Urban Wildlife Through Space and Time

BY

MASON ALLEN FIDINO
B.S., Western Washington University, 2009

THESIS
Submitted as partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biological Sciences in the Graduate College of University of Illinois at Chicago, 2017

Chicago, Illinois

Defense Committee:

Joel S. Brown, Chair and Advisor
Christopher Whelan
Emily S. Minor
Hakan Demirtas, School of Public Health
Seth Magle, Lincoln Park Zoo
This thesis is dedicated to my partner Sophia and our cat Ylla, who both provided support and encouragement in their own unique ways.
ACKNOWLEDGEMENTS

I’ve been fortunate enough to be co-advised by three people, who each deserve my thanks. Seth Magle, I thank you for encouraging me to get my graduate degree, being the best boss I’ve ever had, and teaching me how to think and write like a scientist. Joel Brown, your excitement is contagious, and I thank you for your mentorship and helping me to develop my own ecological worldview. Chris Whelan, I always look forward to our meetings, walks to the Field Museum, or texts about interesting birds during migration. Your insight and knowledge of natural history boosted my often quirky ideas into the realm of ecological relevance.

I thank Hakan Demirtas for taking a chance on an ecologist and mentoring me through computer simulations, Bayesian statistics, and hierarchical modeling. I also thank Emily Minor, whose advice and instruction helped me grow as a scientist. Josie Simonis, your enthusiasm and willingness to mentor me on all things has strongly influenced the way I think about ecology and data analysis. I also thank all of the wonderful folks at the Urban Wildlife Institute (Liza Lehrer, Travis Gallo, Jazmin Rios, Matt Mulligan, and too many interns to count) for their help with data collection, comments on different manuscripts, and their unending support. I could not have done this without you.

Sophia, my partner, I thank you for being there to help pick up the pieces when things have not gone my way over these four years. I also thank my parents who toured my brother, sister, and I through so many different campgrounds as kids. My love for the outdoors had to start with the two of you. Finally, I thank my friends, who helped me take my mind off of things when I needed time to relax, and also texting me photos of animals around Chicago that they wanted me to identify.

MAF
CONTRIBUTIONS OF AUTHORS

Chapter I, my introduction, explains some of my own views towards science, my interest in ecology, and briefly touches on topics of my dissertation. Chapter II is published (Fidino, M., Lehrer, E.W., Magle, S.B.: Habitat dynamics of the Virginia opossum in a highly urban landscape. *Am. Midl. Nat.*, 175;155-167:2016). S. Magle and L. Lehrer helped design the experiment, we all did the field work, and I was responsible for the analysis and wrote the manuscript. Chapter III was done by myself, J. Simonis, and S. Magle. J. Simonis and I developed the statistical framework, I conducted the simulations, subsequent analyses, and was responsible for writing the manuscript. J. Simonis and S. Magle provided guidance on the writing.

Chapter IV is by M. Fidino and S. Magle. I developed the analytical framework to incorporate Fourier series into occupancy models, did the analysis, and wrote the manuscript. Seth looked at previous drafts of this manuscript and supplied the data to illustrate this analytical approach. Chapter V is by M. Fidino, S. Herr, and S. Magle. S. Magle and I developed the idea and S. Herr helped collect the online data. I went through the data, conducted the analysis, and wrote the manuscript with assistance from S. Magle.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>CITED LITERATURE</td>
<td>6</td>
</tr>
<tr>
<td>II. HABITAT DYNAMICS OF THE VIRGINIA OPOSSUM IN A HIGHLY URBAN LANDSCAPE</td>
<td>10</td>
</tr>
<tr>
<td>A. Abstract</td>
<td>10</td>
</tr>
<tr>
<td>B. Background</td>
<td>10</td>
</tr>
<tr>
<td>C. Materials and methods</td>
<td>13</td>
</tr>
<tr>
<td>1. Study area</td>
<td>13</td>
</tr>
<tr>
<td>2. Study design</td>
<td>15</td>
</tr>
<tr>
<td>3. Data collection</td>
<td>17</td>
</tr>
<tr>
<td>4. Predictor variables</td>
<td>18</td>
</tr>
<tr>
<td>5. Statistical analysis</td>
<td>20</td>
</tr>
<tr>
<td>D. Results</td>
<td>21</td>
</tr>
<tr>
<td>E. Discussion</td>
<td>26</td>
</tr>
<tr>
<td>F. Acknowledgements</td>
<td>31</td>
</tr>
<tr>
<td>CITED LITERATURE</td>
<td>31</td>
</tr>
<tr>
<td>III. WHEN IT TAKES MORE THAN TWO TO TANGO: USING MULTI-SPECIES OCCUPANCY MODELS TO ESTIMATE PATTERNS OF CO-OCURRENCE IN COMMUNITY MODULES</td>
<td>37</td>
</tr>
<tr>
<td>A. Abstract</td>
<td>37</td>
</tr>
<tr>
<td>B. Background</td>
<td>38</td>
</tr>
<tr>
<td>C. Model development</td>
<td>40</td>
</tr>
<tr>
<td>1. Incorporating first-order species interactions into occupancy models</td>
<td>40</td>
</tr>
<tr>
<td>2. Extending species interactions to multiple seasons</td>
<td>42</td>
</tr>
<tr>
<td>D. Simulations</td>
<td>44</td>
</tr>
<tr>
<td>1. Simulations conducted</td>
<td>44</td>
</tr>
<tr>
<td>2. Simulation results</td>
<td>48</td>
</tr>
<tr>
<td>E. Analysis of mesocarnivore community dynamics in Chicago, IL</td>
<td>50</td>
</tr>
<tr>
<td>1. Sampling protocol</td>
<td>50</td>
</tr>
<tr>
<td>2. The model</td>
<td>51</td>
</tr>
<tr>
<td>3. Results</td>
<td>52</td>
</tr>
<tr>
<td>F. Discussion</td>
<td>55</td>
</tr>
<tr>
<td>CITED LITERATURE</td>
<td>58</td>
</tr>
</tbody>
</table>
# TABLE OF CONTENTS (continued)

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>IV.</strong> USING FOURIER SERIES TO ESTIMATE PERIODIC TRENDS IN DYNAMIC OCCUPANCY MODELS</td>
<td>64</td>
</tr>
<tr>
<td>A. Abstract</td>
<td>64</td>
</tr>
<tr>
<td>B. Background</td>
<td>65</td>
</tr>
<tr>
<td>C. MODEL DEVELOPMENT</td>
<td>67</td>
</tr>
<tr>
<td>1. Multi-season occupancy models</td>
<td>67</td>
</tr>
<tr>
<td>2. A Brief Foray into Fourier series</td>
<td>70</td>
</tr>
<tr>
<td>3. Single-season pulses</td>
<td>72</td>
</tr>
<tr>
<td>4. Multiple pulses a year</td>
<td>78</td>
</tr>
<tr>
<td>D. Materials and Methods</td>
<td>80</td>
</tr>
<tr>
<td>1. Sampling protocol</td>
<td>80</td>
</tr>
<tr>
<td>2. Models, specification of priors, and model selection</td>
<td>80</td>
</tr>
<tr>
<td>E. Results</td>
<td>84</td>
</tr>
<tr>
<td>1. Colonization</td>
<td>86</td>
</tr>
<tr>
<td>2. Persistence</td>
<td>89</td>
</tr>
<tr>
<td>3. Detection</td>
<td>90</td>
</tr>
<tr>
<td>F. Discussion</td>
<td>93</td>
</tr>
<tr>
<td>CITED LITERATURE</td>
<td>96</td>
</tr>
<tr>
<td><strong>V.</strong> ASSESSING VALUES AND PERCEPTIONS OF WILDLIFE THROUGH SOCIAL MEDIA</td>
<td>103</td>
</tr>
<tr>
<td>A. Abstract</td>
<td>103</td>
</tr>
<tr>
<td>B. Background</td>
<td>103</td>
</tr>
<tr>
<td>C. Materials and Methods</td>
<td>106</td>
</tr>
<tr>
<td>1. Collecting and analyzing comments on social media</td>
<td>106</td>
</tr>
<tr>
<td>2. Categorizing videos and comments</td>
<td>107</td>
</tr>
<tr>
<td>3. Analysis of comments</td>
<td>111</td>
</tr>
<tr>
<td>D. Results</td>
<td>112</td>
</tr>
<tr>
<td>E. Discussion</td>
<td>115</td>
</tr>
<tr>
<td>F. Acknowledgements</td>
<td>119</td>
</tr>
<tr>
<td>CITED LITERATURE</td>
<td>119</td>
</tr>
<tr>
<td><strong>VI.</strong> APPENDICES</td>
<td>126</td>
</tr>
<tr>
<td>APPENDIX A</td>
<td>126</td>
</tr>
<tr>
<td>APPENDIX B</td>
<td>130</td>
</tr>
<tr>
<td>CHAPTER</td>
<td>PAGE</td>
</tr>
<tr>
<td>----------</td>
<td>------</td>
</tr>
<tr>
<td>APPENDIX C</td>
<td>132</td>
</tr>
<tr>
<td>APPENDIX D</td>
<td>151</td>
</tr>
<tr>
<td>APPENDIX E</td>
<td>153</td>
</tr>
<tr>
<td>APPENDIX F</td>
<td>154</td>
</tr>
<tr>
<td>VITA</td>
<td>155</td>
</tr>
</tbody>
</table>
LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. SUMMARY OF VARIABLES USED TO PREDICT OPOSSUM OCCUPANCY, COLONIZATION, AND EXTINCTION IN THE CHICAGO, IL REGION</td>
<td>19</td>
</tr>
<tr>
<td>II. RANKINGS OF OCCUPANCY MODELS FOR OPOSSUM IN CHICAGO, USA BASED ON AKAIKE'S INFORMATION CRITERION CORRECTED FOR SMALL SAMPLE SIZE (AICc). THIS TABLE INCLUDES MODELS WHOSE CUMULATIVE SUMMATION OF MODEL WEIGHTS ARE ≤ 0.95. Δ AICc = AICc FOR A GIVEN MODEL MINUS AICc FOR THE TOP PERFORMING MODEL. K= THE NUMBER OF ESTIMABLE PARAMETERS</td>
<td>23</td>
</tr>
<tr>
<td>III. NOTATION USED IN A DYNAMIC OCCUPANCY MODEL FOR A SINGLE SPECIES.</td>
<td>69</td>
</tr>
<tr>
<td>IV. THE FOUR MODELS FIT TO EACH SPECIES FROM THE CHICAGO, ILLINOIS CAMERA TRAP DATA TO ESTIMATE THE PROBABILITY OF INITIAL OCCUPANCY (Ψ), COLONIZATION (Γ), PERSISTENCE (Φ), AND DETECTION (P). YEAR AND SAMPLING SESSION WERE SPECIFIED AS GROUP-LEVEL PARAMETERS (I.E., RANDOM EFFECTS) THAT ALLOWED THE INTERCEPT TO VARY FOR EACH PROCESS. THE PERIODIC TREND USED WITHIN A MODEL WAS DEPENDENT ON THE SPECIES BEING ANALYZED. A SINGLE-SEASON PULSE WAS USED FOR COYOTE, RED FOX, AND STRIPED SKUNK WHILE A BOOM-BUST SINUSOIDAL TREND WAS USED FOR RACCOON AND OPOSSUM</td>
<td>82</td>
</tr>
<tr>
<td>V. MODEL SELECTION RESULTS OF THE FOUR DIFFERENT MODELS THAT WE FIT TO 9 SEASONS OF CAMERA TRAP DATA FROM CHICAGO, ILLINOIS. MODELS WERE COMPARED WITH THE SUMMARY STATISTIC - Σk,t LOG(CP 0k,t), AND THE LOWEST VALUES INDICATE THE MODEL WITH THE BEST WITHIN-SAMPLE PREDICTIVE PERFORMANCE. THE VALUE OF THE BEST FIT MODEL FOR EACH SPECIES IS BOLDED</td>
<td>85</td>
</tr>
<tr>
<td>VI. PERSISTENCE PARAMETERS FROM THE BEST FIT COYOTE, RED FOX, STRIPED SKUNK, OPOSSUM, AND RACCOON MODELS ESTIMATED FROM 9 SEASONS OF CAMERA TRAP DATA COLLECTED FROM SPRING 2011 TO SUMMER 2013 IN CHICAGO, ILLINOIS. THE TABLE INCLUDES ESTIMATES OF THE INTERCEPT (d0), THE INFLUENCE THAT URB HAS ON EACH SPECIES PERSISTENCE RATES (dURB), AND THE STANDARD DEVIATION OF HOW MUCH d0 VARIANCE OVER TIME STEP (σtime). PARAMETERS ARE ON THE LOGIT-SCALE AND ESTIMATE THE PROBABILITY THAT EACH SPECIES PERSISTS AT A SITE FROM ONE TIME STEP TO THE NEXT. THE 95% CREDIBLE INTERVAL FOR EACH ESTIMATE IS LOCATED WITHIN THE SQUARE BRACKETS NEXT TO THE PARAMETER</td>
<td>90</td>
</tr>
<tr>
<td>VII. DETECTION PARAMETERS FROM THE BEST FIT COYOTE, RED FOX, STRIPED SKUNK, OPOSSUM, AND RACCOON MODELS ESTIMATED FROM 9 SEASONS OF CAMERA TRAP DATA COLLECTED FROM SPRING 2011 TO SUMMER 2013 IN CHICAGO, ILLINOIS. THE TABLE INCLUDES ESTIMATES OF THE INTERCEPT (f0), THE INFLUENCE THAT URB HAS ON EACH SPECIES DETECTION RATES (fURB), AND THE STANDARD DEVIATION OF HOW MUCH f0 VARIANCE OVER TIME (σtime) AND ACROSS SITES (σsite). PARAMETERS ARE ON THE LOGIT-SCALE AND ESTIMATE THE PROBABILITY OF DETECTING A SPECIES PER DAY GIVEN ITS PRESENCE. THE 95% CREDIBLE INTERVAL FOR EACH ESTIMATE IS LOCATED WITHIN THE SQUARE BRACKETS NEXT TO THE PARAMETER</td>
<td>91</td>
</tr>
</tbody>
</table>
TABLE  VIII. CATEGORIES OF THE VIDEO CONTENT IN MOST VIEWED COYOTE, OPOSSUM, AND RACCOON VIDEOS ON YOUTUBE................................................................. 108

TABLE IX. KELLERT CATEGORIES USED IN THIS ANALYSIS TO ASSESS ONLINE ATTITUDES TOWARDS COYOTE, OPOSSUM, AND RACCOON. DEFINITIONS OF KELLERT CATEGORIES WERE TAKEN FROM KELLERT (1980), WHILE COMMENTS WERE TAKEN FROM DIFFERENT VIDEOS THAT WERE COLLECTED FOR THIS ANALYSIS ................................................................. 109

TABLE X. MEAN ABSOLUTE ERROR (MAE), COVERAGE, AND PRECISION OF PERSISTENCE (Φ), COLONIZATION (γ) AND DETECTION PARAMETERS ESTIMATED BY THE CO-OCCURRENCE MODEL FOR THE HORIZONTAL COMPETITION MODULE WHEN SPECIES INFLUENCED EACH OTHER’S PERSISTENCE RATES. SIMULATIONS WERE RUN 200 TIMES WITH VARYING AMOUNTS OF DATA (4, 6, AND 8 SEASONS WORTH). THE MEDIAN ESTIMATE AND 95% QUANTILES ARE LISTED FOR MAE AND PRECISION. COVERAGE WAS CALCULATED AS THE PROPORTION OF SIMULATIONS THAT HAD THE TRUE SIMULATED PARAMETER VALUE WITHIN ITS 95% CREDIBLE INTERVAL ................................................................. 133

TABLE XI. MEAN ABSOLUTE ERROR (MAE), COVERAGE, AND PRECISION OF PERSISTENCE (Φ), COLONIZATION (γ) AND DETECTION PARAMETERS ESTIMATED BY THE CO-OCCURRENCE MODEL FOR THE HORIZONTAL COMPETITION MODULE WHEN SPECIES INFLUENCED EACH OTHER’S COLONIZATION RATES. SIMULATIONS WERE RUN 200 TIMES WITH VARYING AMOUNTS OF DATA (4, 6, AND 8 SEASONS WORTH). THE MEDIAN ESTIMATE AND 95% QUANTILES ARE LISTED FOR MAE AND PRECISION. COVERAGE WAS CALCULATED AS THE PROPORTION OF SIMULATIONS THAT HAD THE TRUE SIMULATED PARAMETER VALUE WITHIN ITS 95% CREDIBLE INTERVAL ................................................................. 138

TABLE XII. MEAN ABSOLUTE ERROR (MAE), COVERAGE, AND PRECISION OF PERSISTENCE (Φ), COLONIZATION (γ) AND DETECTION PARAMETERS ESTIMATED BY THE CO-OCCURRENCE MODEL FOR THE PREDATOR WITH COMPETING PREY MODULE WHEN SPECIES INFLUENCED EACH OTHER’S PERSISTENCE RATES. SIMULATIONS WERE RUN 200 TIMES WITH VARYING AMOUNTS OF DATA (4, 6, AND 8 SEASONS WORTH). THE MEDIAN ESTIMATE AND 95% QUANTILES ARE LISTED FOR MAE AND PRECISION. COVERAGE WAS CALCULATED AS THE PROPORTION OF SIMULATIONS THAT HAD THE TRUE SIMULATED PARAMETER VALUE WITHIN ITS 95% CREDIBLE INTERVAL ................................................................. 143

TABLE XIII. MEAN ABSOLUTE ERROR (MAE), COVERAGE, AND PRECISION OF PERSISTENCE (Φ), COLONIZATION (γ) AND DETECTION PARAMETERS ESTIMATED BY THE CO-OCCURRENCE MODEL FOR THE PREDATOR WITH COMPETING PREY MODULE WHEN SPECIES INFLUENCED EACH OTHER’S PERSISTENCE RATES. SIMULATIONS WERE RUN 200 TIMES WITH VARYING AMOUNTS OF DATA (4, 6, AND 8 SEASONS WORTH). THE MEDIAN ESTIMATE AND 95% QUANTILES ARE LISTED FOR MAE AND PRECISION. COVERAGE WAS CALCULATED AS THE PROPORTION OF SIMULATIONS THAT HAD THE TRUE SIMULATED PARAMETER VALUE WITHIN ITS 95% CREDIBLE INTERVAL ................................................................. 148
# LIST OF TABLES (CONTINUED)

<table>
<thead>
<tr>
<th>TABLE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>XIV.</td>
<td>VIDEOS WHOSE COMMENTS WERE COLLECTED ON YOUTUBE AND USED TO ANALYZE PUBLIC PERCEPTIONS TOWARDS COYOTE, OPOSSUM, AND RACCOON (^1)</td>
</tr>
<tr>
<td>XV.</td>
<td>KEYWORDS ASSOCIATED TO THE SEVEN KELLERT CATEGORIES. ALTHOUGH SOME OF THESE KEYWORDS MAY APPEAR TO BE RELATED TO A SINGLE CATEGORY (E.G., THE KEYWORD 'HUNT' IS LIKELY RELATED TO THE DOMINIONISTIC KELLERT CATEGORY), THE WAY THAT IT IS USED WITHIN A COMMENT COULD DRASTICALLY ALTER HOW IT IS CATEGORIZED. THUS, THESE KEYWORDS ARE ONLY ORGANIZED ALPHABETICALLY BY ROW...</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>My philosopher’s simplex. Each vertex of this triangle has the name of a scholar and a quote from them that aligns with some of my own philosophical views towards science and ecology. .................................................................</td>
</tr>
<tr>
<td>2.</td>
<td>Map of study sites and transects in the Chicago Metropolitan area, Illinois, U.S.A. The dark black lines delineate our three study transects, which follow the Des Plaines River to the northwest, Roosevelt Road to the west, and the Ship and Sanitary Canal to the southwest.................................................................</td>
</tr>
<tr>
<td>3.</td>
<td>Estimated occupancy probability of opossums in the Chicago Metropolitan area, Illinois, U.S.A., across all seasons (n = 10) from spring 2010 to summer 2012 at varying rates of housing density and distances to natural water sources.................................................................</td>
</tr>
<tr>
<td>4.</td>
<td>The A) accuracy, B) precision, and C) coverage of model parameters (1-9) estimated by the co-occurrence model for three-species horizontal competition where species influenced each other’s persistence rates when 4, 6, and 8 seasons of data are supplied (x-axes). Accuracy is measured as the mean absolute error (MAE) of parameter estimates, precision as the measured width of a parameters 95% credible interval, and coverage as the proportion of the 200 models whose posterior distribution contained the true parameter value with its 95% credible interval. The alignment of the model parameters (1-9) reflects the impact matrix with the species-specific intercepts replacing the 0-valued diagonals. For accuracy and precision, the gray dots are median estimates across the 200 simulations for each scenario while the black lines are 95% highest density intervals. Coverage is represented by a single value, and note the scale of the y-axes. ........................................................................................................................................</td>
</tr>
<tr>
<td>5.</td>
<td>Parameter estimates for coyote, opossum, and raccoon colonization (γ) and persistence (Φ) as a function of patch area (m²) and distance from the city center (m) in Chicago, Illinois, USA. Violin plots are posterior densities within the bounds of a parameters 95% credible interval. The vertical dashed line is centered at zero. ........................................................................................................................................</td>
</tr>
<tr>
<td>6.</td>
<td>The estimated influence that the presence of a coyote, opossum, or raccoon have on each other’s probability of persistence from one season to the next throughout Chicago, Illinois. Bolded lines indicate logit-scale parameter estimates whose 95% credible intervals do not bound zero. Species-specific intercept values for coyote, opossum, and raccoon in the absence of other members of the community were respectively 0.24 (-0.20 – 0.72), 0.07 (-0.47 – 0.58), and 0.24 (-0.14 – 0.63). ........................................................................................................................................</td>
</tr>
<tr>
<td>7.</td>
<td>A graphical representation of the periodic pulse function in Eq. 4.12. Here, a pulse occurs with amplitude A once over period P........................................................................................................................................</td>
</tr>
</tbody>
</table>
### LIST OF FIGURES (continued)

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.</td>
<td>Examples of periodic curves when only discrete integers (i.e., t = 1,2,3,4) are supplied for A) a single season pulse with period P = 4 as specified in Eq. 4.17 and B) a boom-bust sinusoidal curve with period P = 2 as specified in Eq. 4.25. Plots A and B are on the logit scale. ( m_0 ) controls the average colonization rate, ( A ) controls the amplitude, and ( \delta ) controls when a pulse occurs.</td>
</tr>
<tr>
<td>9.</td>
<td>Box and whisker plots of the proportion of sites occupied by each species over time estimated by its best fit model from nine seasons of camera trap data collected from spring 2011 to summer 2013 in Chicago, Illinois. The occupancy status of each species at each site was collected by taking the median estimate of the posterior distribution of each species incidence matrix, ( Z ), at each site and time step.</td>
</tr>
<tr>
<td>10.</td>
<td>The probability of colonization over time estimated from the parameters of the best fit model for A) coyote, B) red fox, C) striped skunk, D) raccoon, and E) opossum with 9 seasons of camera trap data collected from spring 2011 to summer 2013 in Chicago, Illinois. The black horizontal line is the median probability of colonization through time, while the gray bars represent the 95% credible interval obtained from the posterior distribution of the species incidence matrix. The black dots are the median estimate of the proportion of sites colonized by each species at each time step from the model, the vertical lines are the 95% credible interval associated to each estimate, and the white dots are the proportion of sites colonized by each species calculated the raw detection data. White dots that are not visible lie directly beneath the model estimate.</td>
</tr>
<tr>
<td>11.</td>
<td>Site-specific detection probabilities estimated by the best fit model for A) coyote, B) red fox, C) striped skunk, D) raccoon, and E) opossum from 9 seasons of camera trap data collected from spring 2011 to summer 2013 in Chicago, Illinois. The black horizontal line is the median estimate of the mean detection rate for each species (( m_0 )), the dark gray bar is the 95% credible interval for ( m_0 ), and the light gray bar is the 95% posterior predictive distribution of site-specific detection rates calculated from each species ( m_0 ) and the site-level standard deviation hyperparameter (( \sigma_{\text{site}} )). Along the x-axis are site-specific detection estimates on the logit scale. The black dots are median estimates for each site while the vertical lines are 95% credible intervals. All sub-figures A-E are sorted by site-specific detection estimates for coyote.</td>
</tr>
<tr>
<td>12.</td>
<td>The estimated proportion of comments that were assigned to Kellert categories from the 10 most viewed coyote (C), opossum (O), and raccoon (R) videos on YouTube. The left panel of the figure illustrates how attitudes differ within a species while the right panels show how species vary across attitudes. The area inside the black tails of each proportion represents the 95% credible interval per species per attitude.</td>
</tr>
<tr>
<td>13.</td>
<td>Distribution of the number of days a camera trap was active from 803 sampling events over 10 seasons between spring 2010 and summer 2012.</td>
</tr>
</tbody>
</table>
LIST OF FIGURES (continued)

14. Comparison of the predicted fits from the bootstrap simulations of 28 (A) and 30 (B) trials from the binomial distribution at daily probabilities calculated from 803 camera trap surveys in Chicago, IL. The black dots are the observed frequency of days a camera trap was active during these surveys, the gray line is the median frequency of days a camera trap was active from the bootstrap simulation, and the dashed lines are 95% confidence intervals................................................................. 129

15. The A) accuracy, B) precision, and C) coverage of model parameters (1-9) estimated by the co-occurrence model for three-species horizontal competition where species influenced each other’s colonization rates when 4, 6, and 8 seasons of data are supplied (x-axes). Accuracy is measured as the mean absolute error (MAE) of parameter estimates, precision as the measured width of a parameters 95% credible interval, and coverage as the proportion of the 200 models whose posterior distribution contained the true parameter value with its 95% credible interval. The alignment of the model parameters (1-9) reflects the impact matrix with the species-specific intercepts replacing the 0-valued diagonals. For accuracy and precision, the gray dots are median estimates across the 200 simulations for each scenario while the black lines are 95% highest density intervals. Coverage is represented by a single value, and note the scale of the y-axes. ................................................................. 137

16. The A) accuracy, B) precision, and C) coverage of model parameters (1-9) estimated by the co-occurrence model for a predator with two competing prey where species influenced each other’s persistence rates when 4, 6, and 8 seasons of data are supplied (x-axes). Accuracy is measured as the mean absolute error (MAE) of parameter estimates, precision as the measured width of a parameters 95% credible interval, and coverage as the proportion of the 200 models whose posterior distribution contained the true parameter value with its 95% credible interval. The alignment of the model parameters (1-9) reflects the impact matrix with the species-specific intercepts replacing the 0-valued diagonals. For accuracy and precision, the gray dots are median estimates across the 200 simulations for each scenario while the black lines are 95% highest density intervals. Coverage is represented by a single value, and note the scale of the y-axes. ................................................................. 142
The A) accuracy, B) precision, and C) coverage of model parameters (1-9) estimated by the co-occurrence model for a predator with two competing prey where species influenced each other’s persistence rates when 4, 6, and 8 seasons of data are supplied (x-axes). Accuracy is measured as the mean absolute error (MAE) of parameter estimates, precision as the measured width of a parameters 95% credible interval, and coverage as the proportion of the 200 models whose posterior distribution contained the true parameter value with its 95% credible interval. The alignment of the model parameters (1-9) reflects the impact matrix with the species-specific intercepts replacing the 0-valued diagonals. For accuracy and precision, the gray dots are median estimates across the 200 simulations for each scenario while the black lines are 95% highest density intervals. Coverage is represented by a single value, and note the scale of the y-axes.
LIST OF ABBREVIATIONS

AIC          Akaike’s information criterion
MAE          Mean absolute error
RMSE         Root mean square error
SUMMARY

Much of my work is focused on developing techniques and methods to get the most ecologically relevant information as possible from observational data. First, I illustrate common approaches used to analyze camera trap data with the analysis of a single species, the Virginia opossum (*Didelphis virginiana*), and show that this species has different habitat requirements throughout Chicago as urbanization increases. My next study shows how to estimate species associations and co-occurrence rates with camera trap data, which I then validate with extensive use of simulations. Following this, I apply the model I developed to estimate rates of co-occurrence between coyote (*Canis latrans*), raccoon (*Procyon lotor*), and the Virginia opossum. While I predicted coyote would negatively influence the two smaller mesocarnivores, I found no evidence of this relationship in the data. In my next chapter, I develop an approach to estimate periodic trends in the spatiotemporal distribution of species by incorporating Fourier series into dynamic occupancy models. Overall, this approach accounts for between 30-73% of the temporal variability in the colonization rates of the species I analyzed. This approach also outperforms other more commonly used approaches that estimate temporal dynamics. In my final chapter, I develop an approach to quantify values and perceptions towards wildlife through comments made on social media. Collectively, I see my dissertation as setting a foundation for future empirical research through the creation of generalizable and robust statistical methods that can be used to answer both basic and applied problems in ecology.
I. INTRODUCTION

If the field of ecology were a building, it would have rooms as diverse as all of the species we have, have yet to, and may never find. The floors of the main hall would be constructed with timber from the Beagle and a statue of Grinnell would occupy a niche near the hall of community ecology, a California thrasher standing proudly on his shoulder (Grinnell, 1917). Looking upwards, you would see Robert Paine’s initials with the depiction of a starfish on the keystone of archways (Paine, 1969). To a keen eye, the distinctive style of many different ecological architects is visible. As the natural world does little to divide itself along the demarcations of each architect’s ecological pursuits and specializations, it has taken many adept and collaborative hands to add form and function to this structure. Given its expansiveness, some may only explore and build within a single wing in their lifetime. Others, seeking generality across disciplines, deftly construct arterial hallways between them. As I wrap up my own Ph.D., I’m excited to walk into this ecological ‘House on the Rock’, knowing that I’ve been taught how to make my own small contributions to it. Most of all, I am excited to see how it will expand and change over my lifetime.

My own research interests primarily lie in understanding why various kinds and numbers of species live in different environments. While it’s a seemingly simple question to ask, the answer tends to be complicated and often begins with ‘well, it depends...’. There are a significant number of processes that determine where species may be, and who they may coexist with. These processes arise from the interaction between one or more axes of environmental heterogeneity, evolutionary trade-offs among organisms, and the supply point
(sensu Tilman, 1981) of a given locality (MacArthur 1958, 1972; Kotler and Brown, 1988; Chase and Leibold, 2003; but see Hubbel 2001). Common trade-offs at the local-scale include differential resource use (MacArthur, 1972; Tilman, 1982), abiotic tolerances (Tilman and Pacala, 1993), stress or disturbance tolerances (Paine and Levin, 1981), or sensitivity to predators (Leibold, 1996). The fact that trade-offs occur is true across environments, from the most natural and far off reaches of planet Earth to my own back yard in Chicago.

If, however, coexistence is not possible locally it may still occur within a given region. For example, environmental heterogeneity shifts supply points across the landscape and may ensure that species have at least one locality to exist (Chesson, 2000). Spatial variability in the abundance of a resource may promote the coexistence of ‘cream skimmer’ and ‘crumb picker’ species (Brown, 2000). Conversely, organisms can simply differ in how they find and exploit habitat patches (Levins and Culver, 1971; Horn and MacArthur, 1972; Tilman, 1994; Calcagno et al., 2006). Therefore, it is important to realize that trade-offs at multiple scales influence local and regional patterns of diversity (Holyoak et al., 2005). To me, this is why community ecology is so interesting.

Similar to the species we study, ecologists face their own set of trade-offs when trying to understand aspects of the world around them. These trade-offs are varied, and most arise from our own research interests and the fact that we will never have the time to know everything we want to. Other trade-offs are methodological and philosophical. For instance, robust causal inference is attainable through controlled experiments while verisimilitude is achieved through observing the natural world. For the former, we randomize, block, and replicate to quell potential confounders (Fisher, 1925). These techniques have been profitably
used to identify a given treatment’s effect and have profoundly influenced the way we answer questions in the sciences. However, many ecological questions and aspects of the natural world do not willingly surrender themselves to randomization, blocking, or replication. We cannot control when it rains nor can we apply and replicate treatments at the regional scale. Instead, we could conduct natural experiments along a gradient of precipitation, controlling for what we can, but inference is not as strong. Yet, such experiments and observations are necessary if we are to understand how real populations in the real world work (Hargrove and Pickering, 1992; Pickett et al., 2010).

All of the research in my dissertation is on the observational side of the research spectrum. As previously stated, inference is weaker, but patterns found via observation arguably imply repeatability and therefore predictability (MacArthur, 1972). Further, observational work is necessary to ethically answer a variety of multicausal and broad-scale questions and therefore is a critical component towards ecological understanding (Pickett et al., 2010). It would be ill-advised, for instance, to intentionally transplant a species beyond its range to determine why such boundaries exist, as such an experiment may result in the inadvertent range expansion of the species if it escapes. We would also likely be hard-pressed today to have an animal care and use committee okay the fumigation of islands (Simberloff and Wilson, 1969). At the same time, technological advancements have also made it easier to collect data at spatial and temporal scales that were previously impossible, and with this has come a resurgence in observational approaches and studies in ecology (Sagarin and Pauchard, 2010; Steenweg et al., 2016).
Be it observational or experimental, ecologists collect vast amounts of data. As a result, we face many choices when it comes time to analyze that information. For controlled experiments, the assumptions of more traditional statistical tests are often fulfilled and subsequent analyses are simplified. Conversely, observational data can be muddy, and the ecological process we are interested in may be hidden by observational noise. As most of my work is related to making better inference with observational data, my personal philosophy towards science reflects how I think this should be done (Fig. 1).

*All models are wrong, but some are useful.*

**Figure 1.** My philosopher’s simplex. Each vertex of this triangle has the name of a scholar and a quote from them that aligns with some of my own philosophical views towards science and ecology.
First, it is my opinion that ecology is a field of probabilities and proportions and that a lot can happen between zero and one. Environments range from homogeneous to heterogeneous, and therefore the relative fitness of different phenotypes in a population differ as a result (Levins, 1968). Species colonize biogeographical islands. Others go extinct (MacArthur and Wilson, 1967). An organism must decide when to stop foraging in a food patch and leaves some proportion of food behind (Stephens et al., 2007). To better understand such problems, I have developed a passion for probability theory, statistical distributions, and modeling. To me, it easier to formulate and answer questions in ecology with such skills. In particular, this can be seen in chapters 3, 4, and 5 of my dissertation.

Second, I feel that we often know more about the natural history of a species than what we include in some statistical or theoretical model. Natural history texts, for example, can tell you when species are likely to breed and when their young will disperse (e.g., Feldhamer et al., 2003). Yet, it is necessary to couch such knowledge in mathematical terms so that it can be explored theoretically (e.g., Skellam, 1967) or estimated from data (Chapter 1 and 4). Doing so will arguably lead to more robust and sound inference. To echo Hali and Järvinen (1982) and add to the quote by John Maynard Smith (Fig. 1): “Sound naturalism is to ecology what legs are to a runner; but antitheoretical naturalists are, quite naturally, like headless runners.” While the field of ecology is becoming more quantitative, we should always do our best to not forget our rooting in natural history.

Finally, the models that we develop, be they graphical, mathematical, or statistical, are abstractions of the truth and function as a means to extract useful information from data (Fig. 1). Maps, for instance, are a model of the geographic features of a region, and good maps can
be very useful. Ecologists are mapmakers of the natural world, and it is our job to simplify and approximate what we can so that we may better conserve and understand the species around us. All of my research reflects this notion.

In conclusion, I hope that this collection of work illustrates my own growth as a scientist and the development of my own ecological world view. We are at a point in ecology where we have the opportunity to answer problems that were previously impossible, and one feature that may limit our ability for research is that we must quantify uncertainty related to the processes we are interested in and the data that is collected. It is my hope that I can play a part in constructing the rooms that help us surpass those limits.

**CITED LITERATURE**


Steenweg, R., Hebblewhite, M., Kays, R., Ahumada, J., Fisher, J.T., Burton, C., Townsend, S.E.,


II. HABITAT DYNAMICS OF THE VIRGINIA OPOSSUM IN A HIGHLY URBAN LANDSCAPE

This research was published in the following article:


A. Abstract

As urban habitats vary in composition and structure along the urban to rural gradient, different degrees of urbanization likely result in a diversity of landscape responses from wildlife. We investigated this relationship with the Virginia opossum (*Didelphis virginiana*), an urban adapted species that is both common and understudied in highly metropolitan landscapes. We investigated which landscape factors affect opossum occupancy, colonization, extinction, and detection by using a large system of motion-triggered camera traps in the Chicago metropolitan area over 10 seasons from spring 2010 to summer 2012. Opossum patch occupancy rates were highest near natural water sources regardless of urbanization, whereas occupancy rates in patches ≥ 1000 m from natural water sources decreased with increasing urbanization. Our results suggest opossums have relaxed habitat needs at intermediate levels of disturbance, as the ability to locate anthropogenic water sources may allow them to occupy previously uninhabitable patches.

B. Background

Though one of the ecological trademarks of cities is decreased species diversity (Aronson et al., 2014), wildlife still exist within these highly urban landscapes and little is known about the
habitat dynamics that drive their distribution. As the world steadily becomes more urban, and more people occupy cities (McDonnell and Hahs, 2008), it is necessary to better understand the habitat needs of species within highly urban landscapes to conserve biodiversity, increase the quality of wildlife management practices, and improve how cities are built. While urban habitats can negatively impact wildlife in many ways (e.g., human-induced disturbance regimes, spatial heterogeneity, novel competitors, etc.), species that adapt to the challenges associated with urban systems often benefit indirectly from these habitats. Of the benefits associated with urban environments, subsidized food and water sources arguably provide the largest influence in that they can increase a species' density (Luniak, 2004; Prange and Gehrt, 2004), individual survival rates (Kanda et al., 2009), and relax selective pressures on natural habitats (Bozek et al., 2007). Yet, as urbanization increases and the landscape is dominated by impervious surfaces, humans, and traffic, energy gained from these subsidized resources may not overcome the energetic cost of survival. For example, our largest cities likely contain a wealth of anthropogenic food for wildlife; however, the ability to successfully acquire these resources may be more difficult as available habitat becomes more fragmented and the human population increases (Magle et al., 2014). Additionally, the lower proportion of green space within cities may decrease the availability of water sources, anthropogenic or otherwise, thereby limiting species to areas where they can locate water. In response, species may make different habitat choices to lessen the impact of increased human presence and habitat fragmentation within highly urban landscapes.

The Virginia opossum (*Didelphis virginiana*, hereafter opossum) is an excellent candidate to study this potential trend. North America’s only marsupial, the opossum is predominantly
nocturnal, roughly the size of a domestic house cat, and largely opportunistic in diet (Gardner and Sunquist, 2003). An urban adapted species (*sensu*, McKinney, 2002), the opossums’ ability to access water sources may be a crucial habitat requirement throughout its distributional range, and many rural studies observe surface water to be a key habitat requirement (Lay, 1942; Reynolds, 1945; Llewellyn and Dale, 1964; Gardner and Sunquist, 2003; Babb et al., 2004). Urban studies of the opossum, however, have not found proximity to surface water to be necessary (Kanda et al., 2006; Markovchick-Nicholls et al., 2008), which may be due to the prevalence of anthropogenic water sources such as irrigation canals, drainage ditches, and bird baths. Therefore, the opossum may have different habitat requirements in urban and rural areas.

Despite their prevalence along the entirety of the urban to rural gradient, opossums have never been studied within highly urban landscapes. To date, urban opossums have only been studied in small towns and suburbs (Meier, 1983; Kanda et al., 2005; Bozek et al., 2007), forest preserves near cities (Prange and Gehrt, 2004; Markovchick-Nicholls et al., 2008), inside of zoos (Harmon et al., 2005), and in fragments of remnant habitat within the steep-sloped canyons of the San Diego Metropolitan region (Crooks, 2002). Due to the opossum’s large distributional range and ubiquity in urban habitats there is a significant knowledge gap regarding the habitat needs of this urban adapted species at high degrees of urbanization. As such, opossums represent a significant opportunity to better understand the ecology of our cities. Additionally, the opossum is an excellent candidate to explore shifting habitat needs in highly urban areas because opossum patch colonization and extinction events can be observed over a short time frame. At the northern portion of their range, opossums may breed up to
twice a year but rarely survive more than one winter (Prange and Gehrt, 2004; Kanda et al., 2009). Moreover, opossums are able to rapidly expand their range and occupy new areas because pregnant females and juvenile opossum often disperse great distances (> 1.5 km) to new habitats in a single night (Gillette, 1980), though dispersal distances may be less within cities due to their highly fragmented nature.

To determine if varying amounts of urbanization affect the habitat dynamics of opossums we set motion-triggered camera traps in city parks, cemeteries, forest preserves, and golf courses along three 50 km transects that originated in downtown Chicago, Illinois, U.S.A. and radiated outward along an urbanization gradient. We used a single species, multiple season occupancy modeling framework (MacKenzie et al., 2006) to predict the distribution of opossum and determine the key landscape, local, and temporal factors that influence their habitat dynamics in a highly urban landscape. Two nonmutually exclusive hypotheses were considered for this study: (i) as opossums have been shown to have a need for green space in less urban environments (Meier, 1983), opossum patch dynamics should be positively related to factors associated to green space and negatively related to factors associated to the built environment and (ii) if anthropogenic water sources are either more difficult to access or less abundant with increasing urbanization, then opossum patch occupancy rates should decrease far from water as urbanization increases, whereas opossum patch occupancy rates close to natural water sources should stay high regardless of the degree of urbanization.

C. **Materials and methods**

1. **Study area**
This study was part of an ongoing camera trap survey to determine the distribution of medium to large mammals throughout the Chicago Metropolitan area (41° 50’ 15’’ N, 87°40’55’’W). Located on the southwestern shore of Lake Michigan, Chicago experiences four distinct seasons that consists of warm summers (average low = 16.8 C), cold winters (average low = 3.6 C), and intermediate springs (average low = 8.1 C) and autumns (average low = 4.7 C; WolframAlpha, 2014). Chicago is the most populous midwestern city in the United States with ≥ 2.7 million residents and an average population density of 7355 people km\(^{-2}\), whereas the greater metropolitan area surrounding the city contains an estimated 9.5 million people (U.S. Census, 2013a).

Four different site types were included in this study, which together represent the majority of potential habitats throughout the Chicago metropolitan area: forest preserves, city parks, cemeteries, and golf courses. We solely focused on these types of habitat as they make up the largest proportion of green space in the metropolitan area. Forest preserves had varying degrees of native and/or natural plant vegetation with some sites being heavily landscaped, highly invaded by invasive plant species (e.g., common buckthorn \((Rhamnus cathartica)\)), or both (Vernon et al., 2014). Small city parks (< 2 ha) consisted of mature trees within a landscape of turf grass, whereas large city parks (> 2 ha) had areas of natural vegetation interspersed throughout the park or around the shoreline of man-made ponds. Cemeteries differed greatly in habitat characteristics with some being large (> 2 ha) but functionally similar to small (< 2 ha) city parks and others small but directly abutting or containing wooded areas. Golf courses were large (mean 82.7 ± 34.5 ha), heavily landscaped and contained man-made water features.
However, most golf courses retained some natural areas within and around their expansive property.

2. **Study design**

A gradient based design was used to determine how varying amounts of urbanization affected the habitat dynamics of opossums. This method has become an increasingly common and useful approach to assess how urbanization alters ecological processes (McDonnell and Hahs, 2008). We delineated three 50 km transects that originated from Union Station in downtown Chicago and radiated outward in three different directions: north following the Des Plaines River, west along Roosevelt Road, and south following the Ship and Sanitary Canal (Fig. 2). Transects included a variety of landscape types such as urban, suburban, exurban, open space, grassland, and forest. Each transect was equally split into ten 5 km sections and a maximum of four sites were sampled within each section. Sites were selected within 2 km of each transect (dependent upon access) with a minimum distance between sites of 1 km. Because the three transects originated from a central location, only four sites were sampled within the first section of all transects combined, but otherwise attempts were made to create an equal distribution of sites across all portions of the gradient, with available sites otherwise randomly chosen (Fig. 2). A total of 118 sites were sampled, and sites represented both public and private ownership.
Figure 2. Map of study sites and transects in the Chicago Metropolitan area, Illinois, U.S.A. The dark black lines delineate our three study transects, which follow the Des Plaines River to the northwest, Roosevelt Road to the west, and the Ship and Sanitary Canal to the southwest.
3. **Data collection**

One Bushnell motion-triggered infrared Trail Camera (Model #119436C) was placed at each site for approximately 4 weeks per season (i.e., spring, summer, autumn, and winter) from spring 2010 to summer 2012 (10 seasons total, details in Vernon et al., 2014). Each season, in addition to the camera trap, one plaster disk impregnated with synthetic fatty acid scent and two randomly selected carnivore-attracting lures were placed at each site within view of the camera trap in order to increase the probability of detecting species (Magle et al., 2015). While lures may draw individuals in from a somewhat larger area, which could alter small scale habitat choices made by species within a patch, this study is solely focused on habitat dynamics at the home range scale and differences in occupancy, colonization, and extinction between patches. Therefore, lure placement should not bias our analysis. Sampling season was considered the primary sampling period, whereas weeks within a season constituted secondary sampling occasions. Because all cameras were not placed on the same date, the start of the first sampling week was defined as the average starting camera date per season and each successive sampling week began 7 days after the previous one. Therefore, each site produced four data points per season and 16 data points per year. Each sampling week was either coded as “1” if an opossum was detected at least once within that time frame, “0” if they were not, or “.” (censored) if the camera was not present or not operable. A site was seasonally censored from the analysis if the camera was not functional for at least 18 days. A subset of sites \((n = 25, 22\%)\) were retired during the study due to repeated vandalism, theft of a camera trap, or change in site accessibility. Data from functional seasons at a site prior to being retired were included in the analysis, whereas seasons following this designation were censored.
4. **Predictor variables**

To identify landscape factors influence opossum patch dynamics, ArcGIS ver.9.3 (ESRI, 2008) was used to create layers of landscape factors within a 500 m buffer centered on the location of the camera trap (Table I). A 500 m buffer (78.5 ha) was chosen based on previous research that found opossum home range sizes varied from 51 ha to 108 ha (Gillette, 1980). All layers used were in raster format, and had 30 m resolution. Two landscape factors were selected to represent both the human and habitat impacts of urbanization: average housing density (derived from U.S. Census, 2010) and average amount of impervious surface (Illinois Geospatial Clearinghouse, 2003). We predicted both factors would be negatively associated to opossum occupancy and colonization but positively associated to extinction because areas with high housing densities have been shown to deter wildlife (Magle et al., 2014) and areas with large tracts of impervious surface may have fewer anthropogenic water sources due to stormwater management practices (i.e., gray infrastructure) as well as higher chances of road mortality.
### TABLE I
SUMMARY OF VARIABLES USED TO PREDICT OPOSSUM OCCUPANCY, COLONIZATION, AND EXTINCTION IN THE CHICAGO, IL REGION

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Housing Density (Units)</td>
<td>776.61</td>
<td>0</td>
<td>4950.81</td>
<td>909.85</td>
</tr>
<tr>
<td>Impervious Surface (%)</td>
<td>38.33</td>
<td>0</td>
<td>85.59</td>
<td>22.52</td>
</tr>
<tr>
<td>Canopy Cover (%)</td>
<td>17.85</td>
<td>0</td>
<td>75.88</td>
<td>19.63</td>
</tr>
<tr>
<td>Distance to Water (m)</td>
<td>1008.95</td>
<td>0</td>
<td>5490.25</td>
<td>1179.44</td>
</tr>
<tr>
<td>Total Available Landscaped Habitat (30 x 30 m units)</td>
<td>168.75</td>
<td>0</td>
<td>862</td>
<td>192.16</td>
</tr>
</tbody>
</table>

Additionally, three factors were selected to represent the natural environment: percent canopy cover of trees (Fry et al., 2011), distance to nearest natural water source (Illinois GAP data), and total available landscaped habitat, which represented the summation of the following land use categories within 500 m of a camera trap: city parks, cemeteries, and golf courses. As opossums have broad habitat tolerances, but a preference for deciduous forest close to water (Gardner and Sunquist, 2003), we predicted these factors would increase patch quality and occupancy. Furthermore, interactions between distance to natural water sources and our two covariates related to urbanization (i.e., housing density and impervious surface) were considered to test if opossums have a stronger need for sites closer to water as urbanization increases. Seasonal variation in detection probability also was investigated because opossums alter their movement patterns seasonally (Gillette, 1980; Kanda et al., 2009), which may increase or decrease the likelihood of encountering a camera trap. Finally,
the impact of canopy cover on detection probability was assessed as it is correlated to the density of nearby vegetation (Fry et al., 2011).

5. **Statistical analysis**

We used single-species, multiple-season occupancy models with the RMark package ver 2.1.8 (Laake, 2013) in Program R ver 3.1.1 (R Core Team, 2014) to determine which landscape and local variables best explained our data with respect to initial occupancy (Ψ), colonization (γ), extinction (ε), and detection (ρ), and then estimated future occupancy rates to determine habitat suitability using the recursive equation \( \Psi_{t+1} = \Psi_t (1 - \epsilon_t) + (1 - \Psi_t) \gamma_t \) (MacKenzie et al., 2003, 2006).

For this analysis, we constructed alternate parameterizations of Ψ, γ, ε, and ρ using our predictor variables, then tested all combinations of each parameterization (including the null model, \( n = 336 \)). To test for multicollinearity, we calculated Pearson’s correlation coefficients between all covariates used in this analysis and did not use any covariates in the same model if \( r \geq 0.60 \). Two sets of covariates were highly correlated: Housing density and impervious surface (\( r = 0.65 \)) and canopy cover and impervious surface (\( r = -0.81 \)), therefore we omitted models that included combinations of these covariates.

As the initial distribution of opossum represents a snapshot of patch dynamics rather than the processes that occurred throughout this survey (MacKenzie et al., 2006), we only considered four parameterizations for Ψ, which included factors associated with natural habitat as singular covariates (i.e., distance to water, total available landscaped habitat, and canopy cover) as well as the null model. We input identical parameterizations for γ and ε within a
model, which resulted in 21 different combinations. To test our first hypothesis, we created additive combinations of our five landscape variables, constrained to a maximum of three covariates in order to decrease over-parameterization (Magle et al., 2010). To test our second hypothesis, we included models with an urbanization covariate (i.e., housing density or impervious surface), distance to water, and an interaction term between the two. Finally, we considered four parameterizations for \( \rho \), which included additive combinations of season and canopy cover as well as the null model.

We used Akaike’s information criterion, adjusted for small sample size (\( \text{AIC}_c \)) to rank models, and models within 2 \( \Delta \text{AIC}_c \) units of the top-ranked model were considered to have substantial support (Burnham and Anderson, 2002). We did not test for overdispersion as currently, there is no adequate method for testing this in multiple-season modeling with covariates (MacKenzie et al., 2006). Before analysis, all variables were z-transformed to ensure comparability.

D. Results

From spring 2010 to summer 2012, a total of 4546 photos of opossum were taken at the 118 sites. Autumn (48.50%) had the greatest number of opossum photos, followed by spring (21.20%), summer (18.30%), and winter (12.00%). Furthermore, opossum exhibited a strong nocturnal trend throughout every season, and almost all (97.30 %) photos were taken between sundown and sunrise in all seasons. Opossums were photographed at 75 of the 118 sites and average naïve occupancy per season was 23.87 ± 14.04%, which increased to 41.41 ± 11.94% after accounting for imperfect detection. A total of 4720 camera trapping data points were


possible (118 sites * 10 seasons * 4 w per season), but not all sites were sampled throughout
the study, and some cameras were stolen or vandalized. Ultimately, 1951 site-weeks of data
were collected for this study. On average, camera traps were functional for 27.10 ± 8.41 nights
per site per season (range = 1 – 125).

Only one model within the model set was considered to have substantial support
(ΔAICc < 2; Table II). Hence, we did not conduct model averaging to address model uncertainty
and made all of our inferences from this top model (Burnham and Anderson, 2002). In this
model, distance to natural water sources was negatively related to initial occupancy (βwater = -
2.14, SE = 1.09), while housing density, distance to water, and the interaction term between the
two influenced colonization and extinction. Housing density was negatively related to
colonization (βhouse = -1.36, SE = 0.29) and extinction (βhouse = -1.67, SE= 0.49), whereas distance
to water was negatively related to colonization (βdist2water = -0.05, SE = 0.23) but positively
related to extinction (βdist2water = 0.51, SE = 0.26). The interaction term between housing density
and distance to water was positively related to colonization (βdist2water*house = 0.34, SE = 0.16) and
extinction (βdist2water*house = 0.40, SE = 0.19). Season was the only term related to detection
probability. Relative to winter, autumn had the highest detection probability (βautumn= 0.54, SE =
0.27), followed by spring (βspring= 0.01, SE = 0.26), and summer (βsummer = -0.25, SE = 0.25).
<table>
<thead>
<tr>
<th>Model</th>
<th>K</th>
<th>( \Delta AIC_c )</th>
<th>Weight</th>
<th>Deviance</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \Psi(\sim \text{water})\varepsilon(\sim \text{water + house + water} \ast \text{house})\gamma(\sim \text{water + house + water} \ast \text{house})\rho(\sim \text{season}) )</td>
<td>14</td>
<td>0.00</td>
<td>0.42</td>
<td>1395.55</td>
</tr>
<tr>
<td>( \Psi(\sim \text{water})\varepsilon(\sim \text{water + house + water} \ast \text{house})\gamma(\sim \text{water + house + water} \ast \text{house})\rho(\sim \text{canopy + season}) )</td>
<td>15</td>
<td>2.06</td>
<td>0.15</td>
<td>1395.52</td>
</tr>
<tr>
<td>( \Psi(\sim \text{water})\varepsilon(\sim \text{water + house})\gamma(\sim \text{water + house})\rho(\sim \text{season}) )</td>
<td>12</td>
<td>2.66</td>
<td>0.11</td>
<td>1402.36</td>
</tr>
<tr>
<td>( \Psi(\sim \text{water})\varepsilon(\sim \text{house})\gamma(\sim \text{house})\rho(\sim \text{season}) )</td>
<td>10</td>
<td>3.20</td>
<td>0.09</td>
<td>1407.05</td>
</tr>
<tr>
<td>( \Psi(\sim 1)\varepsilon(\sim \text{water + house + water} \ast \text{house})\gamma(\sim \text{water + house + water} \ast \text{house})\rho(\sim \text{season}) )</td>
<td>13</td>
<td>3.74</td>
<td>0.07</td>
<td>1401.37</td>
</tr>
<tr>
<td>( \Psi(\sim \text{water})\varepsilon(\sim \text{water + house})\gamma(\sim \text{water + house})\rho(\sim \text{canopy + season}) )</td>
<td>13</td>
<td>4.71</td>
<td>0.04</td>
<td>1402.33</td>
</tr>
<tr>
<td>( \Psi(\sim \text{water})\varepsilon(\sim \text{house})\gamma(\sim \text{house})\rho(\sim \text{canopy + season}) )</td>
<td>11</td>
<td>5.23</td>
<td>0.03</td>
<td>1407.00</td>
</tr>
<tr>
<td>( \Psi(\sim 1)\varepsilon(\sim \text{water + house})\gamma(\sim \text{water + house})\rho(\sim \text{season}) )</td>
<td>11</td>
<td>5.63</td>
<td>0.03</td>
<td>1407.40</td>
</tr>
<tr>
<td>( \Psi(\sim 1)\varepsilon(\sim \text{water + house + water} \ast \text{house})\gamma(\sim \text{water + house + water} \ast \text{house})\rho(\sim \text{canopy + season}) )</td>
<td>14</td>
<td>5.79</td>
<td>0.02</td>
<td>1401.33</td>
</tr>
</tbody>
</table>
With this top model, we quantitatively estimated occupancy rates in the remaining sampling sessions (seasons 2 – 10) using the recursive equation at varying intensities of housing density (0 – 4000 units in intervals of 1000) and distances to natural water sources (0 – 4000 m in intervals of 1000). Then, we averaged these occupancy rates across all estimated seasons to determine how opossum habitat dynamics vary along a gradient of urbanization (Fig. 3). Across all seasons, probability of occupancy near natural water sources averaged ~50% and stayed relatively constant regardless of housing density, although there was a slight increase to 60% when housing density was set to 1000 units (Fig. 3). Additionally, when housing density increased from 0 units to 4000 units occupancy probability in patches further from natural water sources (≥ 1,000 m) decreased an average of 28.15 ± 8.59%. At low housing densities (1,000 units), occupancy rates far from natural water sources (≥ 1000 m) remained relatively high (mean = 33.24±7.28%), while occupancy rates at high housing densities (4000 units) were much lower (mean = 9.90±5.31%).
Figure 3. Estimated occupancy probability of opossums in the Chicago Metropolitan area, Illinois, U.S.A., across all seasons (n = 10) from spring 2010 to summer 2012 at varying rates of housing density and distances to natural water source.
E. Discussion

Opossum patch occupancy rates stayed relatively constant and high near natural water sources regardless of urbanization but greatly decreased in patches ≥ 1000 m from natural water sources. Therefore, we conclude opossum have a strong need for natural water sources in highly urban landscapes. This trend was not found at lower levels of urbanization and likely indicates that characteristics associated with these less urban environments such as drainage ditches and larger yards uncouple this habitat requirement. Additionally, opossum choice of patches close to natural water sources only appears to be relaxed at intermediate levels of disturbance as opossum also require habitat patches close to natural water sources in rural environments (Lay, 1942; Reynolds, 1945; Llewellyn and Dale, 1964; Gardner and Sunquist, 2003). We suggest the most likely explanation for this apparent trend is the abundance of and ease of access to anthropogenic water sources at intermediate levels of urbanization, as evidenced by opossum exhibiting higher occupancy rates further from natural water sources at low housing densities (Fig. 3). Furthermore, this observed relationship accounts for the results of Kanda et al. (2006), who did not find proximity to water to be a key habitat feature of road-killed opossum around Amherst, Massachusetts, a city whose average housing density is an order of magnitude less than Chicago (Amherst = 134.90 houses km\(^{-2}\), Chicago = 1970.50 houses km\(^{-2}\); U.S. Census, 2013b).

Though we could not quantify the abundance of anthropogenic water sources at our study sites, we believe housing density was a sufficient proxy. Irrigation, for example, accounts for a large proportion of external water usage in cities (Syme et al., 2004) that would not be as available in areas with increased housing density because the proportion of green space and
gardens decreases as housing density increases (Tratalos et al., 2007; Fuller and Gaston, 2009). Additionally, the greater density of impervious surfaces in cities shifts stormwater infrastructure from green infrastructure (e.g., drainage ditches) to elaborate sewer systems (i.e., gray infrastructure). As a result, available habitat for the opossum in highly urban landscapes may be constrained to where they can locate water, which appears to be near natural water sources.

While proximity to water strongly predicts the distribution of opossum throughout Chicago, other elements associated with water sources may be responsible for this pattern. Because opossums tend to forage near den sites (Gilette, 1980), having reliable food sources nearby is likely crucial to their survival. As riparian corridors are high in biodiversity and have greater food availability (Naiman et al., 1993), opossum may need such locations to access reliable food sources. Furthermore, opossums may select sites close to natural water sources because these patches are more connected within the city. Just as road networks move humans from place to place, Chicago’s river system, which travels almost to the center of downtown, may function as a crucial corridor for dispersal. As the majority of the world’s largest cities are built on fertile habitats near rivers or oceans, we hypothesize that natural water sources within urban landscapes may function as important habitat for wildlife and allow some species to persist within highly urban environments. In fragmented landscapes riparian corridors have a greater diversity of mammalian predators (Hilty and Merenlender, 2004) and generally are high in biodiversity (Naiman et al., 1993). Though highly urban landscapes are generally low in species diversity (Marzluff, 2005), increasing suitable habitat near rivers and other water
sources may be a useful method for cities to increase biodiversity, facilitate travel through urban cores, and increase city residents’ urban experience.

In addition to the impact of water, opossums were also more likely to occupy patches with lower housing densities, even at distances to water ≥ 1000 m (Fig. 3). This result contradicts previous research on this species (Kanda et al., 2009), but likely is due to the highly metropolitan nature of Chicago. Highly urban landscapes represent a heterogeneous mosaic of habitable (lower housing density) and uninhabitable (higher housing density) patches, regardless of proximity to the urban core. Indeed, opossums were observed throughout the entirety of the gradient (Fig. 2), and housing density at sites within the city of Chicago that had opossums at least once ($n = 14$) averaged $1097 \pm 953$ units, whereas sites that never had opossums present ($n = 22$) averaged $1863 \pm 1,026$ units. A post-hoc analysis of this difference produced significant results; ($t(35) = -2.48, P = 0.02$) which does provide some additional evidence that opossums do indeed select for sites with lower housing densities within the city, but our model suggests the negative aspects associated to areas with higher housing densities may be offset by proximity to natural water sources.

Despite being highly correlated with housing density, degree of impervious surface performed poorly in this analysis as a predictor of opossum colonization and extinction. One possible explanation is that urban mesocarnivores respond more to the presence of humans within cities than the patchy distribution of the natural habitat within the landscape. Many urban adapted mesocarnivores, for instance, alter their temporal activity to avoid people (Ditchkoff et al., 2006). As cities are highly complex and controlled by natural and social factors, wildlife may consider more than just the bricks and mortar of an urban landscape and select
locations where they can maximize the benefits associated with urban life (i.e., subsidized resources) while minimizing the costs. Hence, although impervious surfaces significantly alter ecological processes (Shochat et al., 2006), factors that account for both the landscape and human aspects of the environment, like housing density, may play a more important part in determining the distribution of opossums and potentially other urban adapted mammals, but more research is needed to confirm this generalization.

Furthermore, total available landscaped habitat and canopy cover had less support than distance to water in our analysis. Opossums have been shown to prefer green space in urban areas (Meier, 1983) and were present in all landscaped habitats (i.e., golf courses, city parks, and cemeteries) throughout the study. Yet, total available landscaped habitat may be too imprecise a habitat descriptor because it assumes that all landscaped habitat is equal and that opossum do not select for any local characteristics within a patch. Within cities, the built environment increases habitat heterogeneity between patches but decreases habitat heterogeneity within the patch itself (Band et al., 2005). As such, habitat quality within a patch may have a greater impact on opossum habitat dynamics than the proportion of green space nearby a patch. Additionally, percent canopy cover may represent different habitat types as urbanization increases. Within the city, the majority of landscape below trees is often turf grass that is kept in a constant state of disturbance via landscaping and is less likely to offer opportunities for den sites. Conversely, the landscape below trees in less urban parts of the study area (e.g., forest preserves) is densely vegetated, less likely to be managed, therefore has a higher probability of containing denning locations. As proportion of canopy cover does not
adequately capture this variation, collecting site specific vegetation data around sites in urban studies may be necessary.

Though many studies have shown that habitat associations of mammals in urban habitats differ from their rural counterparts (Bozek et al., 2007; Kanda et al., 2009), very few explore this relationship at high levels of urbanization. We began addressing this gap with an understudied, urban adapted mesocarnivore, the opossum, using a large database of motion-triggered camera images from across the Chicago, Illinois region. Our results suggest opossum habitat dynamics within highly urban landscapes contradict existing knowledge about how this species responds to lower levels of urbanization in two significant ways: opossum in highly urban environments exhibited a strong need for patches near natural water sources, and opossum are negatively impacted by housing density within cities. As cities continue to grow in size and density it is necessary to not only study wildlife at the urban-wild interface but also within urban cores in order to conserve biodiversity in an urbanizing world. Though urban centers today exhibit decreased species diversity (Aronson et al., 2014), this need not be the case for cities in the future with proper research and city planning.

To manage the opossum in urban environments, proximity to natural water sources should be considered. If human-wildlife conflict occurs far from natural water sources it would be necessary to identify and limit potential anthropogenic water sources. As opossums have higher occupancy rates near natural water sources within the city it may be more cost effective to focus educational programs on cohabitating with urban adapted species in these locations, especially in areas with high housing densities. As the world continues to urbanize and metropolitan areas sprawl, discerning the habitat dynamics of urban species can help us to
better understand the ecology of our cities and facilitate cohabitation with wildlife, but in order to manage or conserve wildlife in these areas we must first understand how species select habitat within the landscape itself.

F. Acknowledgements

We thank M. Ambrocio, L. Bainbridge, J. Bender, C. Chung, F. Cortes, J. Frumkin, H. Grant, D. Hughes, M. Kestufskie, J. Kilgour, K. Limbrick, A. Ravera, J. Rios, M. Vernon, and others for assistance with fieldwork. Detailed suggestions provided by reviewers greatly increased the quality of this manuscript. We are grateful to the many landowners who allowed us access to their property, including the forest preserve districts of Cook, DuPage, Lake, and Will counties, the Illinois Department of Natural Resources’ Nature Preserve Commission, the Chicago Park District, and the Archdiocese of Chicago. Funding was provided by the University of Illinois at Chicago, the Lincoln Park Zoo, and the Davee Foundation.

CITED LITERATURE


McKinney, M.L.: Urbanization, biodiversity, and conservation: the impacts of urbanization on native species are poorly studied, but educating a highly urbanized human population about these impacts can greatly improve species conservation in all ecosystems. *BioScience* 52;883-890, 2002.

Meier, K.E., Habitat use by opossums in an urban environment. Master’s Thesis, Oregon State University, Corvallis, 1983.


III. WHEN IT TAKES MORE THAN TWO TO TANGO: USING MULTI-SPECIES OCCUPANCY MODELS TO ESTIMATE PATTERNS OF CO-OCURRENCE IN COMMUNITY MODULES

A. Abstract

While ecology is rife with theory that explores how multiple species co-occur, we lack robust statistical models to parameterize this theory with empirical data, particularly when species are detected imperfectly. We address this need by developing an occupancy modeling framework that estimates multi-species co-occurrence patterns from partial sampling and highlight the utility of our approach with simulated and real datasets. Simulations with three-species community models revealed that our model was able to accurately (calculated as mean absolute error) disentangle interspecific effects from baseline colonization or persistence rates. As a real-world illustration of our model, we estimate patterns of co-occurrence between raccoons (*Procyon lotor*), coyotes (*Canis latrans*), and Virginia opossum (*Didelphis virginiana*) in Chicago, Illinois, USA with data from a multi-year camera trapping study. Coyotes did not influence the persistence rates of the two smaller mesocarnivores, but raccoon and opossum were, on average, 2-3 times more likely to persist in habitat patches when the other species was present. Beyond the current application, our model allows for explicit tests of a wide range of ecological dynamics of interest, such as estimating priority effects, and is a robust and generalizable framework that accurately estimates patterns of co-occurrence and interspecific associations, even when species are detected imperfectly.
B. Background

Many ecological investigations focus on patterns of species occupancy rather than abundance, as occupancy-based sampling designs require less effort and are well-suited to answer a variety of questions about species distributions (MacArthur and Wilson, 1967), metapopulation (Levins, 1969) or metacommunity dynamics (Holyoak et al., 2005), and species invasions (Elton, 1958). However, providing robust estimates of species occupancy rates requires correctly partitioning observed variation between underlying ecological and observational processes. Estimated occupancy rates can be highly biased when detection rate or sampling design are not properly addressed in the model formulation (MacKenzie et al., 2002). Indeed, the realization of ecologists’ fallibility and the impact this may have on our inferential capacity has been a motivating force in the development of statistical methods that account for observational error (MacKenzie et al., 2006; Royle and Dorazio, 2008 and references therein).

Of the statistical techniques that correct for issues of detectability, the patch occupancy modeling framework developed by MacKenzie et al. (2002) has proven to be an effective technique to understand the spatiotemporal dynamics of species distributions (Bailey et al., 2014; hereafter referred to as occupancy models). Although they are similar to other approaches that use repeated samples at a survey location to estimate species detectability and occurrence (e.g., Geissler and Fuller, 1987; Azuma et al., 1990), occupancy models leverage statistical techniques to simultaneously estimate detection and occurrence in a model-based framework. Since their introduction 15 years ago (MacKenzie et al., 2002), occupancy models have been developed to incorporate multiple seasons (MacKenzie et al., 2003), patterns of co-
occurrence for two species (MacKenzie et al., 2004; 2006), and dynamic processes for entire communities through the inclusion of continuous and discrete covariates (Dorazio et al., 2010). These models are versatile and produce robust likelihood estimates, facilitating the testing of a variety of hypotheses on species occupancy or colonization-extinction dynamics, all while accounting for the vagaries of sampling.

Although occupancy models have been modified to accommodate multiple species to test how features of the physical environment influence occupancy rates, they lack the ability to estimate interspecific associations in groups of potentially interacting species through time (i.e., > 2 species, but see Rota et al., 2016 for a single-season formulation). This is unfortunate, as ecology is rife with theory that explores how groups of species may co-occur spatially and temporally (MacArthur, 1972; Rosenzweig, 1995; Chase and Liebold, 2003; Holyoak et al., 2005; Vellend, 2016). Empirical studies exploring the co-occurrence of multiple species either do not account for non-detection and therefore use biased estimates of occurrence and co-occurrence (e.g., Oliver et al., 2009; Byholm et al., 2012), or apply 2-species co-occurrence models iteratively throughout the entire community, which cannot account for the impact of indirect effects such as apparent competition or intraguild predation (e.g., Lesmeister et al., 2015). Thus, there exists a need to develop robust and generalizable statistical models that can accommodate both imperfect detection and interactive effects between multiple species.

How do we untangle the reticulate web of interactions that ecological communities weave so that we may objectively identify patterns of co-occurrence? One approach would be to stride into the quagmire of entire communities and estimate rates of co-occurrence between every species pair. While holistically admirable, this is typically not possible in empirical
communities due to the curse of dimensionality (the near exponential increase in the number of potential direct interactions as the number of species increase), save for cases where the whole community is composed of relatively few species (Holt and Hoopes, 2005). A more tractable approach is to examine community modules, or discrete groups of species that interact in specific manners (e.g., apparent competition) within more complex food webs (Holt and Lawton, 1994). As species in these discrete groups likely influence each other’s distribution (see Holt and Hochberg (2001) and references therein), the presence of one species in a habitat patch may be an important predictor variable for the distribution of another.

Here, we introduce a multivariate occupancy model for the analysis of community modules. We leverage the species incidence matrix to estimate pair-wise patterns of co-occurrence that simultaneously account for the influence of other species and imperfect detection. To confirm the ability of this model to estimate ecological interactions via co-occurrence, we conducted simulation studies using two different simulated community modules: a predator with two competing prey species and three-species horizontal competition. We then apply our model to a large-scale camera trapping study of medium to large mammals in Chicago, Illinois, USA (Magle et al., 2015; Fidino et al., 2016). Overall, our co-occurrence model provides a robust and generalizable framework that accurately estimates patterns of co-occurrence and interspecific associations within community modules.

C. **Model development**

1. **Incorporating first-order species associations into occupancy models**
For a single season, the number of surveys in which a species is observed at a site is modeled as a binomial process:

\[ y_{i,k} \mid Z_{i,k}, \rho_{i,k} \sim \text{Binomial} \left( j_{i,k}, Z_{i,k} \rho_{i,k} \right) \quad (3.1) \]

where \( y_{i,k} \) represents the number of surveys in which species \( i \) was observed at site \( k \), \( Z_{i,k} \) is the true occupancy state for the species at the site (with \( Z \) as the matrix of true occupancy states for all species across sites), \( \rho_{i,k} \) is the conditional probability of detecting species \( i \) at site \( k \) given its presence, and \( j_{i,k} \) is the total number of surveys conducted at site \( k \). For example, if species 2 was present at site 6, then \( Z_{2,6} = 1 \) and the species is detected at probability \( \rho_{2,6} \) across the \( j_{k} \) surveys. The occupancy state of a species, \( Z_{i,k} \), is treated as a random variable and is modeled as a Bernoulli trial with probability of occupancy \( \Psi_{i,k} \).

\[ Z_{i,k} \mid \Psi_{i,k} \sim \text{Bernoulli} \left( \Psi_{i,k} \right) \quad (3.2) \]

This probability can be related to covariates through a variety of link functions. In this case we apply the logistic-link:

\[ \logit( \Psi_{i,k} ) = \beta_{i,0} + \beta_{i,1} x_{k,1} \cdots \beta_{i,q} x_{k,q} \quad (3.3) \]

where \( \beta_{i,0} \) is the intercept value for species \( i \), \( \beta_{i,q} \) is the effect of covariate \( q \) on species \( i \), and \( x_{k,q} \) is the measured value of covariate \( q \) at site \( k \).

We add interspecific associations via the occupancy impact matrix, \( A \), which contains association coefficients:

\[
A = \begin{bmatrix}
0 & \alpha_{1,2} & \alpha_{1,3} \\
\alpha_{2,1} & 0 & \alpha_{2,3} \\
\alpha_{3,1} & \alpha_{3,2} & 0
\end{bmatrix}. \quad (3.4)
\]

In this case, the off-diagonal terms represent pair-wise co-occurrence rates between the other species with species \( i \). Given this specification, we assume that associations do not occur
beyond the first-order. However, the model could be generalized further to include higher orders. As the number of sites species occupy does not change within a single-season the occupancy impact matrix is symmetric. Thus, $\alpha_{1,2}$ represents to what degree species 2 co-varies with species 1, is identical to $\alpha_{2,1}$, and therefore these two quantities need not be estimated separately. For inclusion within the linear model (Eq. 3.3) predicting the logit of the occupancy of species $i$ in site $k$, we multiply the $i$th row of $A$ in Eq. 3.4 by the $k$th column of $Z$:

$$\text{logit}(\Psi_{i,k}) = \beta_{i,0} + \alpha \ Z_k + \beta_{i,1} x_{k,1} \ldots \beta_{i,q} x_{k,q}. \quad (3.5)$$

As the diagonal of the impact matrix is zero, Eq. 3.5 only estimates the effects that other species have on species $i$ (i.e., $Z_{i,k}$ is always multiplied by zero).

We also allow the probability of detecting a species to be a function of covariates:

$$\text{logit}(\rho_{i,k}) = p_{i,0} + p_{i,1} x_{k,1} \ldots p_{i,q} x_{k,q}. \quad (3.6)$$

where $p_{i,0}$ is the intercept value of detecting species $i$ at site $k$, $p_{i,q}$ is the effect of covariate $q$ on species $i$, and $x_{k,q}$ is the measured value of covariate $q$ at site $k$. If it is expected that the presence of one species influences the detectability of another, a detection impact matrix may be created and applied in the same format as the occupancy impact matrix.

2. **Extending species associations to multiple seasons**

Surveying a community over a single season may allow for a brief glimpse at patterns of species occupancy, but it is often of greater interest to understand how patterns of occupancy change via local colonization and extinction events (Hanski, 1999), which requires repeated surveys across multiple seasons. To extend our modeling approach, we consider that species’ occurrences through time are first-order Markov processes, where the probability of occupancy
in the next season \((t + 1)\) depends on the occupancy now \((time \ t)\) and a mixture of local colonization and persistence probabilities. The number of surveys in which a species is observed at a site across multiple seasons is:

\[
y_{i,k,t} \mid Z_{i,k,t}, \rho_{i,k,t} \sim \text{Binomial} \left( j_{i,k,t}, Z_{i,k,t}, \rho_{i,k,t} \right)
\]

where the definitions for \(y, Z, \rho, \) and \(j\) are identical to Eq. 3.1 save for the fact that they may vary over \(T (in \ t = 1...T)\) time steps. We have no information regarding the location of species before sampling began and so the probability of occurrence at the first time step is:

\[
Z_{i,k,t=1} \mid \Psi_{i,k,t=1} \sim \text{Bernoulli} \left( \Psi_{i,k,t=1} \right)
\]

where

\[
\logit(\Psi_{i,k,t}) = \beta_{i,0} + \beta_{i,1} x_{k,1} \cdots \beta_{i,q} x_{k,q}
\]

which is identical to a single-season model (Eq. 3.2 and 3.3). Changes from the first time step to following time steps are modeled as a first-order Markov process

\[
Z_{i,k,t+1} \mid Z_{i,k,t}, \Phi_{i,k,t} \gamma_{i,k,t} \sim \text{Bernoulli} \left( \Phi_{i,k,t} Z_{i,k,t} + \gamma_{i,k,t} (1 - Z_{i,k,t}) \right) \text{ for } t = 1,2,...,T-1,
\]

where

\[
\logit(\gamma_{i,k,t}) = c_{i,0} + c_{i,1} x_{k,1} \cdots c_{i,q} x_{k,q}
\]

and

\[
\logit(\Phi_{i,k,t}) = d_{i,0} + d_{i,1} x_{k,1} \cdots d_{i,q} x_{k,q}
\]

are respectively the logit-scale probability of persistence (i.e., 1 – probability of extinction) and colonization, both of which can be impacted by covariates. Again, a vector of associations can be included in the linear predictor for either process:
\[
\text{logit}(\gamma_{ik,t}) = c_{i,0} + \alpha_{i}Z_{k,t} + c_{i,1}x_{k,1} \ldots c_{i,q}x_{k,q} \quad (3.13)
\]

for persistence and

\[
\text{logit}(\Phi_{ik,t}) = d_{i,0} + \alpha_{i}Z_{k,t} + d_{i,1}x_{k,1} \ldots d_{i,q}x_{k,q} \quad (3.14)
\]

for colonization. Depending on the research question, impact matrices could be included for \(\Phi\), \(\gamma\), or both (in the event that impact matrices are included for both, \(A\) would become a 3-dimensional array with the last dimension indexing which process is being estimated). Unlike single-season models, species may asymmetrically blink on and off at sites through time in multi-season models (e.g., a predator may cause the local extinction of a prey species). Thus, the multi-season formulation allows for the possibility of an asymmetric impact matrices (e.g., \(\alpha_{1,2} \neq \alpha_{2,1}\)).

Finally, the probability of detecting species in the multi-season model is

\[
\text{logit}(\rho_{ik,t}) = p_{i,0} + p_{i,1}x_{k,1} \ldots p_{i,q}x_{k,q} \quad (3.15)
\]

If time-varying covariates are not included within the detection probability than this portion of the model simplifies to the detection function of a single-season model (Eq. 3.6).

The model structure is kept simple and general here for the sake of illustration, but it can be expanded to answer a variety of ecological hypotheses and is robust to various sampling designs. For instance, one could include temporally periodic processes, account for spatial autocorrelation, or estimate how interspecific associations impact species detection rates.

D. Simulations

1. Simulations conducted
We now simulate data sets where species influence each other’s probability of persistence or colonization and compare the accuracy and precision of parameter estimates from our model with varying amounts of data (4, 6, and 8 seasons). To begin exploring the flexibility of our model we apply it to two different community modules: predation on two competing prey species and three-species horizontal competition. For each module, we simulated $Z_{i,k,t}$ for 3 species at 100 sites over 4, 6, and 8 seasons worth of data 200 times, the results of which were then stored in a three-dimensional array, $Z$. To do so, we randomly generated the initial occupancy state (via Eq. 3.8) at the first time step from species-specific occupancy probabilities that were drawn from a beta distribution with $\alpha = 2$ and $\beta = 2$ (mean = 0.5, sd = 0.22). Future occupancy states were then a function of colonization, $\gamma_{i,k}$ (if a site lacked the species), or persistence, $\Phi_{i,k}$ (if the species was present). To keep our examples tractable, we assumed that colonization and persistence rates did not vary across seasons and we respectively drew species-specific values for $\gamma_{i,k}$ and $\Phi_{i,k}$ from normal distributions with means of logit(0.4) and logit(0.7) and standard deviations of 0.5. We imposed community structure along an environmental gradient by including a covariate that influenced the colonization and persistence rates of each species via coefficients that were drawn from a normal distribution with a mean = logit(0.5) and standard deviation = 1. Site-specific values of the environmental variable were drawn from a normal distribution with mean = 0 and standard deviation = 1 and were held constant through time. To ensure that the simulated data had reached a steady state and was not influenced by the random placement of species at the first time step (Eq. 3.8), simulations were run for 10 pre-sampling time-steps before collecting 4, 6, or 8 seasons of data.
In these simulations we included an impact matrix, $A$, for colonization or persistence (but not both at the same time). We represented a community module of predation on two competing prey species via the logit-scaled impact matrix, $A_{p.c.p}$:

$$A_{p.c.p} = \begin{bmatrix} 0 & -1.5 & -2.0 \\ -0.5 & 0 & -1.0 \\ 2.0 & 1.0 & 0 \end{bmatrix}.$$ 

When included with $\Phi$ (Eq. 3.14), $A_{p.c.p}$ signifies that species 2 is more likely to competitively exclude species 1 (i.e., $\alpha_{2,1} > \alpha_{1,2}$) while the predator species, 3, negatively influences species 1 and 2 and benefits from their presence. When included with $\gamma$ (Eq. 3.13), both prey species are less likely to colonize sites when either their competitor or predator are present while the predator is more likely to colonize sites that have its prey. We represented the community module of three-species horizontal competition, $A_{h.c}$ (on the logit-scale) as:

$$A_{h.c} = \begin{bmatrix} 0 & -0.5 & 0 \\ -1.0 & 0 & -0.7 \\ -2.0 & -1.0 & 0 \end{bmatrix}.$$ 

Here, species 1 is the least influenced by the presence of the other species, followed by species 2 and then 3. Because the empirical example presented later comes from a study using camera traps, we formulated the observation component of the simulations to mimic their use. Although camera traps may have a large number of repeated counts (i.e., days of sampling) within a season, five years of camera trapping with over 1 million camera trap images have revealed that it is uncommon to capture images of each species present at a site each day. Thus, we kept species-specific daily detection probabilities low and drew them from a normal distribution with mean = $\text{logit}(0.15)$ and standard deviation = 0.5. The number of days a site was active, $j_{i,k,t}$, was simulated to imitate the variability in sampling effort that is inherent to
camera traps due to issues such as fluctuating battery life. As such, daily probabilities for a
simulated active ‘camera trap’ were calculated directly from 10 seasons of the ongoing study
we introduce in the following section, and the number of days a ‘camera trap’ was active per
season, \( j_{i,k,t} \), was drawn from a binomial distribution at these specific daily probabilities
(Appendix A). Following this, the number of days a species was observed, \( y_{i,k,t} \), was simulated
via Eq. 3.7.

Therefore, for every simulation we generated the true occurrence state, \( Z_{i,k,t} \), the
number of days a site was sampled, \( j_{i,k,t} \), the number of days a species was observed, \( y_{i,k,t} \), and
the true simulated parameter values. We then analyzed the simulated data using our
occupancy model (hereafter \( M_{\text{ins}} \)), which included species-specific responses to detection
probability, colonization, an environmental covariate influencing colonization and persistence,
and an impact matrix for either colonization or persistence (depending on the data that was
simulated). We used a Bayesian approach for parameter estimation and used weakly informed
priors for each parameter (joint posterior distributions of models are available in Appendix B).
The accuracy and precision of parameter estimates over these 12 scenarios was compared by
calculating the mean absolute error (MAE) of parameter estimates from their true value,
determining whether or not the true parameter value was within the 95% credible interval of a
parameter estimate (hereafter, coverage) and measuring the overall width of a parameter’s
credible interval (hereafter, precision).

Models were implemented in JAGS (Plummer, 2003) ver 4.0.0, but written in and
interfaced through program R ver 3.2.3 (R Core Team, 2015) with the package runjags
(Denwood, 2016). All simulations prior to statistical analysis were written in R. Following a
3,000 iteration adaptation and a 10,000 iteration burn-in, the posterior distribution for each parameter was sampled a total of 100,000 times across seven chains in parallel. However, MCMC chains were thinned by 10 to accommodate the computational intensity of tracking each \( Z_{i,k,t} \). To check for model convergence, Gelman-Rubin diagnostics were checked for all parameters to ensure that they were < 1.10 (Gelman et al., 2014).

2. **Simulation results**

Across all simulated scenarios, the off-diagonal elements of the impact matrix became more accurate and precise as the number of seasons of data increased (Fig 4; see Appendix C for a more thorough description of the results from our simulation study). Furthermore, \( M_{\text{inx}} \) was able to separate interspecific effects from a species baseline colonization or persistence rate and estimated species-specific intercepts with high accuracy and precision (Fig 4; Appendix C). Larger interactive effects between species were less accurate and precise than smaller interactive effects (e.g., \( \Phi_{3,1} \) vs \( \Phi_{1,2} \); Fig. 3), though increasing the amount of data did result in better estimates of the true parameter value. Mean coverage of all off-diagonal elements was high and close to the nominal level of 95% across all scenarios, and the median coverage for 4, 6, and 8 seasons of data across all scenarios was respectively 0.93, 0.94, and 0.95 (for coverage of all parameters see Appendix C).
Figure 4. The A) accuracy, B) precision, and C) coverage of model parameters (1-9) estimated by the co-occurrence model for three-species horizontal competition where species influenced each other’s persistence rates when 4, 6, and 8 seasons of data are supplied (x-axes). Accuracy is measured as the mean absolute error (MAE) of parameter estimates, precision as the measured width of a parameters 95% credible interval, and coverage as the proportion of the 200 models whose posterior distribution contained the true parameter value with its 95% credible interval. The alignment of the model parameters (1-9) reflects the impact matrix with the species-specific intercepts replacing the 0-valued diagonals. For accuracy and precision, the gray dots are median estimates across the 200 simulations for each scenario while the black lines are 95% highest density intervals. Coverage is represented by a single value, and note the scale of the y-axes.
Taken together, these simulations indicate that $M_{\text{mixs}}$ is able to correctly estimate interspecific associations when they occur, and it does so with increased accuracy and precision when additional data are supplied. Coverage of interspecific effects is still high with four seasons of data, but highly inaccurate and imprecise. Conversely, model estimates continue to improve as more data is supplied and the off-diagonal terms are estimated more accurately with little to no loss in coverage. For more details on the simulation results, see Appendix C.

E. **Analysis of mesocarnivore community dynamics in Chicago, IL**

In this section we apply our statistical framework to the analysis of three mesocarnivore species from a large-scale long-term camera trap study in the Chicago Metropolitan area (41° 50’ 15” N, 87° 40’ 55” W). Specifically, we analyze the detections of raccoons (*Procyon lotor*), coyote (*Canis latrans*) and Virginia opossum (*Didelphis virginiana*, hereafter opossum) by camera traps over 13 sampling seasons from spring 2010 to spring 2013. We provide brief details of the sampling design here and a more thorough explanation of the sampling protocol can be found in Vernon et al. (2015), Magle et al. (2015), or Fidino et al. (2016).

1. **Sampling protocol**

Cameras were set for 28 days each spring, summer, fall, and winter along three 50 km transects that radiate outwards from downtown Chicago, Illinois, USA. Four different site types were selected for this study, which collectively represents a large proportion of the potential wildlife habitat throughout the city: cemeteries, golf courses, city parks, and forest preserves. In total, 118 sites were sampled, though every site was not sampled each season because some were added or retired throughout the course of the study (i.e., there is missing data). Sites
were only used in this analysis if they had two or more seasons worth of data (n = 98). The number of days a species was observed, $y_{i,k,t}$, and the total number of days a camera trap was active, $j_{i,k,t}$, were calculated from the collected camera trap images.

2. **The model**

As the presence of coyotes may negatively influence raccoons and opossum in local habitat patches (Crooks and Soulé, 1999), our model included an impact matrix for the probability of persistence. In addition to this, we included two habitat covariates on colonization and persistence: patch area ($m^2$) and distance to the city center (m). To collect the covariate data, we used ArcGIS ver.9.3 (ESRI, 2008) to create layers of these landscape factors from Illinois Geospatial Clearinghouse data (2003). Although the distance to the city center does not capture the full complexity of an urban system and assumes urbanization radiates outwards from a central point, in Chicago this covariate shows strong negative correlations with housing density (-0.69, 1 k buffer around sites) and percent impervious surface (-0.80, 1 k buffer around sites) and positive correlation to canopy cover (0.59, 1 k buffer around sites; Illinois Geospatial Clearinghouse data, 2003). Thus, for this example we use this single covariate to represent the cumulative effects of these other factors. Both environmental covariates were z-transformed for this analysis. Patch area was log-transformed before this scaling. Intercept values and habitat covariate beta estimates for colonization and persistence were allowed to vary among species, and habitat variables were allowed to influence both colonization and persistence. We assumed that the conditional probability of detection varied among species and sampling period. As before, we employed a Bayesian approach to parameter estimation and model sampling and used weakly informative priors (Gelman et al. 2008; joint posterior
distribution in Appendix B). In addition to Gelman-Rubin diagnostics, posterior distributions were checked visually to ensure convergence.

3. Results

Across all seasons from spring 2010 to spring 2013 coyote, opossum, and raccoon were respectively photographed at 76.77%, 75.76%, and 86.87% of 99 sites. In the same order, average naïve occupancy rates per season calculated from the observed data for each species were 33.15 ± 9.47%, 29.70 ± 15.58%, and 43.54 ± 10.66%. After accounting for imperfect detection, average occupancy rates per season for coyote, opossum, and raccoon increased to 47.88 ± 7.11%, 36.46 ± 13.52%, and 62.22 ± 12.70%, respectively. Raccoons had the highest average colonization rate and were twice as likely as coyotes or opossum to colonize habitat patches that they did not occupy in the previous time step (Fig. 5).

In general, the habitat patches most often colonized by all species were large in size and further from the center of the city (Fig. 5). Raccoons were the only species more likely to persist in habitat patches further from the city, and patch area had little to no effect on species persistence rates (Fig. 5).

Raccoon and opossum and were respectively 3.32 (2.02 – 5.50) and 2.88 (1.51 – 5.54) times more likely to persist in patches together than alone (Fig. 6). Although the presence of coyotes may be expected to have strong negative effects on smaller mesopredators (Crooks and Soulé, 1999), coyote presence did not alter opossum and raccoon persistence rates. Coyote, on the other hand, were 1.98 (1.11 – 3.66) times more likely to persist at sites with raccoons.
Figure 5. Parameter estimates for coyote, opossum, and raccoon colonization ($\gamma$) and persistence ($\Phi$) as a function of patch area ($m^2$) and distance from the city center (m) in Chicago, Illinois, USA. Violin plots are posterior densities within the bounds of a parameters 95% credible interval. The vertical dashed line is centered at zero.
Figure 6. The estimated influence that the presence of a coyote, opossum, or raccoon have on each other’s probability of persistence from one season to the next throughout Chicago, Illinois. Bolded lines indicate logit-scale parameter estimates from the persistence impact matrix whose 95% credible intervals do not bound zero. Species-specific intercept values for coyote, opossum, and raccoon in the absence of other members of the community were respectively 0.24 (-0.20 – 0.72), 0.07 (-0.47 – 0.58), and 0.24 (-0.14 – 0.63).

As these results are correlative they do not imply facultative mutualism between species, however, they likely indicate the existence of unidentified components of the natural environment that increase species persistence rates within the habitat patches where they co-occur. For example, variation in resource availability at scales smaller than a habitat patch may not be sufficiently captured by our tested covariates. As raccoon and opossum have overlapping diets (Feldhamer et al., 2003), sites with a sufficient resource base may be more
likely to increase the persistence of these species which would then be estimated as a positive co-occurrence between species.

F. Discussion

We have presented here a statistical modeling framework able to analyze species associations within community modules as well as spatiotemporal dynamics and imperfect detection. Furthermore, this approach can quantify patterns of co-occurrence between two species while simultaneously accounting for the influence of others in a highly versatile model-based framework. This, in turn, allows for more explicit tests of dynamics that are of interest in ecological communities (Holyoak et al., 2005; Vellend, 2016).

While we only analyzed three species community modules, the model is general enough to include any number of potentially interacting species. Moreover, multiple community modules or one-way interactions can be specified and estimated within the same model by setting some elements of an impact matrix to zero. Other potentially important ecological quantities can also be estimated within the model or summarized from the posterior distribution as well, such as turnover rates or species richness in different habitat types (Dorazio et al., 2010). Additionally, as it is an extension of Dorazio et al. (2010), this model can also be used to specify how different metacommunity paradigms (e.g., species sorting along environmental gradients) influence local colonization and extinction dynamics. Our approach adds to their model the ability to test for additional relevant components of metacommunity theory, such as priority effects and other types of species interactions (Holyoak et al. 2005). Patch dynamics, one of the four paradigms of metacommunity theory, assumes that
homogeneous habitat patches undergo stochastic colonization events by species but deterministic extinction events via competition (Vellend, 2016). Such a pattern could easily be specified and empirically tested by including an impact matrix in the persistence function of our model.

It is, however, critical to consider the temporal scale at which these associations are estimated when interpreting coefficients from the impact matrix. In our example, raccoon and opossum positively influenced each other’s persistence rates, but these estimates were over 28 days of observation and should not be interpreted at smaller temporal scales (Fig. 6). Within a season, species may temporally avoid each other despite co-occurring in the same habitat patch, and the current model parameterization would not capture this differentiation. If associations at a smaller temporal scale are of interest, sampling methods could easily be altered to reduce the primary sampling period to the relevant temporal scale. For example, multiple observers could conduct independent observations of the same site on the same day, or the primary camera trap sampling period could be reduced to a single day (with secondary sampling units by the hour) if there is sufficient data. Thus, careful consideration needs to be taken when designing studies in order to most effectively answer pertinent questions, parameterize models, and estimate ecologically relevant quantities of a study system (Mackenzie et al., 2002). This generality is a major benefit of the model, as it can estimate quantities over many temporal scales.

Estimating co-occurrence does come at some cost, though. For example, we did observe a general decrease in precision when estimating species associations relative to species baseline colonization or persistence rates, particularly for larger interspecific effects (Fig. 4,
Appendix C). This understandable reduction in precision, however, does not appear to influence parameters outside of the impact matrix (Appendix C). Under a Bayesian approach, a variety of tactics could be used to increase precision by decreasing the variance in the priors. For example, prior knowledge on species associations from other studies could be incorporated into a model or pilot data could be used to constrain the parameter space. Conversely, many new technologies make it increasingly easy to collect large amounts of data. Studies could therefore be designed such that some component of the collected data is used to train model parameters. Regardless of the method used to reduce prior variance, some loss of precision is to be expected as model complexity increases, and thus researchers must carefully consider the relevant ecology of species when constructing occupancy models.

With regards to application, this modeling approach would be a useful tool to, for example, determine how invasive species influence wildlife communities by allowing land managers to estimate species-specific probabilities of colonization, persistence, or occupancy in the absence of some invasive competitor (Sanders et al., 2003). Assuming that continuous covariates are z-transformed, species-specific intercepts represent average persistence or colonization rates in the absence of other species and are therefore more biologically significant. For instance, a prey species may have high persistence rates in the absence of a predator or predators may not be able to persist in the absence of their prey species. Because standard occupancy models assume that species are independent they are unable to elucidate these types of interactive effects. The present modeling technique could also be used to determine the effectiveness of monitoring surrogate species in conservation biology (Caro and O’Doherty, 1999). While understanding the habitat associations of a more common umbrella
species may suggest habitat patches to conserve, estimating the probability of observing one or more umbrella species in a habitat patch given the presence of others while accounting for imperfect detection would provide a more statistically robust metric for use in conservation applications. Finally, we anticipate that this approach can be used to estimate where species are more or less likely to co-occur than random, facilitating follow-up smaller scale experiments to test potential mechanisms of coexistence (e.g., Tilman, 1982; Chesson, 2000).

As data become more accessible and cheaper to collect over larger areas, ecologists are now able to test theories over scales that were previously unattainable (Wolkovich et al., 2014). This is made even easier through advances in computational speed and statistical programming. The proliferation of both data and computational efficiency is a blessing and a curse, more data does not always mean better science, and testing every combination of potentially explanatory variables does little to unify theory to data. One major hurdle that limits Big Data-based ecological research is the ability to understand how to correctly quantify uncertainty related to the patterns and processes underlying the data that are collected. The model we present here is a versatile and robust approach that can fit a multitude of sampling designs and hypotheses that can provide empirical support to ecological theory that explores how potentially interacting species co-occur through space and time.

**CITED LITERATURE**


Geissler, P.H., Fuller, M.R.: Estimation of the proportion of area occupied by animal species. In: 
Proceedings of the Section on Survey Research Methods of the American Statistical 

Gelman, A., Jakulin, A., Pittau, M.G., Sung Su, Y.: A weakly informative default prior distribution 


Holt, R.D., Hochberg, M.E.: Indirect interactions, community modules and biological control: a 
theoretical perspective. In: Evaluating indirect ecological effects of biological control. E. 

Holt, R.D., Hoopes, M.F.: Food web dynamics in a metacommunity context. In: 
Metacommunities: spatial dynamics and ecological communities. M. Holyoak, M.A. 

Holyoak, M., Leibold, M.A., Holt, R.D.: Metacommunities: spatial dynamics and ecological 


IV. USING FOURIER SERIES TO ESTIMATE PERIODIC TRENDS IN DYNAMIC OCCUPANCY MODELS

A. Abstract

Some of the most impressive behavioral and reproductive adaptations of organisms are in response to periodic variability of the region they reside. To capture these temporal dynamics, statistical models that estimate the spatiotemporal distribution of a species currently include categorical seasonal covariates, temporally varying parameters, polynomial terms, or smoothing splines. While these techniques provide a useful starting point, they may require many parameters to estimate, be difficult to interpret or make predictions with, and do not explicitly capture periodic trends and deviations from them. Here, we present a technique that uses Fourier series to identify periodic trends in dynamic occupancy models, and parameterize them with data from a large-scale long-term camera trapping study of medium to large mammals in Chicago, Illinois, USA. Our periodic models accounted for up to 73% of the temporal variability in species colonization rates and outperformed standard occupancy models with temporally varying parameters in 4 of the 5 species analyzed. Overall, this method is able to partition variability between periodic and non-periodic trends in a highly versatile model-based framework. It also allows ecologists to easily estimate the proportion of temporal variability that is attributable to a given periodic trend and provides them the ability to incorporate prior knowledge on the natural history of a species into a statistical model. This should, in turn, create more biologically reasonable models for the conservation and management of species in temporally heterogeneous environments.
B. Background

Meteorological changes throughout the year bring about 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 in which populations of organisms reside. For example, seasonally dependent pulses in resource abundance can have a profound effect on the coexistence of species (Brown, 1989; Chesson et al., 2004). Some of the most impressive behavioral and reproductive adaptations of organisms are in response to seasonal shifts in the local abundance of essential resources (Immelman, 1971). Seasonal migrations to more favorable conditions for breeding or foraging, for example, are one of the most conspicuously observed events that occur across taxa (Berthold, 2001; Lucas et al., 2001; Holland et al., 2006). Likewise, in environments with a restricted time-span of favorable breeding conditions, the birth and subsequent dispersal of young generally follows a periodic trend (Bradshaw and Holzapfel, 2007). For example, in the temperate parts of North America, coyotes (Canis latrans) breed in late February and have pups in late May, many of whom subsequently disperse to new habitats in the Fall (Bekoff and Gese, 2003). Although the proximate mechanisms that facilitated the evolution of such responses in species are varied and debated (Holland et al., 2006; Boyle and Conway, 2007), the fact that the spatial distribution of many species varies temporally, and often periodically, is not.

To quantify such trends, statistical models that estimate the spatiotemporal distribution of species from observational data handle temporal dynamics through the inclusion of
categorical seasonal covariates (e.g., Nielsen et al., 2010), temporally varying parameters (e.g., Dorazio et al., 2010; Kéry et al., 2013), or through the use of polynomial terms or smoothing splines (Thorson et al., 2012). While such analytical techniques provide a useful starting point for understanding patterns that occur through time, they may require many parameters to estimate and can be difficult to interpret or make predictions with. Further, such techniques are unable to partition variability between temporally periodic trends in the data and deviations from them. Temporally varying covariates that follow a periodic pattern (e.g., local temperature) can be included within a model, provided such data exist, but provide inference for the temporal covariate, and not time itself. As periodic patterns are repeatable and therefore predictable, there exists an opportunity to leverage such patterns when they exist so long as it can be explicitly specified within a statistical model.

Here, we present a technique that uses Fourier series (Bloomfield, 2004) to identify periodic trends in dynamic occupancy models, a statistical framework developed by MacKenzie et al. (2003) that simultaneously estimates the detectability and spatiotemporal dynamics of a species with repeated samples of detection / non-detection data (Bailey et al., 2014). Fourier series have been used to quantify a variety of complex periodic trends in ecology (e.g., population dynamics, Grover et al., 2000; diel foraging patterns, Flury and Levri, 1999; stream discharge rates, Sabo and Post, 2008), and do so by converting a periodic signal into terms of frequencies driven by the possibly infinite sum of sines and cosines (Shumway and Stoffer, 2010). We first illustrate how to incorporate Fourier series into dynamic occupancy models, then we develop and parameterize a model with data from a large-scale camera trapping study
of medium to large mammals in Chicago, Illinois, USA (Vernon et al., 2015; Magle et al., 2015; Fidino et al., 2016).

Although Fourier series can be used to construct any type of periodic trend, we focus our attention on those informed by the life history of a target species. Such information can be easily collected in natural history texts (e.g., Feldhamer et al., 2003) and formulated as a simple Fourier series that represents an explicit formulation of a periodic pattern. Ecologists can use this method to incorporate prior knowledge on the natural history of a species into a statistical model. Further, this technique can also be used to partition variability between periodic and non-periodic trends in a highly versatile model-based framework. This should, in turn, create more biologically reasonable models for the conservation and management of species in temporally heterogeneous environments.

C. MODEL DEVELOPMENT

1. Multi-season occupancy models

As the formulation of dynamic occupancy models has been covered in depth (see MacKenzie et al. 2003; Dorazio et al. 2010), we briefly cover the topic here for a single species potentially located at $k$ in $1,2,...,K$ sites and sampled during $t$ in $1,2,...,T$ seasons using matrix notation. Our notation is described in Table III. For simplicity, we assume that we would like to quantify the effect of $Q$ covariates collected at all $K$ sites on the probability of initial occupancy ($\Psi$), colonization ($\gamma$), and persistence ($\Phi$) and the effect of $R$ covariates collected at all $K$ sites on the conditional probability of detection ($p$). However, the model is general enough such that different covariates could be used for each process, including those that temporally vary.
The number of surveys a species is detected at a site across multiple seasons is:

\[ y_{k,t} | Z_{k,t}, p_{k,t} \sim \text{Binomial}(j_{k,t}, Z_{k,t}p_{k,t}) \]  

(4.1)

The probability of occupancy at the first time step is

\[ Z_{k,1} | \Psi_k \sim \text{Bernoulli}(\Psi_k) \]  

(4.2)

Which can be made a function of \( Q \) covariates through the logit link

\[ \logit(\Psi_k) = x_k b \]  

(4.3)

Following the first time step we assume a first-order autoregressive relationship in that the future occupancy state in the next season (\( t+1 \)) depends on the current occupancy status (time \( t \)) via a mixture of local colonization (\( \gamma \)) and persistence (\( \Phi \)) probabilities.

\[ Z_{k,t} | Z_{k,t-1}, \Phi_{k,t}, \gamma_{k,t} \sim \]  

\( \text{Bernoulli} \left( \Phi_{k,t}Z_{k,t-1} + \gamma_{k,t}(1 - Z_{k,t-1}) \right) \; \text{for } t = 2,3,\ldots,T \]  

(4.4)

where

\[ \logit(\Phi_{k,t}) = x_k d \]  

(4.5)

and

\[ \logit(\gamma_{k,t}) = x_k m \]  

(4.6)

are the logit-scale probabilities for persistence (i.e., \( 1 - \) probability of extinction, Eq. 4.5) and colonization (Eq. 4.6), which are both functions of \( Q \) covariates. Finally, the conditional probability of detecting a species given its presence can also be made a function of \( R \) covariates such that
### TABLE III
NOTATION USED IN A DYNAMIC OCCUPANCY MODEL FOR A SINGLE SPECIES

<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$K$</td>
<td>The total number of sites.</td>
</tr>
<tr>
<td>$T$</td>
<td>The total number of seasons of data.</td>
</tr>
<tr>
<td>$y_{kt}$</td>
<td>The number of surveys a species was detected at site $k$ at time $t$.</td>
</tr>
<tr>
<td>$Z_{kt}$</td>
<td>The occupancy state of a species at site $k$ at time $t$.</td>
</tr>
<tr>
<td>$p_{kt}$</td>
<td>The conditional probability of detecting a species given its presence at site $k$ at time $t$.</td>
</tr>
<tr>
<td>$j_{kt}$</td>
<td>The total number of surveys conducted at site $k$ at time $t$.</td>
</tr>
<tr>
<td>$\Psi_k$</td>
<td>The probability of initial occupancy at site $k$ at time $t = 1$.</td>
</tr>
<tr>
<td>$Q$</td>
<td>The total number of covariates thought to influence $\Psi$, $\gamma$, and $\Phi$.</td>
</tr>
<tr>
<td>$X$</td>
<td>A $K \times (Q + 1)$ matrix of covariates. The first column of $X$ is a column of 1’s for the constant term in the regression. If covariates temporally vary $X$ becomes a three-dimensional array.</td>
</tr>
<tr>
<td>$x_k$</td>
<td>A row vector of length $Q + 1$. The first value in this vector is 1, which is multiplied by the constant term in the regression. The rest the elements in $x_k$ are the $Q$ covariate values collected at site $k$.</td>
</tr>
<tr>
<td>$b$</td>
<td>A column vector of $Q + 1$ regression coefficients which includes the constant term and the effect of $Q$ covariates on $\Psi$.</td>
</tr>
<tr>
<td>$\Phi_{kt}$</td>
<td>The probability of persistence at site $k$ at time $t$.</td>
</tr>
<tr>
<td>$d$</td>
<td>A column vector of $Q + 1$ regression coefficients which includes the constant term and the effect of $Q$ covariates on $\Phi$.</td>
</tr>
<tr>
<td>$\gamma_{kt}$</td>
<td>The probability of colonization at site $k$ at time $t$.</td>
</tr>
<tr>
<td>$m$</td>
<td>A column vector of $Q + 1$ regression coefficients which includes the constant term and the effect of $Q$ covariates on $\gamma$.</td>
</tr>
<tr>
<td>$R$</td>
<td>The number of covariates thought to influence $p_k$</td>
</tr>
<tr>
<td>$S$</td>
<td>A $K \times (R + 1)$ matrix of covariates. The first column of $G$ being a column of 1’s for the constant term in the regression. If covariates temporally vary $S$ becomes a three-dimensional array.</td>
</tr>
<tr>
<td>$s_k$</td>
<td>A row vector of length $R + 1$. The first value in this vector is 1, which is multiplied by the constant term in the regression, while the rest of the elements in $s_k$ are the $R$ covariate values collected at site $k$.</td>
</tr>
<tr>
<td>$f$</td>
<td>A column vector of $R + 1$ regression coefficients which includes the constant term and the effect of $R$ covariates on $p$.</td>
</tr>
<tr>
<td>$U_t$</td>
<td>Coefficient that allows the constant term of the regression to vary temporally. The deviation around the intercept is assumed to be Normal($0$, $\sigma^2$), where $\sigma^2$ is estimated from the data.</td>
</tr>
</tbody>
</table>
The above model allows for the estimation of how spatial features of a landscape influence $\Psi$, $\Phi$, $\gamma$, and $p$ but could easily be generalized to include time-varying parameters and covariates. For example, an additional term could be included to Eq. 4.5 or 4.6 so that intercepts vary over time. We illustrate this with Eq. 4.6.

$$\operatorname{logit}(y_{k,t}) = x_k m + U_t$$

$$U_t \sim \text{Normal}(0, \sigma^2).$$

As long as the continuous covariates are z-transformed, Eq. 4.8 allows the intercept to vary at each sampling season around an overall mean (the colonization intercept, $m_0$), and the variance in this parameter across sampling sessions can be estimated. Thus, Eq. 4.8 accounts for general temporal variability but does not allow for explicit tests of periodic components. We address this issue by briefly reviewing Fourier series and then illustrate and test two different periodic trends using data from a large-scale long-term camera trapping project (Fidino et al., 2016; Magle et al., 2015).

2. **A Brief Foray into Fourier series**

When a mathematical function with some periodic trend, $x(t)$, is known it can be approximated as a Fourier series with a single frequency over time, $t$,

$$x_p(t) = a_0 + \sum_{n=1}^{\infty} (a_n \cos(nw_0 t) + b_n \sin(nw_0 t))$$

Where $a_n$ and $b_n$ are Fourier coefficients that are 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 $P$ that measures the number of cycles per unit time (Kreyszig, 2010). For example, if $P = 8$, then the cycle repeats itself every eight time steps (or 0.125 cycles per observation). However, in the statistical analysis of time series, the summation over $n$ is generally not possible as there is no periodic function, $x_p(t)$, to derive $a_n$ and $b_n$; there is only data with a periodic pattern.

Instead, a periodic series to be estimated from data leverages the mathematical properties of Eq. 4.9 and is generally represented as:

$$x(t) = a_0 + (a \cos(nw_0 t) + b \sin(nw_0 t))$$

(4.10)

Where it can be shown that $a$ and $b$ are independent normally distributed random variables with mean zero and variances $\sigma^2$ (Shumway and Stoffer, 2010) that are estimated from the data. 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, is $\delta = \tan^{-1}(-b/a)$. Eq. 4.10 can be fit to data to estimate the presence of repetitive sinusoidal oscillations through time. To fit more complicated trends, however, Eq. 4.10 can be generalized to allow for up to $K$ mixtures of periodic series at varying frequencies and amplitudes,

$$x(t) = a_0 + \sum_{k=1}^{K} (a_k \cos(nw_k t) + b_k \sin(nw_k t))$$

(4.11)

Where $a_k$ and $b_k$ are independent normally distributed random variables with mean zero and variances $\sigma_k^2$, and $w_k$ are different frequencies. Through the summation of sines and cosines at varying frequencies 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 trends). Background on the theory and application of such analytical techniques are covered in depth in many texts on the statistical analysis of time series, and such methods are widely used across the sciences (e.g., Shumway and Stoffer, 2010; Prado and West, 2010; Box et al., 2015). Yet, while Eq. 4.11 can theoretically fit any periodic shape, it does so at the cost of 2 additional parameters for each \( k \). If a certain periodic trend is hypothesized to be present within a dataset and can be written as a periodic function, \( x_P(t) \), it is then possible to derive \( a_n \) and \( b_n \) and quantify the presence of such a pattern with only 2 parameters. When modeling the spatiotemporal distribution of a species we believe that there are many cases that this is possible if the natural history of a species is known. We illustrate how to do this in the following section, making use of the life history strategies of different species.

3. **Single-season pulses**

Let us assume we want to model periodic dispersal rates in species who breed once per year and estimate that trend in a dynamic occupancy model. From our camera trapping survey data in Chicago, Illinois, this categorization would describe the coyote (Bekoff and Gese, 2003), red fox (*Vulpes vulpes*; Cypher, 2003), and striped skunk (*Mephitis mephitis*; Rosatte and Larivière, 2003). Each year these species breed in the late winter to early spring and the resulting offspring disperse in the fall. Given this life history strategy, we would predict that habitat patches would experience a periodic pulse in colonization rates once per year as juveniles disperse to new habitats in the fall. Such a trend could be estimated with many parameters at different frequencies as in Eq. 4.11, but could not be estimated with a single frequency as in Eq. 4.10. However, it is possible to develop a periodic function of such a pulse
and derive its Fourier coefficients so that we may estimate the presence of this trend with only two parameters. Therefore, let us consider a periodic pulse (Fig. 7) that illustrates such a biological pattern and define it over a single period:

\[
x_P(t) = \begin{cases} 
A, |t| \leq \frac{P_u}{2} \\
-\frac{P}{2} < t < \frac{P}{2} \\
0, |t| > \frac{P_u}{2}
\end{cases}
\]  

(4.12)

\(x_P(t)\)

\(\frac{P_u}{2}

\(\frac{P}{2}

\(\frac{P}{2}

\(\frac{P}{2}

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)

\)
This function has a period of $P$, an amplitude of $A$, and a pulse width of $P_u$. Over a single period when the absolute value of $t$ is less than or equal to $P_u/2$, Eq. 4.12 takes the value $A$. Otherwise, it is zero. This represents a short pulse that is separated by a longer trough over time (Fig. 7). With our species and data, Eq. 4.12 could represent periodic pulses of increased colonization rates as juveniles disperse at a particular time over the period $P = 4$ (i.e., from sampling that occurs in the spring, summer, fall, and winter). As Eq. 4.12 is symmetric around its origin of zero, it is an even function (Kreyszig, 2010). Therefore, $b_n = 0$ for all $n$ and this function can be approximated with a Fourier cosine series

$$
\begin{align*}
    x_p(t) &= a_0 + \sum_{n=1}^{\infty} (a_n \cos(nw_0t)) \\
    \text{(4.13)}
\end{align*}
$$

Which, following Kreyszig (2010), means that $a_n$ can be written as

$$
\begin{align*}
    a_n &= \frac{2}{P} \int_{x}^{y} x_p(t) \cos(nw_0t) \, dt, \quad n \neq 0 \\
    \text{(4.14)}
\end{align*}
$$

Given $x_p(t) = A$ at any time between $-P_u/2$ to $P_u/2$ and 0 at any other point (Eq. 4.12), $a_n$ can be simplified and then solved. We illustrate this over the interval $-P/2$ to $P/2$

$$
\begin{align*}
    a_n &= \frac{2}{P} \int_{-\frac{P}{2}}^{\frac{P}{2}} x_p(t) \cos(nw_0t) \, dt \\
    \text{(4.15)}
\end{align*}
$$
\[ = \frac{2}{P} \frac{A}{nw_0} \sin(nw_0 t) \left|_{\frac{P_u}{2P}}^{\frac{P_u}{2}} \right. \]

\[ = \frac{2}{P} \frac{A}{nw_0} \left( \sin(nw_0 \frac{P_u}{2}) - \sin(-nw_0 \frac{P_u}{2}) \right) \]  

(4.15)

Furthermore, as \(\sin(a) - \sin(-a) = 2\sin(a)\) and \(w_0 = 2\pi/P\), Eq. 4.15 can be reduced to

\[ a_n = \frac{2A}{n\pi} \sin \left( \frac{n\pi P_u}{P} \right) \]  

(4.16)

As \(n \to \infty\), Eq. 4.13 more closely approximates the graphical representation of the pulse in Fig. 7. However, when included within a statistical model it is impossible to include infinite terms and computational efficiency is reduced as more \(n\) are included. In general, when estimating such a trend with discrete time steps (e.g., \(t = 1, 2, ..., T\)), our own trials suggest that \(n\) should at least range from 1, 2, ..., \(P\) for model fitting. Moreover, the pulse width, \(P_u\), is only a parameter of interest if one expects the pulse to occur over multiple time steps, which may be the case if \(P\) is large and the time between steps is small. If the pulse should only happen at one specific time step, as in our case, then \(P_u\) can be reduced to a value such that \(t + P_u/2 < t + 1\) so the pulse does not reach into the surrounding time steps. The natural choice is to set \(P_u\) to 1 so that it can then be removed from the equation. Further, Eq. 4.13 allows us to estimate the amplitude of a pulse, \(A\), but it does not let us estimate when the pulse occurs over \(P\). We add a phase shift term, \(\delta\), to the equation:

\[ x_P(t) = a_0 + \sum_{n=1}^{P} \frac{2A}{\pi n} \sin \left( \frac{\pi n}{P} \right) \cos \left( \frac{2\pi n}{P} (t - \delta) \right) \]  

(4.17)

So that the model can then estimate the presence of a periodic pulse at any time step over period \(P\) with two parameters, \(A\) and \(\delta\) (Fig. 8). In our example, this would provide a
general equation that can estimate if a periodic pulse occurs in the spring, summer, fall, or winter \((P = 4)\).

The next step is to incorporate Eq. 4.17 into the colonization function of a dynamic occupancy model so that we may estimate \(a_0, A,\) and \(\delta\) from the data. As \(a_0\) is just the average colonization rate across time, this can be used in place of the constant term in the regression (the intercept, \(m_0\)). Further, assuming that \(P\) is known a priori we can use the trigonometric formula \(\cos(u-v) = \cos(u)\cos(v) + \sin(u)\sin(v)\) and do some rearranging to separate the terms that contain \(A\) and \(\delta\) in Eq. 4.17 and specify them as parameters

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

\[
\beta_{2,n} = A\sin\left(\frac{2\pi n}{P} \delta\right), \quad \text{for } n = 1, 2, ..., P
\]

and combine the rest of the terms into two \(T \times P\) matrices of covariates that can be calculated outside of the model and included like any other temporally varying covariate

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

\[
S = \sin\left(\frac{2\pi n}{P} t\right)\sin\left(\frac{\pi n}{P}\right)\frac{2}{\pi n}, \text{for } n = 1, 2, ..., P \text{ and } t = 1,2, ..., T
\]  

These parameters and covariates can then be included within the linear predictor of our model.
Figure 8. Examples of periodic curves over discrete time for A) a single season pulse with period $P = 4$ as specified in Eq. 4.17 and B) a boom-bust sinusoidal curve with period $P = 2$ as specified in Eq. 4.25. Plots A and B are on the logit scale. $m_0$ controls the average colonization rate, $A$ controls the amplitude, and $\delta$ controls when a pulse occurs.

For example, including this in Eq. 4.8 would result in the linear predictor

$$
\text{logit}(\gamma_{k,t}) = x_k m + U_t + \sum_{n=1}^{P} \beta_{1,n} C_{t,n} + \beta_{2} S_{t,n}
$$

(4.20)

Which results in a model that can estimate a periodic pulse and deviations from it via the variance term in $U_t$. The above formulation only requires two parameters to estimate the pulse:
the amplitude \((A)\) and the phase \((\delta; \text{Fig. } 8)\). As \(A\) must be a positive logit-scale parameter to generate a pulse and \(\delta\) is an integer, we give them these vague priors:

\[
A \sim \text{Gamma}(1, 1) \quad (4.21)
\]

\[
\delta \sim \text{Categorical}(\rho_1, ..., \rho_p) \text{ for } \delta \in \{0, ..., P - 1\} \quad (4.22)
\]

\[
\rho_{1:p} = 1/P \quad (4.23)
\]

The variance term in Eq. 4.20 associated to \(U_t\) is a parameter of interest as it represents the temporal variability that is not accounted for by the periodic trend. By fitting a secondary model without the periodic component (e.g., Eq. 4.8) we can then calculate the proportion of temporal variability explained by the periodic trend

\[
1 - \frac{E(\sigma_1^2)}{E(\sigma_2^2)} \quad (4.24)
\]

Where \(\sigma_1^2\) and \(\sigma_2^2\) are respectively the expected values of the variance estimates from separate models that use Eq. 4.20 and 4.8 in their linear predictor with the same dataset.

**4. Multiple pulses a year**

Other species can breed multiple times a year or have young that disperse at different times in a year. Raccoons \((Procyon lotor)\) typically have a bimodal pattern in their mating distribution and juvenile dispersal occurs in the spring or fall (Gehrt, 2003). Virginia opossum \((Didelphis virginiana\), hereafter opossum) ordinarily have two litters per year, but occasionally have three (Gardner and Sunquist, 2003). Although opossum are primarily a nomadic species, males born in the first litter typically disperse in the fall of the same year while the second cohort disperses during the next breeding season in the spring (Gardner and Sunquist, 2003). Thus, we would predict a boom-bust pattern in these species colonization rates that peak in the
spring and fall, which we could approximate with a special case of Eq. 4.17 when \( n = 1 \). The reason for this is that the additional \( n \) terms in Eq. 4.17 help flatten the trough of the periodic pulse, while the first \( n \) is a simple sinusoidal curve that can estimate this pattern. When \( n = 1 \), Eq. 4.17 simplifies into a model that is identical to what Flury and Levri (1999) used to estimate periodic trends in logistic regression:

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

(4.25)

where the additional terms in \( a_n \) (Eq. 4.16) are removed as they are linear transformations of the random variable \( A \). Therefore, Eq. 4.25 is a special case of eq. 4.17 when \( n = 1 \), and can have the same priors for \( A \) and \( \delta \). However, in this case we would set \( P = 2 \) in order to capture two peaks within a single year (Fig. 8). Similar to the last section, Eq. 4.25 can be simplified down to parameters that are estimated from the data

\[
\beta_1 = A \cos \left( \frac{2\pi}{P} \delta \right)
\]

\[
\beta_2 = A \sin \left( \frac{2\pi}{P} \delta \right)
\]  

(4.26)

and two covariate vectors of length \( T \)

\[
c_t = \cos \left( \frac{2\pi}{P} t \right), \text{ for } t = 1, 2, ..., T
\]

\[
s_t = \sin \left( \frac{2\pi}{P} t \right), \text{ for } t = 1, 2, ..., T.
\]  

(4.27)

Which can then be included in a dynamic occupancy model that can estimate the presence of a boom-bust pattern in colonization rates, where colonization rates periodically switch from high to low as the seasons advance.
\[ \text{logit}(y_{k,t}) = x_k m + U_t + \beta_1 c_t + \beta_2 s_t. \]  

D. **Materials and Methods**

1. **Sampling protocol**

Starting in the spring of 2011, we set Bushnell motion-triggered camera traps (model 
#119436C) in the spring, summer, fall, and winter for 28 days per season along three 50km 
transects. These transects begin in downtown, Chicago, Illinois, USA and radiate outwards to 
the northwest, west, and southwest. Cameras were placed in four different site types, which 
together represent the bulk of potential wildlife habitat throughout the city: cemeteries, golf 
courses, city parks, and forest preserves. Furthermore, camera trapping sites were \( \geq 1 \) km from 
each other, but multiple sites could be placed within a single habitat patch if it was of sufficient 
size. Between spring 2011 and fall of 2013 118 sites were sampled, although every site was not 
sampled each season because some were added or retired over the course of the survey. As 
such, there is missing data. Furthermore, sites were only included in our analysis if they had at 
least two or more seasons worth of data \( n = 95 \). From the collected camera trap images we 
calculated the number of days a species was detected at a site \( y_{k,t} \) and the total number of 
days a camera trap was active \( j_{k,t} \). For more specific details of our sampling protocol, see 
Magle et al. (2016) or Vernon et al. (2016).

2. **Models, specification of priors, and model selection**
We included five species in our analysis: coyote, red fox, striped skunk, raccoon, and opossum. We fit four models for each species, which were identical save for what was included in the linear predictor of the colonization function (Table IV). We intentionally kept the estimation of initial occupancy simple as the colonization and persistence rates of the system control future occupancy states and are of more ecological importance. To account for spatial factors that may influence the probability of detection, colonization, and persistence we calculated the mean tree cover, mean impervious cover, and mean housing density within a 500m buffer around a site using QGIS ver. 2.14 (QGIS Development Team 2009). These values were then z-transformed. However, these covariates are all highly correlated. Pearson’s correlation coefficients between these covariates were 0.67 between mean impervious cover and housing density, -0.59 between mean housing density and canopy cover, and -0.71 between mean canopy cover and impervious cover. To address issues with multicollinearity in these spatial covariates, we used principal components analysis (PCA) to generate orthogonal principal components and used the first principal component in our model (hereafter URB), as it explained 77.31% of the variation in the data. Positive values of this metric 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 persistence in all models, and varied detection probabilities by site and time in the same fashion.
As the distribution of the species at the first time step (i.e., initial occupancy) is only a snapshot of the ecological dynamics of interest, we used empirical Bayes methods to specify the prior of the intercept for each model. To do so, we used our detection / non-detection data to calculate the shape parameters of a beta prior from the mean ($\mu$) and standard deviation ($\sigma$) of the proportion of sites a species was detected at each time step. Thus, the prior for the initial probability of occupancy for species $i$ is:

$$
\alpha_i = \left( \frac{1 - \mu_i}{\sigma_i^2} - \frac{1}{\mu_i} \right) \mu_i^2
$$

$$
\beta_i = \alpha_i \left( \frac{1}{\mu_i} - 1 \right)
$$

$$
\Psi_i \sim Beta(\alpha_i, \beta_i)
$$

(4.29)
For models that include a periodic component we used the priors specified in Eq. 4.21 for $A$ and 4.22 for $\delta$. We used the colonization function in Eq. 4.20 to predict coyote, red fox, and striped skunk colonization rates and set $P = 4$. Eq. 4.28 was used for raccoon and opossum with $P = 2$. Group-level parameters (i.e., random effects) were drawn from a Normal distribution with mean zero and a standard deviation hyperparameter set to half-Cauchy(0, 25). Following this, we used weakly informative priors on the logit-scale for all other parameters that were specified as Normal distributions with mean zero and standard deviation 1.83. Such a specification places 99% of the probability mass between -4.25 and 4.25, which represents realistic logit-scale parameters estimates (Gelman et al., 2008), particularly for occupancy models (Broms et al., 2016).

With regards to Bayesian model selection, many methods exist that vary in their assumptions, requirements, and ease of use (for an excellent review see Hooten and Hobbs, 2015). However, because multi-season occupancy models are mixture models that have the potential for spatial or temporal auto-correlation, options are more limited. Commonly used methods such as the deviance information criterion (DIC), Bayesian information criterion (BIC), or the Watanabe-Akaike information criterion (WAIC) should not be used with such data or models (Hooten and Hobbs, 2015). Instead, leave-one-out or k-fold cross validation techniques could be used to compare the out of sample predictive accuracy of a model, but such methods take significant amounts of computational time as the model must be refit many times. Instead, we take a more computationally effective approach and approximate the conditional predictive ordinates (CPO; Geisser, 1993) of each data point within an MCMC algorithm.
CPO_{k,t} \approx \frac{R}{\sum_{r=1}^{R} \log \left( \frac{1}{y_{k,t} | \theta^{(r)}} \right)}

(4.30)

where w = 1, ..., R are the MCMC iterations and \( y_{k,t} | \theta^{(r)} \) is the likelihood of an observed data point given the parameter estimates at that particular MCMC iteration of a given model. Thus, high values of CPO_{k,t} indicate observations that are more likely given the current model and parameter set. To estimate the overall performance of each model the summary statistic \(- \sum_{k,t} \log (CPO_{k,t})\) can be calculated, and the best fit model has the lowest value. Such an approximation closely mimics other more computationally intensive cross-validation techniques (Hooten and Hobbs, 2005).

All models were executed with JAGS (Plummer, 2003) ver 4.2.0, through program R ver 3.2.3 (R Core Team, 2015) and the runjags package (Denwood, 2016). The posterior distribution of each model was sampled 100,000 times after a 3,000 step adaptation and 3,000 step burn-in. MCMC chains were thinned by 10 to accommodate the computational intensity of tracking each parameter estimate. To verify model convergence, we visually inspected the MCMC chains to ensure proper mixing and ensured that all Gelman-Rubin diagnostics for each parameter were < 1.10 (Gelman et al., 2014).

E. Results

Over a total of 20,025 traps nights between spring 2011 and fall 2013, coyote were photographed on 1,195 days, red fox 146 days, striped skunk 357 days, raccoon 2,348 days, and opossum 1,877 days. With these data, the periodic time model had the best fit for coyote, red fox, and striped skunk, the stochastic time model performed best for raccoon, and the full model that included a periodic and stochastic component had the best fit for opossum (Table
For each species, we discuss the colonization, persistence, and detection results from the best fit model in the sections below.

\textbf{TABLE V}

\textbf{MODEL SELECTION RESULTS OF THE FOUR DIFFERENT MODELS THAT WE FIT TO 9 SEASONS OF CAMERA TRAP DATA FROM CHICAGO, ILLINOIS. MODELS WERE COMPARED WITH THE SUMMARY STATISTIC $-\sum \log(CPO_{k,t})$, AND THE LOWEST VALUES INDICATE THE MODEL WITH THE BEST WITHIN-SAMPLE PREDICTIVE PERFORMANCE. THE VALUE OF THE BEST FIT MODEL FOR EACH SPECIES IS BOLDED}

<table>
<thead>
<tr>
<th>Model</th>
<th>y covariates</th>
<th>Coyote</th>
<th>Red fox</th>
<th>Skunk</th>
<th>Raccoon</th>
<th>Opossum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full</td>
<td>URB + session + periodic</td>
<td>1021.73</td>
<td>215.55</td>
<td>438.14</td>
<td>1731.24</td>
<td>\textbf{1332.20}</td>
</tr>
<tr>
<td>Stochastic time</td>
<td>URB + session</td>
<td>1023.44</td>
<td>220.20</td>
<td>442.96</td>
<td>\textbf{1729.25}</td>
<td>1342.44</td>
</tr>
<tr>
<td>Periodic time</td>
<td>URB + periodic</td>
<td>\textbf{1020.74}</td>
<td>\textbf{213.59}</td>
<td>\textbf{437.35}</td>
<td>1736.86</td>
<td>1334.51</td>
</tr>
<tr>
<td>Homogeneous time</td>
<td>URB</td>
<td>1028.12</td>
<td>221.27</td>
<td>442.48</td>
<td>1731.97</td>
<td>1340.34</td>
</tr>
</tbody>
</table>

Of the five species analyzed, raccoon occupied the most sites at each time step, followed by the coyote, opossum, striped skunk, and then the red fox (Fig. 8). Red fox were estimated never to be present across all time steps at 69.47% of sites. Striped skunk, though they had similarly low overall occupancy rates, were estimated never to be present at 42.11% of sites. Raccoon, the most common species, were never observed at 5.26% of sites, while coyote and opossum were respectively never estimated to be present at 16.84% and 18.95% of sites across all time steps.
Figure 9. Box and whisker plots of the proportion of sites occupied by each species over time estimated by its best fit model from nine seasons of camera trap data collected from spring 2011 to summer 2013 in Chicago, Illinois. The occupancy status of each species at each site was collected by taking the median estimate of the posterior distribution of each species incidence matrix, $Z$, at each site and time step.

1. Colonization

The probability that colonization rates peaked in the fall for coyote, red fox, and striped skunk was respectively 95.89%, 97.03%, and 85.22%, while the probability that opossum colonization rates peaked in the spring and fall was 97.97%. On average, coyote, red fox, and striped skunk colonization rates respectively increased 2.2, 3.23, and 1.75 times in the fall (Fig. 9). Opossum colonization rates increased 1.76 times in the spring and fall, but this increase fluctuated some by a standard deviation of 0.17 (95% CI [0.01 – 0.89]) due to the varying intercept of its best fit
model (Fig. 9). Raccoon colonization rates did not vary in a periodic fashion but did vary through time by a standard deviation of 0.35 (95% CI [0.01 – 1.79]). At 62.42%, raccoons had the highest average colonization rate of all species ($m_{0, \text{raccoon}} = 0.51$, 95% CI [-0.19 – 1.51]). URB positively influenced the colonization rate of coyotes ($m_{\text{URB, coyote}} = 0.99$, 95% CI [0.31 – 1.62]), striped skunk ($m_{\text{URB, skunk}} = 1.14$, 95% CI [0.63 – 1.70]), opossum ($m_{\text{URB, opossum}} = 0.49$, 95% CI [0.05 – 0.94]), and raccoon ($m_{\text{URB, raccoon}} = 0.92$, 95% CI [0.46 – 1.46]), but did not have a definite positive or negative influence on red fox colonization rates ($m_{\text{URB, fox}} = 0.34$, 95% CI [-0.31 – 0.90]). Using Eq. 4.24, the proportion of temporal variability explained by a single-season pulse for coyote, red fox, and striped skunk was respectively 66%, 73%, and 30%, while the boom-bust sinusoidal trend accounted for 50% of the temporal variability in opossum colonization rates. Including the sinusoidal trend in the raccoon models actually increased the amount of temporal variability by 13% as the varying intercept in the full model drifted more to correct for unnecessary periodic trends.
Figure 10. The probability of colonization over time estimated from the parameters of the best fit model for A) coyote, B) red fox, C) striped skunk, D) raccoon, and E) opossum with 9 seasons of camera trap data collected from spring 2011 to summer 2013 in Chicago, Illinois. The black horizontal line is the median probability of colonization through time, while the gray bars represent the 95% credible interval obtained from the posterior distribution of the species incidence matrix. The black dots are the median estimate of the proportion of sites colonized by each species at each time step from the model, the vertical lines are the 95% credible interval associated to each estimate, and the white dots are the proportion of sites colonized by each species calculated the raw detection data. White dots that are not visible lie directly beneath the model estimate.
2. Persistence

Coyotes had both the highest average persistence rate of the species analyzed and the lowest amount of temporal variance in their persistence rates through time (Table VI). Red fox and striped skunk had similar persistence rates that were lower than coyote, opossum, and raccoon. Persistence rates varied more through time for the red fox than all other species, but given their rarity the precision of this estimate was the lowest of all species (Table VI). On average, the probability of persistence was highest for coyote (82.64% [95% CI 75.40 – 88.39]), followed by raccoon (79.74% [95% CI 71.50 – 86.99]), opossum (74.46% [95% CI 65.25 – 82.92]), red fox (63.41% [95% CI 25.54 – 92.76]), and striped skunk (60.35% [95% CI 42.31 – 77.56]). Raccoon persistence rates increased with URB, but 95% credible intervals for $d_{URB}$ bounded zero for all other species (Table VI).
TABLE VI
PERSISTENCE PARAMETERS FROM THE BEST FIT COYOTE, RED FOX, STRIPED SKUNK, OPOSSUM, AND RACCOON MODELS ESTIMATED FROM 9 SEASONS OF CAMERA TRAP DATA COLLECTED FROM SPRING 2011 TO SUMMER 2013 IN CHICAGO, ILLINOIS. THE TABLE INCLUDES ESTIMATES OF THE INTERCEPT ($d_0$), THE INFLUENCE THAT URB HAS ON EACH SPECIES PERSISTENCE RATES ($d_{URB}$), AND THE STANDARD DEVIATION OF HOW MUCH $d_0$ VARIES OVER TIME STEP ($\sigma_{time}$). PARAMETERS ARE ON THE LOGIT-SCALE AND ESTIMATE THE PROBABILITY THAT EACH SPECIES PERSISTS AT A SITE FROM ONE TIME STEP TO THE NEXT. THE 95% CREDIBLE INTERVAL FOR EACH ESTIMATE IS LOCATED WITHIN THE SQUARE BRACKETS NEXT TO THE PARAMETER.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$d_0$</th>
<th>$d_{URB}$</th>
<th>$\sigma_{time}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coyote</td>
<td>1.56 [1.12 – 2.03]</td>
<td>0.11 [-0.15 – 0.40]</td>
<td>0.11 [0.01 – 0.53]</td>
</tr>
<tr>
<td>Red fox</td>
<td>0.55 [-1.07 – 2.55]</td>
<td>-0.02 [-0.93 – 1.00]</td>
<td>1.00 [0.02 – 6.11]</td>
</tr>
<tr>
<td>Striped skunk</td>
<td>0.42 [-0.31 – 1.24]</td>
<td>-0.09 [-0.70 – 0.54]</td>
<td>0.21 [0.01 – 1.32]</td>
</tr>
<tr>
<td>Opossum</td>
<td>1.07 [0.63 – 1.58]</td>
<td>0.17 [-0.12 – 0.46]</td>
<td>0.55 [0.42 – 0.68]</td>
</tr>
<tr>
<td>Raccoon</td>
<td>1.37 [0.92 – 1.90]</td>
<td>0.52 [0.17 – 0.85]</td>
<td>0.23 [0.01 – 0.90]</td>
</tr>
</tbody>
</table>

3. Detection

The daily probability of detecting each species was low and varied more across sites than it did across time (Table VII). Coyote, opossum, and raccoon were more detectable in sites with higher levels of URB (Table VII). Raccoon had the lowest temporal variability in their detection rate, while opossum detection rates varied most across sites (Table VII). The average daily probability of detecting a species given its presence was 5.36% for coyote [95% CI 3.44 – 8.34], 2.00% for red fox [95% CI 0.30 – 6.50], 4.43% for striped skunk [95% CI 1.65 – 9.20], 7.42% for opossum [95% CI 4.03 – 12.74], and 7.92% for raccoon [95% CI 5.44 – 11.45].
**TABLE VII**

DETECTION PARAMETERS FROM THE BEST FIT COYOTE, RED FOX, STRIPED SKUNK, OPOSSUM, AND RACCOON MODELS ESTIMATED FROM 9 SEASONS OF CAMERA TRAP DATA COLLECTED FROM SPRING 2011 TO SUMMER 2013 IN CHICAGO, ILLINOIS. THE TABLE INCLUDES ESTIMATES OF THE INTERCEPT ($f_0$), THE INFLUENCE THAT URB HAS ON EACH SPECIES DETECTION RATES ($f_{URB}$), AND THE STANDARD DEVIATION OF HOW MUCH $f_0$ VARIES OVER TIME ($\sigma_{time}$) AND ACROSS SITES ($\sigma_{site}$). PARAMETERS ARE ON THE LOGIT-SCALE AND ESTIMATE THE PROBABILITY OF DETECTING A SPECIES PER DAY GIVEN ITS PRESENCE. THE 95% CREDIBLE INTERVAL FOR EACH ESTIMATE IS LOCATED WITHIN THE SQUARE BRACKETS NEXT TO THE PARAMETER.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$f_0$</th>
<th>$f_{URB}$</th>
<th>$\sigma_{time}$</th>
<th>$\sigma_{site}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coyote</td>
<td>-2.87 [-3.33 – -2.39]</td>
<td>0.33 [0.12 – 0.54]</td>
<td>0.41 [0.26 – 0.75]</td>
<td>1.04 [0.80 – 1.35]</td>
</tr>
<tr>
<td>Red fox</td>
<td>-3.90 [-5.70 – -2.67]</td>
<td>0.30 [-0.29 – 1.02]</td>
<td>0.57 [0.11 – 1.35]</td>
<td>1.14 [0.44 – 2.18]</td>
</tr>
<tr>
<td>Striped skunk</td>
<td>-3.07 [-4.09 – -2.29]</td>
<td>-0.40 [-0.90 – 0.21]</td>
<td>0.52 [0.30 – 1.00]</td>
<td>1.26 [0.85 – 1.81]</td>
</tr>
<tr>
<td>Opossum</td>
<td>-2.52 [-3.17 – -1.92]</td>
<td>0.38 [0.05 – 0.76]</td>
<td>0.57 [0.37 – 1.00]</td>
<td>1.37 [1.05 – 1.82]</td>
</tr>
<tr>
<td>Raccoon</td>
<td>-2.45 [-2.86 – -2.06]</td>
<td>0.63 [0.40 – 0.87]</td>
<td>0.29 [0.18 – 0.52]</td>
<td>1.10 [0.89 – 1.39]</td>
</tr>
</tbody>
</table>

Given their general rarity (Fig. 8) and the fact that they were never detected in >40% of sites, site-specific detection estimates for red fox and striped skunk tended toward their mean detection rate and were less precise than the other species (Fig. 10). As species became more common, estimates of site-specific detection probabilities and mean detection rates increased in precision.
Figure 11. Site-specific detection probabilities estimated by the best fit model for A) coyote, B) red fox, C) striped skunk, D) raccoon, and E) opossum from 9 seasons of camera trap data collected from spring 2011 to summer 2013 in Chicago, Illinois. The black horizontal line is the median estimate of the mean detection rate for each species \( (m_0) \), the dark gray bar is the 95% credible interval for \( m_0 \), and the light gray bar is the 95% posterior predictive distribution of site-specific detection rates calculated from each species \( m_0 \) and the site-level standard deviation hyperparameter \( (\sigma_{site}) \). Along the x-axis are site-specific detection estimates on the logit scale. The black dots are median estimates for each site while the vertical lines are 95% credible intervals. All sub-figures A-E are sorted by site-specific detection estimates for coyote.
F. Discussion

We have presented here a highly versatile method that incorporates Fourier series into dynamic occupancy models in order to partition variability between periodic trends and deviations from those trends. Furthermore, we do so with types of data that are typically available from most ecological surveys. Models with a periodic component best predicted colonization rates in 4 of the 5 species we analyzed (Table V) and, depending on the species, accounted for between 30 – 73% of the overall temporal variability in their colonization rates. Such models provide a way to explicitly specify periodic trends in ecological time-series that are both easier to interpret, make future predictions with, and uses few parameters. For example, we should expect coyote colonization rates to roughly double in the Fall around Chicago, from 25.97% in the ‘trough’ seasons to 57.56%. Furthermore, these rates will increase in areas with greater canopy cover and decrease in locations with more impervious surfaces and/or human dwellings. Most importantly, this method makes it easier to incorporate prior knowledge on the natural history of a species into a statistical model and represents an explicit and biologically reasonable formulation regarding the rates that determine a species spatiotemporal distribution in temporally heterogeneous environments.

While we only altered the shape of the periodic trend given the particular breeding history of a species, prior distributions on $A$ and $\delta$ could also be adjusted to better reflect prior knowledge in how large the amplitude ($A$) of a pulse is or when it occurs ($\delta$). For example, as juvenile coyote, red fox, and striped skunk primarily disperse in the fall (Bekoff and Gese, 2003; Cypher, 2003; Rosatte and Lariviére, 2003), the categorical prior on $\delta$ could be altered to better
acknowledge this information by putting a higher prior probability on fall dispersal. Although the use of strong or informed priors tends to be avoided in ecology, they can be exceptionally useful to help constrain model parameters, account for multi-collinearity, and incorporate prior knowledge into a model (Hooten and Hobbs, 2015). Moreover, while noninformative priors (e.g., large uniform priors) are often employed in applied textbooks on the Bayesian analysis of ecological data (e.g., Royle and Dorazio, 2008; Kéry and Schaub, 2012; Kéry and Royle, 2016), their use is often improper in many cases (Gelman, 2006; Van Dongen, 2006; Seaman et al., 2012). As Bayesian analyses in ecology continue to increase in popularity (Hooten and Hobbs, 2015), more care should be taken in the prior specification of a model to better reflect biologically reasonable bounds of parameter estimates. The benefit of this is multi-faceted, as it often leads to increases in the algorithmic stability of MCMC samplers and encourages ecologists to more explicitly consider the natural history of species in their analyses.

Our results provide support to the notion that Fourier series can be used and parameterized in dynamic occupancy models, but our approach also has implications for ecological theory and empirical tests of it. For instance, metacommunity theory naturally focuses on the ecological consequences that spatial heterogeneity, competition, and dispersal have on community composition and structure in fragmented habitats (Holyoak et al., 2005 and references therein). Often, the mathematical models used to explore the four paradigms of metacommunity theory (i.e., patch dynamics, species sorting, mass effects, and neutral theory, see Leibold et al., 2004 for an overview) make the necessary simplification that dispersal is constant and happens at the same time for all species. A notable exception is Matias et al.’s (2012) temporally stochastic model for source-sink environments. As species with varying life
history strategies breed and disperse at different times, empirical data gathered to test such
theories are likely a function of when they are collected throughout the year and as such may
be more variable than theoretical predictions supplied by a metacommunity model. Thus, to
test such models with species that have evolved in temporally heterogeneous environments
there is a need to more fully connect current metacommunity theory to the evolutionary
context and life history strategies of species to explore how space, time, and their interaction
influence community dynamics (Holt et al., 2005).

The spatial covariate that we calculated and used in these models, URB, had a positive
effect on the colonization rate of all species except the red fox, the species with the least
amount of data. As most mammals in urban environments must traverse through the urban
matrix to locate suitable habitat patches, it is naturally intuitive that the spatial composition
surrounding sites can influence the effective isolation between them and therefore their
probability of colonization (Ricketts, 2001). In this case, species were less likely to colonize sites
that had more human dwellings and impervious surfaces around them, but more likely to
colonize sites with greater levels of canopy cover. Conversely, URB had little to no effect on all
species persistence rates save for the raccoon (Table VI). While it is possible that there are
other spatial forces that influence these species, factors at much smaller scales could modify
their probability of persistence. For instance, the ability of each species to locate appropriate
den sites or food sources likely occurs as at a smaller scale than many of the environmental
covariates collected by geographical information systems software. Although this analysis was
primarily conducted to assess the ability of Fourier series to predict colonization rates, it also
indicates that the habitat matrix around sites in Chicago likely influences wildlife colonization rates, but other factors influence their ability to persist once they are there.

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 more explicitly captures periodic patterns and illustrate its use with types of data that are typically available from wildlife surveys. Models that incorporated periodic colonization trends best predicted the dynamics of four of the five species in our analysis, and were consistent with readily available information on their natural history. While we only explored the implications of two simple periodic formulations within dynamic occupancy models, such patterns likely apply to many other species. Further, Fourier series are highly generalizable and can be used to fit any kind of periodic trend. Our results show the potential of using Fourier series in dynamic occupancy models to robustly estimate periodic trends and makes it easier for ecologists 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.

**CITED LITERATURE**


V. ASSESSING VALUES AND PERCEPTIONS OF WILDLIFE THROUGH SOCIAL MEDIA

A. Abstract

As a global repository of facts and opinions, the Internet may track public attitudes towards a variety of topics. While it is commonplace to use online data in other fields, relatively few studies have used such information to assess public attitudes towards the natural world. Here, we quantitatively compare comments made by users on YouTube, one of the most popular social media websites in the world, on the most viewed wildlife videos of three common North American mammalian species: coyote (*Canis latrans*), Virginia opossum (*Didelphis virginiana*), and raccoon (*Procyon lotor*). In general, coyote were viewed negatively, whereas raccoon and opossum were primarily viewed as ‘cute’ or as potential pets. The technique we developed to analyze these comments could be used by researchers to assess any number of questions regarding online viewpoints or perceptions and has implications regarding how scientists communicate information online.

B. Background

Urban areas are expanding worldwide, and although urban environments show decreased biodiversity (Aronson et al., 2014), many species have adapted to human-dominated habitats. As wildlife, humans, and their pets encounter each other, conflicts may occur which require action from land managers. Because these management actions are increasingly under the scrutiny of a diverse constituency, public awareness of and attitudes toward wildlife inherently influence wildlife management practices (Decker and Enck, 1996; Adams and Lindsey, 2010). For example, negative public sentiment and increased hunting of coyotes (*Canis*
The presence of American invasive species (latrans) in the southeastern United States has proven to be a major obstacle in the recovery of the red wolf (Canis rufus), one of North America’s most endangered mammals (Hinton et al., 2013). Conversely, increased public education and positive public attitudes can significantly benefit species (Adams and Lindsey, 2010). The city of Austin, Texas researches, celebrates, and facilitates the continued growth of the largest urban bat colony of Mexican free-tailed bats (Tadarida brasiliensis) in North America (Keeley and Tuttle, 1999; Adams and Lindsey, 2010). Today, this colony of bats is a major tourist attraction and generates $8 million in tourism revenue each year (Adams and Lindsey, 2010). Public attitudes towards wildlife thus strongly influence how species are managed.

At the same time, the exponential growth of information at our fingertips alters the way people learn and relate to the world around them and thus influences public attitudes (Dahlgren, 2005; Brundidge, 2010; O’Conner et al., 2010). For instance, the Google search engine processes over 40 thousand queries every second worldwide (Internet Live Stats, 2016a). Furthermore, over 40% of the world’s population has Internet access (Internet Live Stats, 2016b). The immense popularity and convenience of the Internet have primed it to be the primary source of public information. For example, the majority of the adult population in the United States used Internet sources to get their political news during the 2008 presidential election (Pew Research, 2009), and nearly 60% of the U.S. population use the Internet as a primary source of information to learn about scientific issues (National Science Board, 2012). Scientists must therefore understand and engage with this new media landscape that exists online (Brossand and Scheufele, 2013).
Aside from increased access to mainstream media, user-generated content and online forums (i.e., social media) allow people to interact with and share ideas or content with others based on their own identities, ideologies, interests, and morals (Boyd et al., 2007). As online forums alleviate the restrictions of distance and time, computer-mediated communication can be beneficial. As one example, social media forums have been shown to provide emotional support to patients with Huntington’s Disease (Coulson et al., 2007). Most importantly, social media has also proven successful in altering public perception (Voinov and Costanza, 1999; Mason and Rennieb, 2007; Mwambui, 2010). For example, while the initial comments made on a viral video of a slow loris (Lorisidae, *Nycticebus*) primarily focused on people wanting slow lorises as pets, later comments increasingly discussed the negative influence that the illegal pet trade has had on their wild populations and illustrated an increased knowledge of their ecology and conservation status (Nekaris et al., 2013). As such, it is critical for our conservation and wildlife management efforts to better understand what is being said online about the natural world (Burgess and Green, 2009; Castells, 2011).

Online data are commonly used in economics (Ockenfels and Roth, 2006) and epidemiology (Cook et al., 2011; Richiardi et al., 2014), but relatively few studies have used such information to assess public attitudes towards the natural world (Proulx et al., 2014). However, the few studies that do reveal striking trends in the publics’ awareness of different conservation issues (McCallum and Bury, 2013; Proulx et al., 2014), threatened species (Nekaris et al., 2013), and the varying attitudes of hunters and non-hunters (El Bizri et al., 2015). Yet, these studies either focus on abstract concepts (e.g., ecosystem services) or on rare and exotic
species. Little is known regarding online perceptions towards common species that people are more likely to encounter in their daily lives.

We studied online perceptions of three widespread North American mammals commonly found near people: the coyote, Virginia opossum (*Didelphis virginiana*, hereafter opossum), and raccoon (*Procyon lotor*). All three of these species are able to persist at high population densities in human-dominated landscapes due in large part to the presence of anthropogenic food, water, and denning sites (Gehrt et al., 2010). As such, both direct (e.g., attacks; Poessel et al., 2013) and indirect (e.g., zoonotic disease; Bradley and Altizer, 2007) interactions occur between these species and people, their pets, and their livestock, making them ideal study candidates to assess public perceptions of wildlife online. To understand public perceptions regarding these species, we ask two specific research questions: 1) what are the most commonly held attitudes toward each species online? And 2), do the types of attitudes held towards coyotes, opossum, and raccoon differ?

C. **Materials and Methods**

1. **Collecting and analyzing comments on social media**

To assess public sentiment towards coyote, raccoon, and opossum we collected and analyzed the online comments of the top ten most viewed YouTube videos for each species. YouTube is a video-sharing website that easily allows and encourages users to upload, share, rate, and comment on user-generated content (Freeman et al., 2007; Heckner and Wolf, 2009). YouTube is also one of the most popular social media websites in the world with nearly 1 billion active users each month (YouTube, 2016). To locate these videos, we used the YouTube search
feature and searched for videos about each species, filtering by number of views. After verifying that each video contained the target species, we recorded the upload date, number of views, and number of comments made. We then accessed the mobile YouTube website and programmatically collected all comments made about each video by iteratively using Asynchronous JavaScript (AJAX) Language calls to retrieve 10 comments at a time (Nekaris et al., 2013). We collected the name of the user who made the comment, the date the comment was made, and the comment itself. The mobile YouTube site was used because the comments had a formatting structure that was easier to access but retained the same information as the desktop site. Following this, the comments for each video were stored as a JavaScript Object Notation (JSON) stream. In order to give individual users one ‘vote’ per video, we only saved their first comment. The vast majority of YouTube users only make one comment per video (Nekaris et al., 2013), so we do not expect this to bias our results.

At the time of our analysis the mobile YouTube website would not correctly load if a video had > 20 thousand comments. Because we could not guarantee a representative sample of the comments for these videos, we had to exempt videos with > 20 thousand comments. As only two videos had this many comments (a video of a man defending his dog from a raccoon titled “My dog Toaster was attacked by a raccoon” and a video of a coyote biting on a man’s boot in the snow titled “Coyote Attack Best Footage Ever”), we simply moved to the next most popular video and collected its comments until we had a total of 10 videos per species (see Appendix D for a list of the videos analyzed).

2. **Categorizing videos and comments**
Videos were placed into one of five different categories based on their content, with respect to the species of interest in each video (Table VII). Categories were created after viewing the videos and cover the full range of activities observed.

**TABLE VIII**
CATEGORIES OF THE VIDEO CONTENT IN MOST VIEWED COYOTE, OPOSSUM, AND RACCOON VIDEOS ON YOUTUBE

<table>
<thead>
<tr>
<th>Video category</th>
<th>The target species in the video is...</th>
</tr>
</thead>
<tbody>
<tr>
<td>As a pet or in captivity</td>
<td>either a pet or wild individual but in captivity.</td>
</tr>
<tr>
<td>Actively feeding wildlife</td>
<td>wild and actively fed by people.</td>
</tr>
<tr>
<td>Human-wildlife conflict</td>
<td>either directly or indirectly coming into conflict with people or their domestic pets.</td>
</tr>
<tr>
<td>Hunting</td>
<td>being hunted by people.</td>
</tr>
<tr>
<td>Wildlife interactions</td>
<td>interacting with another wild species in a natural environment.</td>
</tr>
</tbody>
</table>

For each of the ten videos per species, comments were categorized based on their content. To ensure that these categories represented distinct attitudes towards wildlife, we used categories defined by Kellert which describe contrasting attitudes toward wildlife (Table IX; Kellert, 1980). However, some attitudes described by Kellert (1980) were both rare and difficult to distinguish from a single comment (utilitarian and aesthetic). Due to this, we combined dominionistic and utilitarian attitudes into a single category as well as naturalistic and aesthetic. Thus, we considered seven of the nine Kellert categories.
TABLE IX
KELLERT CATEGORIES USED IN THIS ANALYSIS TO ASSESS ONLINE ATTITUDES TOWARDS COYOTE, OPOSSUM, AND RACCOON. DEFINITIONS OF KELLERT CATEGORIES WERE TAKEN FROM KELLERT (1980), WHILE COMMENTS WERE TAKEN FROM DIFFERENT VIDEOS THAT WERE COLLECTED FOR THIS ANALYSIS.

<table>
<thead>
<tr>
<th>Kellert category</th>
<th>Definition</th>
<th>Examples from YouTube comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Naturalistic</td>
<td>Primary interest and affection for wildlife and the outdoors</td>
<td>“I love all coyotes”</td>
</tr>
<tr>
<td>Ecologistic</td>
<td>Primary concern for the environment as a system, for interrelationships between wildlife species and natural habitats</td>
<td>“Just like every animal...they are vital to keep this world working properly. These little dudes eat animals we class as 'pests' like rats, mice, bugs etc etc”</td>
</tr>
<tr>
<td>Scientific</td>
<td>Primary interest in the physical attributes and biological functioning of animals.</td>
<td>“Opossums are extremely resistant to zoonotic diseases.”</td>
</tr>
<tr>
<td>Humanistic</td>
<td>Primary interest and strong affection for individual animals, principally pets. Regarding wildlife, focus on large attractive animals with strong anthropomorphic associations.</td>
<td>“I wish I had that baby possum to raise. I love possums and I would love another baby possum to raise. They make lovely affectionate pets. The one I had was named Ziggy Piggy.”</td>
</tr>
<tr>
<td>Moralistic</td>
<td>Primary concern for the right and wrong treatment of animals, with strong opposition to exploitation or cruelty towards animals.</td>
<td>“Why did u kill that?? it was a poor creature.....”</td>
</tr>
<tr>
<td>Dominionistic</td>
<td>Primary interest in the mastery and control of animals, typically in sporting situations</td>
<td>“Destroy all coyote who ruin deer hunts”</td>
</tr>
<tr>
<td>Negativistic</td>
<td>Primary orientation an avoidance of animals due to indifference, dislike, or fear.</td>
<td>“If I ever saw a possum, I would freak the hell out!! Those things are gross and disgusting.”</td>
</tr>
</tbody>
</table>
To decrease the number of irrelevant comments (e.g., job offers posted by “bots” or links to other YouTube pages), we programmatically searched through the comments for a list of 64 common keywords that represent these Kellert categories and conversations related to these species (Appendix E). Keywords were selected both by 1) looking through a random subset of comments for each species across videos and choosing words that were attributable to a Kellert category, and 2) using our professional opinion to include additional words related to the ecology, conservation, and management of these three species. We used approximate string matching with the ‘agrep’ function in program R to locate comments containing keywords and variant spellings or misspellings of keywords (R Core Team, 2016).

We then categorized each of the comments selected through this process with the exception of comments that were difficult to interpret or irrelevant. For example, the comment “it’s so cute” would not be categorized as humanistic if there were multiple species present in the video whereas “that raccoon is so cute” would be categorized for a raccoon video even if multiple species were present. Comments could also be placed into multiple categories if they covered varying topics. The comment “opossum are the only marsupial native to North America, can play dead, and are commonly found in urban environments” would fit into the ‘scientific’ and ‘ecologicistic’ categories because the comment focuses on the species biology (being a marsupial that can play ‘dead’) as well as its ecology (found in urban environments). Finally, the accuracy of statements made in a comment was not considered when assigning them to Kellert categories. After classifying all comments associated with a single video, the number of comments from each category was summed and the total number of comments categorized was collected.
3. **Analysis of comments**

To estimate the proportion of comments that fit a particular attitude towards coyotes, opossum, and raccoons we fit a multinomial model to these data for each species. We used a Bayesian framework to parameterize the model, and vague priors were used as we had no prior information to include (Gelman et al. 2014). Our posterior distribution for the \( j \) videos, \( k \) species, and \( i \) Kellert categories is:

\[
[\theta|N] \propto \prod_{j=1}^{10} \prod_{k=1}^{3} \text{Multinomial}(N_{j,k}|\theta_{1,j,k} ... \theta_{I,j,k}) \text{Dirichlet}(\theta_{1,j,k} ... \theta_{I,j,k}|a_{1,j,k} ... a_{I,j,k})
\]

where \( \theta_{i,j,k} \) is the estimated proportion of comments for each category per video per species and \( N_{j,k} \) is the number of comments in each category per video per species. As we had no prior information to include within the model we used a vague Dirichlet prior centered on a uniform multinomial distribution. To do so, all alpha values \( (a_{1,j,k} ... a_{I,j,k}) \) for the Dirichlet prior were set to \( 1/7 \) (~0.14) for the 7 Kellert categories (i.e., we assumed an equal prior probability for all comment types in a video). Given this specification, our model estimates the proportion of comments in each category per species per video. However, in order to compare and contrast the most common types of comments made about each species, it is of greater interest to calculate the global proportion of each comment type over all videos for each species. Following Gelman et al. (2014), we calculate these values with the posterior simulations of the above model and the known values of \( N_{j,k}/N_k \):
\[
\theta_{i,k} = \sum_{j=1}^{10} \frac{N_{j,k}}{N_k} \theta_{i,j,k}
\]

Where \(\theta_{i,j,k}\) and \(N_{j,k}\) are the same as above, \(N_k\) is the total number of comments with a preference, and \(\theta_{i,k}\) is the proportion of comments made in the \(i\)th category for the \(k\)th species. Furthermore, another estimate of interest is the proportional difference of comments in category \(i\) between species, which can readily be calculated by subtracting \(\theta_{i,k}\) values between species from the equation above. For example, to estimate the proportional difference in comments of category \(j\) between raccoon and opossum one can simply subtract \(\theta_{i,\text{raccoon}}\) from \(\theta_{i,\text{opossum}}\). To fit this model and estimate all relevant quantities we used JAGS version 4.0.0 (Plummer, 2003) through the runjags package (Denwood, 2016) on version 3.2.3 of R (R core team 2016). Following a 20,000 step adaptation and a 20,000 step burn-in, the posterior distribution was sampled 60,000 times. To ensure model convergence we calculated Gelman-Rubin diagnostics and visually inspected the conditional posteriors for proper mixing (Gelman et al., 2014).

D. Results

The average number of views for videos was 1,680,000 for coyote (sd =600,000), 960,000 for opossum (sd =770,000), and 1,750,000 for raccoons (sd = 610,000). Overall, the ten most viewed coyote, opossum, and raccoon videos were respectively seen > 16 million, 17 million, and 9 million times. For coyote, 80% of their top viewed videos were of hunting and 20% were of interference competition between coyotes and wolves. On the other hand, 70% of the most popular opossum videos focused on pet or wildlife opossum in captivity and 30% on
human-wildlife conflict. Finally, 50% of raccoon videos were of human-wildlife conflict, 40% featured pet or wild individuals in captivity, and 10% included people actively feeding raccoons.

In total, 19,298, 13,636, and 19,748 comments were collected from coyote, opossum, and raccoon videos, respectively on 31 October 2013. Filtering by keyword left us with 10,532, 5,058, and 7,772 comments for coyote, opossum, and raccoon of which 18.96%, 54.74%, and 24.22% were assignable to a Kellert category. For coyotes, dominionistic attitudes were most commonly recorded (49.49%, 95% CI = 47.24 – 51.77) and were 2.31 times (95% CI 2.22 – 2.39) more likely than ecologistic attitudes, the second most common category (Fig. 12). Moralistic attitudes, or those that expressed concern over the right and wrong treatment of animals, were almost as common as ecologistic attitudes (Fig. 12). Humanistic attitudes (i.e., comments that anthropomorphized the species and/or declared an interest in having it as a pet) were the rarest for coyotes (1.00%, 95% CI = 0.63 – 1.49; Fig. 12).

Conversely, humanistic attitudes were the most common for opossum (46.63%, 95% CI = 44.92 – 48.37; Fig. 12), followed by negativistic attitudes (18.76%, 95% CI = 17.36 – 20.22). Naturalistic attitudes (i.e., comments that proclaim a positive interest in the entire species, not just a particular individual), were the third most common and accounted for 12% of categorized comments (95% CI = 11.15 – 13.56). Ecologistic attitudes were the rarest Kellert category for opossum (Fig. 12).
Figure 12. The estimated proportion of comments that were assigned to Kellert categories from the 10 most viewed coyote (C), opossum (O), and raccoon (R) videos on YouTube. The left panel of the figure illustrates how attitudes differ within a species while the right panels show how species vary across attitudes. The area inside the black tails of each proportion represents the 95% credible interval per species per attitude.
Similar to opossum, humanistic attitudes were the most frequent Kellert category for raccoon and accounted for 34.35% (95% CI = 32.30 – 36.53) of comments in their videos (Fig. 12). However, unlike coyote and opossum, the spread between the first and second most frequent category was far smaller for raccoon, with humanistic attitudes only 1.39 (95% CI = 1.38 – 1.43) times as common as scientific (i.e. comments related to a species biology or physical attributes). Comparing between species, coyote had 4.87 (95% CI = 4.14 – 5.63) times more comments that signified dominionistic attitudes than opossum and raccoon (Fig. 12). For all species, humanistic, dominionistic, and negativistic attitudes were the most common categories across all videos and species and respectively accounted for 23.80% (95% CI = 23.01 – 24.61), 19.44% (95% CI = 18.64 – 20.26) and 14.56% (95% CI = 13.79 – 15.35) of comments. Across all videos and species, naturalistic, ecologistic, and scientific attitudes were the rarest Kellert categories and respectively accounted for 8.70 (95% CI = 8.10 – 9.32), 9.92% (95% CI = 9.26 – 10.61), and 11.36% (95% CI = 10.67 – 12.61) of all comments.

E. Discussion

In an increasingly urbanizing world, humans are bound to interact with the wildlife residing in towns and cities. To facilitate coexistence between humans and nature, it is crucial to characterize public attitudes towards wildlife and identify common trends or areas of misinformation. We used novel techniques to better understand online attitudes about three common urban-adapted species throughout North America (coyote, opossum, and raccoon). Our research provides a look into the public perception of these species as recorded on social media and highlights a number of similarities and differences between them (Fig. 12).
Based on our results, the coyote is clearly the pariah of the three. Over two-thirds of the comments made about coyote either offered a negative opinion, encouraged shooting it, or both (Fig. 12). Historically, predators like the coyote were emblematic of the wild that humans sought to control and were therefore viewed as a roadblock to human progress (Nash, 1967). Even online, this ‘American pioneer’ mentality toward species like the coyote is prevalent, and coyotes appear to be seen as competitors for wild game at best or pests at worst. However, when comments were not negative or dominionistic, they either focused on the coyotes’ ecological role within the environment or viewed harm to this species as immoral (Fig. 12).

Raccoon and opossum, on the other hand, were primarily seen as charismatic or as prospective pets (Fig. 12). While this may be primarily good news for these species’ ability to persist in human-dominated environments, there is some cause for concern about the prevalence of humanistic attitudes. Increased contact with humans and resource provisioning via supplemental feeding (a common topic in popular raccoon and opossum videos) can increase the spread of zoonotic diseases (Bradley and Altizer, 2007), bloat population densities (Prange et al. 2003), and alter where these species occur (Prange et al., 2003, Fidino et al., 2016), and as such increase the likelihood of human-wildlife conflict (Gehrt et al., 2010). It is therefore unsettling that many of the top most viewed videos and associated attitudes towards these species focused on feeding. However, our results demonstrate that a clear opportunity exists, if care is taken, to harness the charisma of these species to educate the public about their ecological role.

A goal of many wildlife professionals is to better educate people and foster an understanding of wildlife management, ecology, and the environment (Adams and Lindsey,
2010). However, the three Kellert categories most aligned with wildlife agencies goals (i.e.,
naturalistic, ecologistic, and scientific) represented the minority of comments made by people
online. As social media itself represents a new environment where scientists are interacting
with the public, there is a need to critically consider the Internet’s role in our educational and
outreach efforts (Brossard and Scheufele, 2013; Bombaci et al., 2015). As coyote, opossum, and
raccoon were primarily either anthropomorphized or hated, our results suggest a disconnect
between what is said about these species online and the inherent goals of science education.
People may be more likely to learn to coexist with these species if they knew that opossum may
decrease the spread of Lyme disease (Keesing et al., 2010) or coyotes may increase the diversity
of songbirds (Crooks and Soulé, 1999). While the more traditional route for online
dissemination is science journalism, we must also learn how to better interact with the public in
a new media landscape (Brossard and Scheufele, 2013).

There are important limitations to our work that should be carefully considered. Similar
to other online studies, the trends observed are not a truly random sample of the population,
but are based on a specific sample of Internet users. Regardless, the comments in our analysis
were made in response to the most popular online videos of these three species and seen
collectively over 42 million times. Therefore, these comments do provide a general depiction of
common attitudes of these three species. Moreover, there did appear to be some circularity
between the content of the video and the comments that are made. For example, 80% of the
top viewed coyote videos were of hunting, and almost half of the comments made about
coyote were dominionistic (Fig. 12). For raccoon and opossum, 55% of videos were of
individuals who were pets or wild individuals in captivity, and most comments made were
humanistic. Although these correlations should be noted, given the fact that these videos are in the 10 most viewed for these species, our results still track the publics’ perception of what they are likely to see and say about these three species on the Internet. Finally, while our study reveals a mean response in comment types, it does not assess if such responses vary geographically or between urban and rural users. As people’s viewpoints towards wildlife likely differ along an urbanization gradient (Kellert, 1984), further insight could be gained with analyses that account for geographical location.

In our study, we illustrate a novel approach to quantify attitudes towards different wildlife species on social media. While we employed Kellert’s categories of attitudes towards wildlife (Kellert, 1980), this approach is easily generalizable and can be used on different species or topics entirely. Furthermore, our approach could also be used as a complimentary resource to more traditional techniques, such as surveys, to quantify how much information people receive online, where they find this information, and how their views are shaped by these different sources of information.

For the first time ever, more people live in urban areas than anywhere else (United Nations, 2014), and the wildlife assemblages in cities and towns are subject to the whims of wildlife managers and urban design, both of which track public attitudes and perception. As people increasingly look to the Internet to learn about the world around them, their attitudes can be influenced by what is available to them online (Voinov and Costanza, 1999; Mason and Rennieb, 2007; Mwambui, 2010). Therefore, it behooves biologists and conservation scientists to understand what types of traditional or non-traditional information may be disseminated online (Brossard and Scheufele, 2013). Using comments from YouTube videos, we found that
Internet users had more dominionistic attitudes towards coyotes, but more humanistic attitudes towards opossums and raccoons. Naturalist, ecological, and scientific attitudes were the least common attitudes overall, suggesting gaps in knowledge that could be addressed with educational programs. Our results show the potential of using online resources to robustly estimate human perceptions towards wildlife and also can provide information on common attitudes to target education and management efforts for wildlife.

F. Acknowledgements

We thank K. A. I. Nekaris for discussing their methods to collect comments from YouTube and J. Frumkin for assistance with data entry. We also thank T. Gallo, A. Halloway, and J. Peplinski whose useful comments greatly improved earlier versions of this manuscript. Funding was provided by the Lincoln Park Zoo and the Abra Prentice Wilkin Foundation.

CITED LITERATURE


VI. APPENDICES

APPENDIX A

To model the sampling method of a camera trap survey, we used 10 seasons of camera trap data from 118 sites in Chicago, IL to determine the daily probability of a camera trap being active (see Vernon et al., 2015 for more details on camera trapping project). Although there was a possibility of 1180 surveys (10 seasons * 118 sites), not all sites were sampled each season due to staggered entry. Additionally, camera traps were at times vandalized, stolen, or malfunctioned, which would result in missing data. In total, we had 803 surveys from spring 2010 to summer 2012. For a 28-day sampling period the median number of days a camera trap was active was 27 (range 2 – 28). The distribution was also highly left-skewed, with over 75% of the camera traps being active ≥ 24 days (Fig. 13). We used these data to (1) calculate the daily probability of a camera trap being active, (2) bootstrapped 803 draws from the binomial distribution at these probabilities 1000 times, and (3) assessed the fit of this method by calculating root mean squared error (RMSE).

The average probability a camera trap was active per day was 0.85 (range 0.45 – 0.94). After determining the daily probability of a camera trap being active we simulated 803 draws from the binomial distribution 1000 times with 28 trails per draw, with each trial at its own daily probability. To further encourage a left-skew in our simulated data, we also tried sampling from the binomial distribution with 30 trials (instead of 28), using the average daily probability over all days (0.85) for the two additional days. For this secondary approach we then converted any draws > 28 to 28. A contingency table was made for each simulation,
marking the frequency of days that camera traps were active over the 803 draws. To assess fit, we calculated root mean squared error (RMSE) between each simulated contingency table and a contingency table for the observed data.

**Figure 13.** Distribution of the number of days a camera trap was active from 803 sampling events over 10 seasons between spring 2010 and summer 2012.
APPENDIX A (Continued)

Average RMSE for 28 trials was 64.43 (95% confidence intervals = 61.84 – 67.07) while average RMSE for 30 trials 29.60 (95% confidence intervals = 29.60 – 33.10). A visual inspection of the predicted fit of the 28 trial model shows that it does not adequately capture the left-skew of the observed data, while the 30 trial model performs better (Fig. 14). Therefore, we used the 30 trials instead of 28 to simulate the $J$ matrix.
Figure 14. Comparison of the predicted fits from the bootstrap simulations of 28 (A) and 30 (B) trials from the binomial distribution at daily probabilities calculated from 803 camera trap surveys in Chicago, IL. The black dots are the observed frequency of days a camera trap was active during these surveys, the gray line is the median frequency of days a camera trap was active from the bootstrap simulation, and the dashed lines are 95% confidence intervals.
APPENDIX B

For the co-occurrence model used in simulations, $M_{\text{inxS}}$, the posterior density for $i$ species, $k$ sites, and $t$ time steps is:

$$
[p, c, d, \Psi, \alpha, Z | Y] \propto \prod_{i=1}^{3} \prod_{k=1}^{100} \prod_{t=1}^{T} \text{binomial}(Y_{i,k,t} | j_{k,t}, Z_{i,k,t} \rho_{i}) \text{Bernoulli}(Z_{i,k,t=1} = 1 | \Psi_{i}) \\
* \text{Bernoulli}[Z_{i,k,t+1} | (1 - Z_{i,k,t})(c_{0,i}c_{i}x_{k}) + Z_{i,k,t}d_{0,i}a_{i}Z_{i,k,t}d_{i}x_{k}] \\
* \text{Beta}(\Psi_{i} | 1, 1) \text{Cauchy}(\rho_{i} | 0, 2.5) \text{Cauchy}(c_{0,i} | 0, 2.5) \text{Cauchy}(c_{i} | 0, 2.5) \\
* \text{Cauchy}(d_{0,i} | 0, 2.5) \text{Cauchy}(d_{i} | 0, 2.5) \text{Cauchy}(a_{i} | 0, 5)
$$

where $d_{0,i}$ are the species-specific intercepts for persistence and $a_{i}$ is the $i$th row of the impact matrix, $A$. The number of surveys a species is observed ($y_{i,k,t}$) is made with error based off a species probability of detection ($p_{i}$) and whether or not the species is estimated to be present ($z_{i,k,t} = 1$ = present, 0 = absent). At the first time step the probability a species occupies a site ($\Psi_{i}$) is treated as i.i.d. Bernoulli trails, whereas after the first time step a mixture model is used to determine the factors that influence a species probability of colonization ($c_{0,i}$) or persistence. Covariate effects ($c_{i}$ and $d_{i}$) are respectively included in the colonization and persistence portions of this model. The total number of seasons included in a model varied based off the particular simulation and was either 4, 6, or 8. The posterior distribution for the co-occurrence model that included the impact matrix on the probability of colonization instead of persistence was constructed similarly.

For the co-occurrence model fit to the raccoon, coyote, and opossum data the posterior density is:
APPENDIX B (Continued)

\[ [p, c, d, \Psi, \alpha, Z] \; \mathcal{N} \propto \prod_{i=1}^{3} \prod_{k=1}^{9} \prod_{t=1}^{13} \text{binomial}(Y_{ikt} \mid j_{kt}, Z_{ikt}, \rho) \; \text{Bernoulli}(Z_{ikt=1} \mid \Psi) \]

* Bernoulli \( [Z_{ikt+1} \mid (1 - Z_{ikt})(c_{i,0}c_{i,1}x_{k,1}c_{i,2}x_{k,2}) \]
+ \( Z_{ikt}(d_{0,i}\alpha_{r}Z_{r,k,t}d_{i,1}x_{k,1}d_{i,2}x_{k,2}) \) \( \beta(\Psi \mid 1,1) \)
* \( \text{Cauchy}(p_{i,t} \mid 0, 2.5) \; \text{Cauchy}(c_{i,0} \mid 0, 2.5) \; \text{Cauchy}(c_{i,1} \mid 0, 2.5) \)
* \( \text{Cauchy}(c_{i,2} \mid 0, 2.5) \; \text{Cauchy}(d_{i,0} \mid 0, 2.5) \; \text{Cauchy}(d_{i,1} \mid 0, 2.5) \)
* \( \text{Cauchy}(d_{i,t} \mid 0, 2.5) \; \text{Cauchy}(a \mid 0, 5) \)

Where all definitions are the same as above except for the inclusion of two environmental covariates (patch area = \( x_1 \) and distance to the city center = \( x_2 \)).
Here, we more fully describe the results of our simulation study and provide additional figures and tables to more explicitly illustrate the ability of our co-occurrence model to recover interspecific associations. We break down our 12 different sampling scenarios into four groups: 1) three-species horizontal competition where species influence each other’s persistence rates, 2) three-species horizontal competition where species influence each other’s colonization rates, 3) a predator with competing prey who influence each other’s persistence rates, and 4) a predator with competing prey who influence each other’s colonization rates. Within each group of scenarios, we provide results when different amounts of data are supplied to the model (i.e., 4, 6 and 8 seasons).

**Three-species horizontal competition via persistence**

Within this scenario, the accuracy and precision of all parameter estimates increased as more data were supplied (Table X). With 8 seasons of data, coverage was high for all parameters and ranged from 0.93 – 0.98. The average distance between the estimated interspecific effects and the true parameter values ranged from 0.29 – 0.66, 0.21 – 0.5, 0.18 – 0.43 when 4, 6, and 8 seasons of data are used, respectively (Table X). Larger interspecific effects (e.g., $\alpha_{3,1}$) were less accurate and precise than smaller interspecific effects (e.g., $\alpha_{2,3}$). Overall, these results suggest that our model was able to recover the true parameter estimates with reasonable accuracy and precision under this scenario.
## TABLE X

Mean absolute error (MAE), coverage, and precision of persistence (Φ), colonization (γ) and detection parameters estimated by the Co-occurrence Model for the horizontal competition module when species influenced each other’s persistence rates. Simulations were run 200 times with varying amounts of data (4, 6, and 8 seasons worth). The median estimate and 95% quantiles are listed for MAE and precision. Coverage was calculated as the proportion of simulations that had the true simulated parameter value within its 95% credible interval.

<table>
<thead>
<tr>
<th>Summary statistic</th>
<th>Process</th>
<th>Parameter</th>
<th>4</th>
<th>6</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE Φ</td>
<td>d₀,₁</td>
<td>0.25 (0.17 - 0.48)</td>
<td>0.21 (0.14 - 0.4)</td>
<td>0.18 (0.11 - 0.36)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>α₁,₂</td>
<td>0.36 (0.25 - 0.74)</td>
<td>0.3 (0.2 - 0.56)</td>
<td>0.25 (0.17 - 0.47)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>α₁,₃</td>
<td>0.29 (0.21 - 0.56)</td>
<td>0.24 (0.17 - 0.46)</td>
<td>0.22 (0.15 - 0.43)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>α₂,₁</td>
<td>0.52 (0.32 - 0.96)</td>
<td>0.37 (0.25 - 0.72)</td>
<td>0.32 (0.2 - 0.67)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>d₀,₂</td>
<td>0.35 (0.2 - 0.67)</td>
<td>0.25 (0.16 - 0.48)</td>
<td>0.22 (0.14 - 0.45)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>α₂,₃</td>
<td>0.44 (0.31 - 0.73)</td>
<td>0.36 (0.23 - 0.66)</td>
<td>0.3 (0.21 - 0.56)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>α₃,₁</td>
<td>0.66 (0.39 - 1.44)</td>
<td>0.5 (0.28 - 1.22)</td>
<td>0.43 (0.27 - 0.85)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>α₃,₂</td>
<td>0.61 (0.36 - 1.08)</td>
<td>0.46 (0.26 - 0.84)</td>
<td>0.39 (0.23 - 0.75)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>d₀,₃</td>
<td>0.35 (0.2 - 0.73)</td>
<td>0.27 (0.18 - 0.64)</td>
<td>0.24 (0.15 - 0.49)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>covariate₁</td>
<td>0.23 (0.14 - 0.59)</td>
<td>0.17 (0.1 - 0.38)</td>
<td>0.15 (0.08 - 0.33)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>covariate₂</td>
<td>0.25 (0.13 - 0.54)</td>
<td>0.2 (0.12 - 0.34)</td>
<td>0.16 (0.09 - 0.37)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>covariate₃</td>
<td>0.29 (0.16 - 0.64)</td>
<td>0.23 (0.12 - 0.53)</td>
<td>0.19 (0.1 - 0.38)</td>
<td></td>
</tr>
<tr>
<td>γ</td>
<td>c₀,₁</td>
<td>0.17 (0.12 - 0.37)</td>
<td>0.14 (0.09 - 0.3)</td>
<td>0.12 (0.08 - 0.26)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c₀,₂</td>
<td>0.18 (0.12 - 0.34)</td>
<td>0.13 (0.09 - 0.26)</td>
<td>0.11 (0.08 - 0.23)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c₀,₃</td>
<td>0.17 (0.12 - 0.32)</td>
<td>0.14 (0.09 - 0.33)</td>
<td>0.11 (0.07 - 0.23)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>covariate₁</td>
<td>0.19 (0.12 - 0.41)</td>
<td>0.15 (0.09 - 0.31)</td>
<td>0.14 (0.08 - 0.32)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>covariate₂</td>
<td>0.17 (0.11 - 0.4)</td>
<td>0.15 (0.09 - 0.34)</td>
<td>0.12 (0.07 - 0.3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>covariate₃</td>
<td>0.19 (0.11 - 0.39)</td>
<td>0.15 (0.09 - 0.34)</td>
<td>0.12 (0.07 - 0.24)</td>
<td></td>
</tr>
<tr>
<td>Detection</td>
<td>ρ₁</td>
<td>0.006 (0.004 - 0.012)</td>
<td>0.004 (0.003 - 0.009)</td>
<td>0.004 (0.002 - 0.008)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ρ₂</td>
<td>0.006 (0.004 - 0.011)</td>
<td>0.005 (0.003 - 0.01)</td>
<td>0.004 (0.003 - 0.008)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ρ₃</td>
<td>0.005 (0.004 - 0.012)</td>
<td>0.005 (0.003 - 0.011)</td>
<td>0.004 (0.003 - 0.009)</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX C (Continued)

TABLE XI (continued)
MEAN ABSOLUTE ERROR (MAE), COVERAGE, AND PRECISION OF PERSISTENCE (Φ), COLONIZATION (γ) AND DETECTION PARAMETERS ESTIMATED BY THE CO-OCCURRENCE MODEL FOR THE HORIZONTAL COMPETITION MODULE WHEN SPECIES INFLUENCED EACH OTHER’S PERSISTENCE RATES. SIMULATIONS WERE RUN 200 TIMES WITH VARYING AMOUNTS OF DATA (4, 6, AND 8 SEASONS WORTH). THE MEDIAN ESTIMATE AND 95% QUANTILES ARE LISTED FOR MAE AND PRECISION. COVERAGE WAS CALCULATED AS THE PROPORTION OF SIMULATIONS THAT HAD THE TRUE SIMULATED PARAMETER VALUE WITHIN ITS 95% CREDIBLE INTERVAL.

<table>
<thead>
<tr>
<th>Summary statistic</th>
<th>Process</th>
<th>Parameter</th>
<th>Number of seasons</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Coverage</td>
<td>Φ</td>
<td>d_{0,1}</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α_{1,2}</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α_{1,3}</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α_{2,1}</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d_{0,2}</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α_{2,3}</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α_{3,1}</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α_{3,2}</td>
<td>0.84</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d_{0,3}</td>
<td>0.90</td>
</tr>
<tr>
<td></td>
<td></td>
<td>covariate₁</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>covariate₂</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>covariate₃</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>γ</td>
<td>c_{0,1}</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c_{0,2}</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c_{0,3}</td>
<td>0.97</td>
</tr>
<tr>
<td></td>
<td></td>
<td>covariate₁</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>covariate₂</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>covariate₃</td>
<td>0.94</td>
</tr>
<tr>
<td>Detection</td>
<td>p₁</td>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>p₂</td>
<td>0.96</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>p₃</td>
<td>0.96</td>
<td>0.94</td>
</tr>
</tbody>
</table>
APPENDIX C (Continued)

TABLE XII (continued)
MEAN ABSOLUTE ERROR (MAE), COVERAGE, AND PRECISION OF PERSISTENCE (Φ), COLONIZATION (γ) AND DETECTION PARAMETERS ESTIMATED BY THE CO-OCCURRENCE MODEL FOR THE HORIZONTAL COMPETITION MODULE WHEN SPECIES INFLUENCED EACH OTHER’S PERSISTENCE RATES. SIMULATIONS WERE RUN 200 TIMES WITH VARYING AMOUNTS OF DATA (4, 6, AND 8 SEASONS WORTH). THE MEDIAN ESTIMATE AND 95% QUANTILES ARE LISTED FOR MAE AND PRECISION. COVERAGE WAS CALCULATED AS THE PROPORTION OF SIMULATIONS THAT HAD THE TRUE SIMULATED PARAMETER VALUE WITHIN ITS 95% CREDIBLE INTERVAL.

<table>
<thead>
<tr>
<th>Summary statistic</th>
<th>Process</th>
<th>Parameter</th>
<th>4</th>
<th>6</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>4.00 (3.84 - 1.36)</td>
<td>0.81 (0.66 - 1.06)</td>
<td>0.69 (0.54 - 0.92)</td>
</tr>
<tr>
<td>Precision Φ</td>
<td></td>
<td>d₀,1</td>
<td>1.36 (1.07 - 1.88)</td>
<td>1.09 (0.9 - 1.5)</td>
<td>0.96 (0.75 - 1.27)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α₁,₂</td>
<td>1.28 (1.05 - 1.64)</td>
<td>1.02 (0.86 - 1.35)</td>
<td>0.91 (0.74 - 1.14)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α₁,₃</td>
<td>1.66 (1.22 - 2.44)</td>
<td>1.36 (1.07 - 1.79)</td>
<td>1.15 (0.95 - 1.54)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d₀,₂</td>
<td>1.19 (0.92 - 1.67)</td>
<td>0.97 (0.75 - 1.29)</td>
<td>0.83 (0.66 - 1.02)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α₂,₃</td>
<td>1.53 (1.12 - 2.23)</td>
<td>1.32 (0.96 - 1.83)</td>
<td>1.11 (0.90 - 1.51)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α₃,₁</td>
<td>2.3 (1.55 - 3.66)</td>
<td>1.89 (1.39 - 3.23)</td>
<td>1.57 (1.11 - 2.33)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α₃,₂</td>
<td>1.78 (1.22 - 3.17)</td>
<td>1.55 (1.11 - 2.25)</td>
<td>1.35 (1.07 - 1.96)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d₀,₃</td>
<td>1.27 (0.98 - 1.83)</td>
<td>1.06 (0.8 - 1.38)</td>
<td>0.9 (0.68 - 1.21)</td>
</tr>
<tr>
<td>covariate₁</td>
<td></td>
<td>0.87 (0.58 - 1.57)</td>
<td>0.65 (0.46 - 1.05)</td>
<td>0.53 (0.4 - 0.83)</td>
<td></td>
</tr>
<tr>
<td>covariate₂</td>
<td></td>
<td>0.93 (0.61 - 1.56)</td>
<td>0.72 (0.51 - 0.97)</td>
<td>0.61 (0.44 - 0.88)</td>
<td></td>
</tr>
<tr>
<td>covariate₃</td>
<td></td>
<td>1.05 (0.69 - 1.67)</td>
<td>0.84 (0.58 - 1.31)</td>
<td>0.7 (0.49 - 1.01)</td>
<td></td>
</tr>
<tr>
<td>Detection</td>
<td></td>
<td>p₁</td>
<td>0.022 (0.017 - 0.027)</td>
<td>0.017 (0.013 - 0.022)</td>
<td>0.015 (0.011 - 0.019)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>p₂</td>
<td>0.023 (0.018 - 0.028)</td>
<td>0.019 (0.014 - 0.023)</td>
<td>0.016 (0.012 - 0.021)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>p₃</td>
<td>0.022 (0.017 - 0.03)</td>
<td>0.019 (0.014 - 0.024)</td>
<td>0.016 (0.013 - 0.021)</td>
</tr>
</tbody>
</table>
Three-species horizontal competition via colonization

Similar to the previous scenario, the accuracy, and precision of all parameter estimates (both inside and outside of the impact matrix) increased as more data was supplied and the model was able to disentangle interspecific effects from each species baseline colonization rate (Fig. 15). With 8 seasons of data, coverage was high for all parameters and ranged from 0.91 – 0.98 (Table XI). With only 4 seasons of data, accuracy and precision of the largest interspecific effect ($a_{3,1}$) were both low relative to all other parameter estimates. However, increasing the amount of data made this parameter more estimable (Fig. 15). Overall, these results suggest that our model was able to recover the true parameter estimates with reasonable accuracy and precision under this scenario, but requires more data if interspecific effects are large (on the logit scale).
Figure 15. The A) accuracy, B) precision, and C) coverage of model parameters (1-9) estimated by the co-occurrence model for three-species horizontal competition where species influenced each other’s colonization rates when 4, 6, and 8 seasons of data are supplied (x-axes). Accuracy is measured as the mean absolute error (MAE) of parameter estimates, precision as the measured width of a parameters 95% credible interval, and coverage as the proportion of the 200 models whose posterior distribution contained the true parameter value with its 95% credible interval. The alignment of the model parameters (1-9) reflects the impact matrix with the species-specific intercepts replacing the 0-valued diagonals. For accuracy and precision, the gray dots are median estimates across the 200 simulations for each scenario while the black lines are 95% highest density intervals. Coverage is represented by a single value, and note the scale of the y-axes.
### TABLE XIII

Mean Absolute Error (MAE), Coverage, and Precision of Persistence ($\Phi$), Colonization ($\gamma$) and Detection Parameters Estimated by the Co-occurrence Model for the Horizontal Competition Module When Species Influenced Each Other’s Colonization Rates. Simulations were run 200 times with varying amounts of data (4, 6, and 8 seasons worth). The median estimate and 95% quantiles are listed for MAE and precision. Coverage was calculated as the proportion of simulations that had the true simulated parameter value within its 95% credible interval.

<table>
<thead>
<tr>
<th>Summary statistic</th>
<th>Process</th>
<th>Parameter</th>
<th>4</th>
<th>6</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>$\gamma$</td>
<td>$c_{0,1}$</td>
<td>0.29 (0.17 - 0.64)</td>
<td>0.2 (0.13 - 0.39)</td>
<td>0.17 (0.11 - 0.32)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$c_{1,2}$</td>
<td>0.41 (0.25 - 0.79)</td>
<td>0.3 (0.2 - 0.63)</td>
<td>0.26 (0.18 - 0.49)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$c_{1,3}$</td>
<td>0.33 (0.21 - 0.67)</td>
<td>0.24 (0.18 - 0.59)</td>
<td>0.21 (0.15 - 0.42)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$c_{2,1}$</td>
<td>0.53 (0.28 - 0.94)</td>
<td>0.34 (0.24 - 0.75)</td>
<td>0.28 (0.19 - 0.64)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$c_{0,2}$</td>
<td>0.33 (0.18 - 0.71)</td>
<td>0.23 (0.14 - 0.48)</td>
<td>0.2 (0.12 - 0.4)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$c_{2,3}$</td>
<td>0.48 (0.28 - 0.81)</td>
<td>0.35 (0.23 - 0.59)</td>
<td>0.28 (0.2 - 0.57)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$c_{1,1}$</td>
<td>0.67 (0.32 - 1.65)</td>
<td>0.39 (0.24 - 1.01)</td>
<td>0.33 (0.21 - 0.67)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$c_{1,2}$</td>
<td>0.58 (0.31 - 0.99)</td>
<td>0.38 (0.24 - 0.81)</td>
<td>0.31 (0.21 - 0.68)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$c_{0,3}$</td>
<td>0.34 (0.19 - 0.86)</td>
<td>0.23 (0.15 - 0.56)</td>
<td>0.19 (0.12 - 0.4)</td>
</tr>
<tr>
<td></td>
<td>covariate</td>
<td>1</td>
<td>0.25 (0.13 - 0.64)</td>
<td>0.17 (0.1 - 0.4)</td>
<td>0.14 (0.08 - 0.31)</td>
</tr>
<tr>
<td></td>
<td>covariate</td>
<td>2</td>
<td>0.25 (0.12 - 0.57)</td>
<td>0.17 (0.1 - 0.38)</td>
<td>0.15 (0.09 - 0.3)</td>
</tr>
<tr>
<td></td>
<td>covariate</td>
<td>3</td>
<td>0.28 (0.13 - 0.58)</td>
<td>0.18 (0.1 - 0.38)</td>
<td>0.15 (0.08 - 0.3)</td>
</tr>
<tr>
<td>$\Phi$</td>
<td>$d_{0,1}$</td>
<td></td>
<td>0.21 (0.12 - 0.45)</td>
<td>0.14 (0.1 - 0.34)</td>
<td>0.13 (0.08 - 0.25)</td>
</tr>
<tr>
<td></td>
<td>$d_{0,2}$</td>
<td></td>
<td>0.23 (0.12 - 0.48)</td>
<td>0.16 (0.1 - 0.29)</td>
<td>0.13 (0.09 - 0.26)</td>
</tr>
<tr>
<td></td>
<td>$d_{0,3}$</td>
<td></td>
<td>0.26 (0.14 - 0.6)</td>
<td>0.16 (0.11 - 0.35)</td>
<td>0.15 (0.09 - 0.31)</td>
</tr>
<tr>
<td></td>
<td>covariate</td>
<td>1</td>
<td>0.24 (0.11 - 0.54)</td>
<td>0.17 (0.09 - 0.33)</td>
<td>0.14 (0.08 - 0.3)</td>
</tr>
<tr>
<td></td>
<td>covariate</td>
<td>2</td>
<td>0.25 (0.12 - 0.61)</td>
<td>0.17 (0.09 - 0.33)</td>
<td>0.16 (0.1 - 0.3)</td>
</tr>
<tr>
<td></td>
<td>covariate</td>
<td>3</td>
<td>0.3 (0.14 - 0.75)</td>
<td>0.19 (0.13 - 0.42)</td>
<td>0.17 (0.09 - 0.33)</td>
</tr>
<tr>
<td>Detection</td>
<td>$p_1$</td>
<td></td>
<td>0.006 (0.004 - 0.015)</td>
<td>0.004 (0.003 - 0.01)</td>
<td>0.004 (0.003 - 0.008)</td>
</tr>
<tr>
<td></td>
<td>$p_2$</td>
<td></td>
<td>0.007 (0.003 - 0.019)</td>
<td>0.005 (0.003 - 0.011)</td>
<td>0.004 (0.003 - 0.009)</td>
</tr>
<tr>
<td></td>
<td>$p_3$</td>
<td></td>
<td>0.008 (0.004 - 0.021)</td>
<td>0.005 (0.003 - 0.01)</td>
<td>0.005 (0.003 - 0.01)</td>
</tr>
</tbody>
</table>
**APPENDIX C** (Continued)

**TABLE XIV** (continued)

Mean absolute error (MAE), coverage, and precision of persistence ($\Phi$), colonization ($\gamma$) and detection parameters estimated by the co-occurrence model for the horizontal competition module when species influenced each other’s colonization rates. Simulations were run 200 times with varying amounts of data (4, 6, and 8 seasons worth). The median estimate and 95% quantiles are listed for MAE and precision. Coverage was calculated as the proportion of simulations that had the true simulated parameter value within its 95% credible interval.

<table>
<thead>
<tr>
<th>Summary statistic</th>
<th>Process</th>
<th>Parameter</th>
<th>4</th>
<th>6</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coverage $\gamma$</td>
<td></td>
<td>$c_{0,1}$</td>
<td>0.98</td>
<td>0.97</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\alpha_{1,2}$</td>
<td>0.95</td>
<td>0.93</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\alpha_{1,3}$</td>
<td>0.99</td>
<td>0.96</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\alpha_{2,1}$</td>
<td>0.89</td>
<td>0.93</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$c_{0,2}$</td>
<td>0.92</td>
<td>0.94</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\alpha_{2,3}$</td>
<td>0.96</td>
<td>0.96</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\alpha_{3,1}$</td>
<td>0.92</td>
<td>0.92</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\alpha_{3,2}$</td>
<td>0.89</td>
<td>0.93</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$c_{0,3}$</td>
<td>0.91</td>
<td>0.91</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>covariate $1$</td>
<td>0.97</td>
<td>0.94</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>covariate $2$</td>
<td>0.97</td>
<td>0.96</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>covariate $3$</td>
<td>0.96</td>
<td>0.98</td>
<td>0.96</td>
</tr>
<tr>
<td>$\Phi$</td>
<td></td>
<td>$d_{0,1}$</td>
<td>0.98</td>
<td>0.93</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$d_{0,2}$</td>
<td>0.97</td>
<td>0.97</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$d_{0,3}$</td>
<td>0.94</td>
<td>0.97</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>covariate $1$</td>
<td>0.97</td>
<td>0.94</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>covariate $2$</td>
<td>0.94</td>
<td>0.96</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>covariate $3$</td>
<td>0.96</td>
<td>0.94</td>
<td>0.97</td>
</tr>
<tr>
<td>Detection</td>
<td></td>
<td>$p_{1}$</td>
<td>0.98</td>
<td>0.95</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$p_{2}$</td>
<td>0.94</td>
<td>0.94</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$p_{3}$</td>
<td>0.94</td>
<td>0.96</td>
<td>0.93</td>
</tr>
</tbody>
</table>
APPENDIX C (Continued)

**TABLE XVI (continued)**

Mean absolute error (MAE), coverage, and precision of persistence ($\Phi$), colonization ($\gamma$) and detection parameters estimated by the co-occurrence model for the horizontal competition module when species influenced each other’s colonization rates. Simulations were run 200 times with varying amounts of data (4, 6, and 8 seasons worth). The median estimate and 95% quantiles are listed for MAE and precision. Coverage was calculated as the proportion of simulations that had the true simulated parameter value within its 95% credible interval.

<table>
<thead>
<tr>
<th>Number of seasons</th>
<th>Summary statistic</th>
<th>Process</th>
<th>Parameter</th>
<th>4</th>
<th>6</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Precision $\gamma$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$c_{0,1}$</td>
<td></td>
<td></td>
<td>1.05 (0.72 - 1.83)</td>
<td>0.78 (0.64 - 1.06)</td>
<td>0.66 (0.54 - 0.9)</td>
</tr>
<tr>
<td></td>
<td>$\alpha_{1,2}$</td>
<td></td>
<td></td>
<td>1.5 (1.02 - 2.48)</td>
<td>1.09 (0.86 - 1.54)</td>
<td>0.94 (0.76 - 1.19)</td>
</tr>
<tr>
<td></td>
<td>$\alpha_{1,3}$</td>
<td></td>
<td></td>
<td>1.44 (1.02 - 2.57)</td>
<td>1.04 (0.86 - 1.51)</td>
<td>0.9 (0.73 - 1.12)</td>
</tr>
<tr>
<td></td>
<td>$\alpha_{2,1}$</td>
<td></td>
<td></td>
<td>1.66 (1.2 - 3.02)</td>
<td>1.23 (0.97 - 1.61)</td>
<td>1.03 (0.84 - 1.34)</td>
</tr>
<tr>
<td></td>
<td>$c_{0,2}$</td>
<td></td>
<td></td>
<td>1.15 (0.82 - 2.02)</td>
<td>0.84 (0.68 - 1.09)</td>
<td>0.7 (0.57 - 0.89)</td>
</tr>
<tr>
<td></td>
<td>$\alpha_{2,3}$</td>
<td></td>
<td></td>
<td>1.65 (1.11 - 3.06)</td>
<td>1.21 (0.92 - 1.66)</td>
<td>1.06 (0.78 - 1.33)</td>
</tr>
<tr>
<td></td>
<td>$\alpha_{3,1}$</td>
<td></td>
<td></td>
<td>2.22 (1.39 - 6.3)</td>
<td>1.55 (1.14 - 2.57)</td>
<td>1.26 (0.98 - 1.93)</td>
</tr>
<tr>
<td></td>
<td>$\alpha_{3,2}$</td>
<td></td>
<td></td>
<td>1.79 (1.16 - 3.43)</td>
<td>1.35 (0.97 - 1.88)</td>
<td>1.16 (0.89 - 1.45)</td>
</tr>
<tr>
<td></td>
<td>$c_{0,3}$</td>
<td></td>
<td></td>
<td>1.17 (0.88 - 2.15)</td>
<td>0.87 (0.65 - 1.13)</td>
<td>0.72 (0.59 - 0.96)</td>
</tr>
<tr>
<td></td>
<td>covariate$_1$</td>
<td></td>
<td></td>
<td>1 (0.6 - 2.08)</td>
<td>0.63 (0.48 - 1.12)</td>
<td>0.53 (0.38 - 0.82)</td>
</tr>
<tr>
<td></td>
<td>covariate$_2$</td>
<td></td>
<td></td>
<td>0.97 (0.6 - 1.87)</td>
<td>0.64 (0.46 - 0.97)</td>
<td>0.54 (0.38 - 0.81)</td>
</tr>
<tr>
<td></td>
<td>covariate$_3$</td>
<td></td>
<td></td>
<td>1 (0.64 - 1.96)</td>
<td>0.67 (0.5 - 1.05)</td>
<td>0.56 (0.44 - 0.81)</td>
</tr>
<tr>
<td></td>
<td>$\Phi$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$d_{0,1}$</td>
<td></td>
<td></td>
<td>0.78 (0.57 - 1.43)</td>
<td>0.57 (0.46 - 0.75)</td>
<td>0.49 (0.4 - 0.66)</td>
</tr>
<tr>
<td></td>
<td>$d_{0,2}$</td>
<td></td>
<td></td>
<td>0.86 (0.57 - 1.4)</td>
<td>0.61 (0.48 - 0.83)</td>
<td>0.52 (0.42 - 0.67)</td>
</tr>
<tr>
<td></td>
<td>$d_{0,3}$</td>
<td></td>
<td></td>
<td>0.95 (0.6 - 1.61)</td>
<td>0.67 (0.5 - 0.89)</td>
<td>0.57 (0.42 - 0.76)</td>
</tr>
<tr>
<td></td>
<td>covariate$_1$</td>
<td></td>
<td></td>
<td>0.9 (0.58 - 1.8)</td>
<td>0.62 (0.43 - 0.89)</td>
<td>0.54 (0.39 - 0.79)</td>
</tr>
<tr>
<td></td>
<td>covariate$_2$</td>
<td></td>
<td></td>
<td>0.97 (0.62 - 1.87)</td>
<td>0.67 (0.49 - 1.09)</td>
<td>0.57 (0.39 - 0.81)</td>
</tr>
<tr>
<td></td>
<td>covariate$_3$</td>
<td></td>
<td></td>
<td>1.08 (0.58 - 2.13)</td>
<td>0.76 (0.49 - 1.08)</td>
<td>0.63 (0.46 - 0.97)</td>
</tr>
<tr>
<td></td>
<td>Detection</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$p_1$</td>
<td></td>
<td></td>
<td>0.023 (0.017 - 0.041)</td>
<td>0.017 (0.013 - 0.023)</td>
<td>0.015 (0.011 - 0.02)</td>
</tr>
<tr>
<td></td>
<td>$p_2$</td>
<td></td>
<td></td>
<td>0.025 (0.017 - 0.049)</td>
<td>0.018 (0.014 - 0.024)</td>
<td>0.016 (0.013 - 0.021)</td>
</tr>
<tr>
<td></td>
<td>$p_3$</td>
<td></td>
<td></td>
<td>0.027 (0.017 - 0.049)</td>
<td>0.02 (0.014 - 0.027)</td>
<td>0.017 (0.013 - 0.025)</td>
</tr>
</tbody>
</table>
APPENDIX C (continued)

Predator with prey competing via persistence

While accuracy and precision did increase in this scenario as more data was supplied, there was a general tendency for these simulations to be less accurate and precise than the previous community module (Fig. 16; Table XII). This effect was most exacerbated by the pair-wise interaction between species 3 (the predator) and species 1 (who was most negatively influenced by species 3). The second largest negative effect, $\sigma_{1,2}$, was also less accurate and precise. However, this decrease in accuracy and precision is likely attributable to our choice in large parameter values as the strong negative effects that species 2 and 3 had on species 1 greatly decreased the possibility of co-occurrence (and therefore rendered estimation of the parameter more difficult). This is further evidenced by the ability of the model to easily estimate all of the other parameter values. With 8 seasons of data, coverage ranged from 0.91 – 0.97 for all parameter values (Table XII). These results suggest that our model was able to recover the true parameter estimates with reasonable accuracy and precision under this scenario, but requires more data if interspecific effects are large (on the logit scale).
Figure 16. The A) accuracy, B) precision, and C) coverage of model parameters (1-9) estimated by the co-occurrence model for a predator with two competing prey where species influenced each other’s persistence rates when 4, 6, and 8 seasons of data are supplied (x-axes). Accuracy is measured as the mean absolute error (MAE) of parameter estimates, precision as the measured width of a parameters 95% credible interval, and coverage as the proportion of the 200 models whose posterior distribution contained the true parameter value with its 95% credible interval. The alignment of the model parameters (1-9) reflects the impact matrix with the species-specific intercepts replacing the 0-valued diagonals. For accuracy and precision, the gray dots are median estimates across the 200 simulations for each scenario while the black lines are 95% highest density intervals. Coverage is represented by a single value, and note the scale of the y-axes.
TABLE XVIII
MEAN ABSOLUTE ERROR (MAE), COVERAGE, AND PRECISION OF PERSISTENCE (Φ),
COLONIZATION (γ) AND DETECTION PARAMETERS ESTIMATED BY THE CO-OCCURRENCE MODEL
FOR THE PREDATOR WITH COMPETING PREY MODULE WHEN SPECIES INFLUENCED EACH
OTHER’S PERSISTENCE RATES. SIMULATIONS WERE RUN 200 TIMES WITH VARYING AMOUNTS
OF DATA (4, 6, AND 8 SEASONS WORTH). THE MEDIAN ESTIMATE AND 95% QUANTILES ARE
LISTED FOR MAE AND PRECISION. COVERAGE WAS CALCULATED AS THE PROPORTION OF
SIMULATIONS THAT HAD THE TRUE SIMULATED PARAMETER VALUE WITHIN ITS 95% CREDIBLE
INTERVAL.

<table>
<thead>
<tr>
<th>Summary statistic</th>
<th>Process</th>
<th>Parameter</th>
<th>4</th>
<th>6</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE Φ</td>
<td>d₀,₁</td>
<td>0.41 (0.21 - 0.99)</td>
<td>0.31 (0.19 - 0.74)</td>
<td>0.26 (0.14 - 0.57)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>α₁,₂</td>
<td>0.79 (0.47 - 1.33)</td>
<td>0.6 (0.3 - 1.2)</td>
<td>0.47 (0.27 - 0.95)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>α₁,₃</td>
<td>0.69 (0.4 - 1.46)</td>
<td>0.56 (0.3 - 1.16)</td>
<td>0.44 (0.26 - 0.99)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>α₂,₁</td>
<td>0.4 (0.28 - 0.74)</td>
<td>0.32 (0.21 - 0.64)</td>
<td>0.29 (0.2 - 0.54)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>d₀,₂</td>
<td>0.36 (0.2 - 0.88)</td>
<td>0.28 (0.17 - 0.68)</td>
<td>0.23 (0.13 - 0.49)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>α₂,₃</td>
<td>0.5 (0.31 - 0.87)</td>
<td>0.38 (0.23 - 0.78)</td>
<td>0.31 (0.2 - 0.59)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>α₃,₁</td>
<td>0.65 (0.37 - 2.85)</td>
<td>0.51 (0.28 - 1.58)</td>
<td>0.41 (0.24 - 0.84)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>α₃,₂</td>
<td>0.49 (0.29 - 0.88)</td>
<td>0.35 (0.23 - 0.76)</td>
<td>0.31 (0.2 - 0.62)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>d₀,₃</td>
<td>0.25 (0.17 - 0.57)</td>
<td>0.2 (0.14 - 0.39)</td>
<td>0.16 (0.11 - 0.35)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>covariate₁</td>
<td>0.32 (0.16 - 0.74)</td>
<td>0.25 (0.14 - 0.55)</td>
<td>0.2 (0.12 - 0.47)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>covariate₂</td>
<td>0.24 (0.14 - 0.59)</td>
<td>0.19 (0.1 - 0.43)</td>
<td>0.16 (0.09 - 0.36)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>covariate₃</td>
<td>0.22 (0.12 - 0.42)</td>
<td>0.17 (0.1 - 0.37)</td>
<td>0.15 (0.09 - 0.31)</td>
<td></td>
</tr>
<tr>
<td>γ</td>
<td>c₀,₁</td>
<td>0.17 (0.11 - 0.37)</td>
<td>0.14 (0.09 - 0.28)</td>
<td>0.12 (0.08 - 0.25)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c₀,₂</td>
<td>0.17 (0.12 - 0.36)</td>
<td>0.14 (0.09 - 0.31)</td>
<td>0.12 (0.08 - 0.25)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c₀,₃</td>
<td>0.22 (0.13 - 0.49)</td>
<td>0.17 (0.11 - 0.42)</td>
<td>0.15 (0.09 - 0.34)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>covariate₁</td>
<td>0.19 (0.11 - 0.39)</td>
<td>0.14 (0.09 - 0.29)</td>
<td>0.12 (0.07 - 0.24)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>covariate₂</td>
<td>0.19 (0.11 - 0.49)</td>
<td>0.16 (0.08 - 0.35)</td>
<td>0.13 (0.07 - 0.28)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>covariate₃</td>
<td>0.22 (0.14 - 0.53)</td>
<td>0.18 (0.11 - 0.45)</td>
<td>0.16 (0.09 - 0.33)</td>
<td></td>
</tr>
<tr>
<td>Detection</td>
<td>p₁</td>
<td>0.006 (0.004 - 0.013)</td>
<td>0.005 (0.003 - 0.011)</td>
<td>0.004 (0.003 - 0.009)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p₂</td>
<td>0.006 (0.004 - 0.011)</td>
<td>0.005 (0.003 - 0.009)</td>
<td>0.004 (0.003 - 0.009)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p₃</td>
<td>0.005 (0.003 - 0.009)</td>
<td>0.004 (0.002 - 0.009)</td>
<td>0.003 (0.002 - 0.007)</td>
<td></td>
</tr>
</tbody>
</table>
### APPENDIX C (continued)

**TABLE XVIII (continued)**

Mean absolute error (MAE), coverage, and precision of persistence ($\Phi$), colonization ($\gamma$) and detection parameters estimated by the co-occurrence model for the predator with competing prey module when species influenced each other’s persistence rates. Simulations were run 200 times with varying amounts of data (4, 6, and 8 seasons worth). The median estimate and 95% quantiles are listed for MAE and precision. Coverage was calculated as the proportion of simulations that had the true simulated parameter value within its 95% credible interval.

<table>
<thead>
<tr>
<th>Summary statistic</th>
<th>Process</th>
<th>Parameter</th>
<th>Number of seasons</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Coverage $\Phi$</td>
<td></td>
<td>$d_{0,1}$</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\alpha_{1,2}$</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\alpha_{1,3}$</td>
<td>0.91</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\alpha_{2,1}$</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$d_{0,2}$</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\alpha_{3,1}$</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$\alpha_{3,2}$</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$d_{0,3}$</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>covariate 1</td>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>covariate 2</td>
<td>0.96</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>covariate 3</td>
<td>0.96</td>
<td>0.98</td>
</tr>
<tr>
<td>Coverage $\gamma$</td>
<td></td>
<td>$c_{0,1}$</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$c_{0,2}$</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$c_{0,3}$</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>covariate 1</td>
<td>0.98</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>covariate 2</td>
<td>0.95</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>covariate 3</td>
<td>0.94</td>
<td>0.95</td>
</tr>
<tr>
<td>Detection $p$</td>
<td></td>
<td>$p_1$</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$p_2$</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td></td>
<td>$p_3$</td>
<td>0.96</td>
</tr>
</tbody>
</table>
APPENDIX C (continued)

TABLE XVIII (continued)

Mean Absolute Error (MAE), Coverage, and Precision of Persistence (Φ), Colonization (γ) and Detection Parameters Estimated by the Co-Occurrence Model for the Predator with Competing Prey Module When Species Influenced Each Other’s Persistence Rates. Simulations Were Run 200 Times with Varying Amounts of Data (4, 6, and 8 Seasons Worth). The Median Estimate and 95% Quantiles Are Listed for MAE and Precision. Coverage Was Calculated as the Proportion of Simulations That Had the True Simulated Parameter Value Within Its 95% Credible Interval.

<table>
<thead>
<tr>
<th>Summary statistic</th>
<th>Process</th>
<th>Parameter</th>
<th>4</th>
<th>6</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision Φ</td>
<td>d0,1</td>
<td>1.54 (1.06 - 2.37)</td>
<td>1.26 (0.9 - 1.8)</td>
<td>0.93</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a1,2</td>
<td>2.45 (1.54 - 4.44)</td>
<td>1.97 (1.37 - 3.28)</td>
<td>1.71 (1.18 - 2.32)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a1,3</td>
<td>2.49 (1.86 - 4.5)</td>
<td>2 (1.41 - 2.99)</td>
<td>1.67 (1.21 - 2.47)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a2,1</td>
<td>1.5 (1.08 - 2.25)</td>
<td>1.22 (0.93 - 1.67)</td>
<td>1.05 (0.86 - 1.43)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>d0,2</td>
<td>1.31 (0.97 - 1.71)</td>
<td>1.05 (0.8 - 1.41)</td>
<td>0.92 (0.7 - 1.23)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a2,3</td>
<td>1.69 (1.19 - 2.35)</td>
<td>1.34 (1.02 - 1.84)</td>
<td>1.16 (0.92 - 1.54)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a3,1</td>
<td>2.14 (1.43 - 22.24)</td>
<td>1.79 (1.22 - 9.72)</td>
<td>1.5 (1.06 - 2.44)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a3,2</td>
<td>1.61 (1.21 - 2.34)</td>
<td>1.3 (0.98 - 1.82)</td>
<td>1.14 (0.88 - 1.59)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>d0,3</td>
<td>0.94 (0.76 - 1.27)</td>
<td>0.76 (0.62 - 0.97)</td>
<td>0.65 (0.51 - 0.82)</td>
<td></td>
</tr>
<tr>
<td>covariate1</td>
<td>1.18 (0.84 - 2.17)</td>
<td>0.95 (0.62 - 1.48)</td>
<td>0.77 (0.53 - 1.19)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>covariate2</td>
<td>0.94 (0.63 - 1.57)</td>
<td>0.73 (0.53 - 1.08)</td>
<td>0.62 (0.44 - 0.86)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>covariate3</td>
<td>0.81 (0.57 - 1.35)</td>
<td>0.65 (0.48 - 1)</td>
<td>0.55 (0.4 - 0.83)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>γ</td>
<td>c0,1</td>
<td>0.64 (0.55 - 0.93)</td>
<td>0.5 (0.43 - 0.73)</td>
<td>0.43 (0.36 - 0.6)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c0,2</td>
<td>0.66 (0.57 - 0.97)</td>
<td>0.53 (0.44 - 0.77)</td>
<td>0.44 (0.37 - 0.61)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>c0,3</td>
<td>0.83 (0.65 - 1.12)</td>
<td>0.65 (0.51 - 0.95)</td>
<td>0.55 (0.45 - 0.78)</td>
<td></td>
</tr>
<tr>
<td>covariate1</td>
<td>0.7 (0.48 - 1.1)</td>
<td>0.53 (0.4 - 0.84)</td>
<td>0.46 (0.35 - 0.7)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>covariate2</td>
<td>0.75 (0.51 - 1.19)</td>
<td>0.57 (0.39 - 0.97)</td>
<td>0.48 (0.35 - 0.74)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>covariate3</td>
<td>0.91 (0.63 - 1.45)</td>
<td>0.7 (0.46 - 1.15)</td>
<td>0.6 (0.41 - 0.88)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Detection</td>
<td>p1</td>
<td>0.023 (0.018 - 0.03)</td>
<td>0.017 (0.013 - 0.022)</td>
<td>0.016 (0.013 - 0.022)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p2</td>
<td>0.022 (0.017 - 0.029)</td>
<td>0.019 (0.014 - 0.023)</td>
<td>0.016 (0.012 - 0.021)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p3</td>
<td>0.019 (0.016 - 0.025)</td>
<td>0.019 (0.014 - 0.024)</td>
<td>0.013 (0.011 - 0.017)</td>
<td></td>
</tr>
</tbody>
</table>
APPENDIX C (continued)

Predator with prey competing via colonization

Like all of the other scenarios, accuracy and precision of parameter estimated increased as more data were supplied: the coverages of parameter estimates were around the nominal level of 95% with 8 seasons of data (range = 0.91 – 0.97). Similar to the third scenario (predator with prey competing via persistence), accuracy and precision for $a_{3,1}$ (the largest positive interactive effect) was lower, but all other parameter effects performed similarly to the first two scenarios (Fig. 17). We attribute this decrease in performance to our selection of parameter values, as smaller interspecific effects were estimated both with high accuracy and precision (Table XIII). Overall, these results suggest that our model was able to recover the true parameter estimates with reasonable accuracy and precision under this scenario, but requires more data if interspecific effects are large (on the logit scale).
APPENDIX C (continued)

Figure 17. The A) accuracy, B) precision, and C) coverage of model parameters (1-9) estimated by the co-occurrence model for a predator with two competing prey where species influenced each other’s persistence rates when 4, 6, and 8 seasons of data are supplied (x-axes). Accuracy is measured as the mean absolute error (MAE) of parameter estimates, precision as the measured width of a parameters 95% credible interval, and coverage as the proportion of the 200 models whose posterior distribution contained the true parameter value with its 95% credible interval. The alignment of the model parameters (1-9) reflects the impact matrix with the species-specific intercepts replacing the 0-valued diagonals. For accuracy and precision, the gray dots are median estimates across the 200 simulations for each scenario while the black lines are 95% highest density intervals. Coverage is represented by a single value, and note the scale of the y-axes.
APPENDIX C (continued)

TABLE XIXII
MEAN ABSOLUTE ERROR (MAE), COVERAGE, AND PRECISION OF PERSISTENCE (Φ), COLONIZATION (γ) AND DETECTION PARAMETERS ESTIMATED BY THE CO-OCCURRENCE MODEL FOR THE PREDATOR WITH COMPETING PREY MODULE WHEN SPECIES INFLUENCED EACH OTHER’S PERSISTENCE RATES. SIMULATIONS WERE RUN 200 TIMES WITH VARYING AMOUNTS OF DATA (4, 6, AND 8 SEASONS WORTH). THE MEDIAN ESTIMATE AND 95% QUANTILES ARE LISTED FOR MAE AND PRECISION. COVERAGE WAS CALCULATED AS THE PROPORTION OF SIMULATIONS THAT HAD THE TRUE SIMULATED PARAMETER VALUE WITHIN ITS 95% CREDIBLE INTERVAL.

<table>
<thead>
<tr>
<th>Summary statistic</th>
<th>Process</th>
<th>Parameter</th>
<th>4</th>
<th>6</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAE</td>
<td>γ</td>
<td>c_{0,1}</td>
<td>0.31 (0.18 - 0.77)</td>
<td>0.24 (0.15 - 0.56)</td>
<td>0.21 (0.14 - 0.5)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α_{1,2}</td>
<td>0.6 (0.34 - 1.27)</td>
<td>0.45 (0.28 - 0.99)</td>
<td>0.37 (0.23 - 0.82)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α_{1,3}</td>
<td>0.53 (0.32 - 1.1)</td>
<td>0.39 (0.25 - 0.87)</td>
<td>0.33 (0.2 - 0.77)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α_{2,1}</td>
<td>0.38 (0.27 - 0.71)</td>
<td>0.32 (0.21 - 0.64)</td>
<td>0.27 (0.18 - 0.53)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c_{0,2}</td>
<td>0.28 (0.18 - 0.63)</td>
<td>0.23 (0.14 - 0.48)</td>
<td>0.19 (0.12 - 0.42)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α_{2,3}</td>
<td>0.45 (0.27 - 0.81)</td>
<td>0.32 (0.21 - 0.68)</td>
<td>0.28 (0.18 - 0.52)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α_{3,1}</td>
<td>0.9 (0.4 - 4.76)</td>
<td>0.64 (0.29 - 4.73)</td>
<td>0.55 (0.25 - 2.28)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α_{3,2}</td>
<td>0.57 (0.32 - 1.02)</td>
<td>0.44 (0.26 - 0.81)</td>
<td>0.37 (0.23 - 0.8)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>c_{0,3}</td>
<td>0.32 (0.17 - 0.66)</td>
<td>0.23 (0.16 - 0.56)</td>
<td>0.2 (0.13 - 0.52)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>covariate₁</td>
<td>0.24 (0.12 - 0.51)</td>
<td>0.19 (0.11 - 0.39)</td>
<td>0.16 (0.1 - 0.33)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>covariate₂</td>
<td>0.21 (0.12 - 0.48)</td>
<td>0.17 (0.1 - 0.37)</td>
<td>0.14 (0.08 - 0.3)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>covariate₃</td>
<td>0.27 (0.15 - 0.6)</td>
<td>0.21 (0.13 - 0.44)</td>
<td>0.17 (0.09 - 0.41)</td>
</tr>
<tr>
<td>Φ</td>
<td>d_{0,1}</td>
<td>0.23 (0.13 - 0.52)</td>
<td>0.18 (0.12 - 0.4)</td>
<td>0.16 (0.1 - 0.32)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>d_{0,2}</td>
<td>0.2 (0.12 - 0.4)</td>
<td>0.16 (0.1 - 0.36)</td>
<td>0.14 (0.08 - 0.28)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>d_{0,3}</td>
<td>0.17 (0.12 - 0.34)</td>
<td>0.13 (0.09 - 0.28)</td>
<td>0.11 (0.08 - 0.24)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>covariate₁</td>
<td>0.27 (0.15 - 0.63)</td>
<td>0.21 (0.11 - 0.43)</td>
<td>0.18 (0.1 - 0.4)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>covariate₂</td>
<td>0.24 (0.12 - 0.52)</td>
<td>0.18 (0.1 - 0.39)</td>
<td>0.15 (0.08 - 0.3)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>covariate₃</td>
<td>0.18 (0.11 - 0.44)</td>
<td>0.14 (0.09 - 0.38)</td>
<td>0.12 (0.08 - 0.23)</td>
<td></td>
</tr>
<tr>
<td>Detection</td>
<td>p₁</td>
<td>0.007 (0.004 - 0.016)</td>
<td>0.006 (0.003 - 0.013)</td>
<td>0.005 (0.003 - 0.009)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p₂</td>
<td>0.006 (0.004 - 0.013)</td>
<td>0.005 (0.003 - 0.01)</td>
<td>0.004 (0.003 - 0.01)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>p₃</td>
<td>0.005 (0.003 - 0.01)</td>
<td>0.004 (0.002 - 0.009)</td>
<td>0.004 (0.002 - 0.009)</td>
<td></td>
</tr>
</tbody>
</table>
**APPENDIX C** (continued)

**TABLE XXIII** (continued)

MEAN ABSOLUTE ERROR (MAE), COVERAGE, AND PRECISION OF PERSISTENCE ($\Phi$), COLONIZATION ($\gamma$) AND DETECTION PARAMETERS ESTIMATED BY THE CO-OCCURRENCE MODEL FOR THE PREDATOR WITH COMPETING PREY MODULE WHEN SPECIES INFLUENCED EACH OTHER’S PERSISTENCE RATES. SIMULATIONS WERE RUN 200 TIMES WITH VARYING AMOUNTS OF DATA (4, 6, AND 8 SEASONS WORTH). THE MEDIAN ESTIMATE AND 95% QUANTILES ARE LISTED FOR MAE AND PRECISION. COVERAGE WAS CALCULATED AS THE PROPORTION OF SIMULATIONS THAT HAD THE TRUE SIMULATED PARAMETER VALUE WITHIN ITS 95% CREDIBLE INTERVAL.

<table>
<thead>
<tr>
<th>Summary statistic</th>
<th>Process</th>
<th>Parameter</th>
<th>Number of seasons</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Coverage $\gamma$</td>
<td>$c_{0,1}$</td>
<td>0.9</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>$\alpha_{1,2}$</td>
<td>0.92</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>$\alpha_{1,3}$</td>
<td>0.94</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>$\alpha_{2,1}$</td>
<td>0.96</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>$c_{0,2}$</td>
<td>0.91</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>$\alpha_{2,3}$</td>
<td>0.93</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>$\alpha_{3,1}$</td>
<td>0.86</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>$\alpha_{3,2}$</td>
<td>0.89</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>$c_{0,3}$</td>
<td>0.91</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>covariate$_1$</td>
<td>0.93</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>covariate$_2$</td>
<td>0.94</td>
<td>0.92</td>
</tr>
<tr>
<td></td>
<td>covariate$_3$</td>
<td>0.94</td>
<td>0.95</td>
</tr>
<tr>
<td>$\Phi$</td>
<td>$d_{0,1}$</td>
<td>0.95</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>$d_{0,2}$</td>
<td>0.97</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>$d_{0,3}$</td>
<td>0.93</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>covariate$_1$</td>
<td>0.96</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>covariate$_2$</td>
<td>0.93</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>covariate$_3$</td>
<td>0.95</td>
<td>0.94</td>
</tr>
<tr>
<td>Detection $\rho$</td>
<td>$\rho_1$</td>
<td>0.93</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td>$\rho_2$</td>
<td>0.94</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>$\rho_3$</td>
<td>0.96</td>
<td>0.95</td>
</tr>
</tbody>
</table>
APPENDIX C (continued)

TABLE XXIII (continued)
MEAN ABSOLUTE ERROR (MAE), COVERAGE, AND PRECISION OF PERSISTENCE (Φ), COLONIZATION (γ) AND DETECTION PARAMETERS ESTIMATED BY THE CO-OCCURRENCE MODEL FOR THE PREDATOR WITH COMPETING PREY MODULE WHEN SPECIES INFLUENCED EACH OTHER’S PERSISTENCE RATES. SIMULATIONS WERE RUN 200 TIMES WITH VARYING AMOUNTS OF DATA (4, 6, AND 8 SEASONS WORTH). THE MEDIAN ESTIMATE AND 95% QUANTILES ARE LISTED FOR MAE AND PRECISION. COVERAGE WAS CALCULATED AS THE PROPORTION OF SIMULATIONS THAT HAD THE TRUE SIMULATED PARAMETER VALUE WITHIN ITS 95% CREDIBLE INTERVAL.

<table>
<thead>
<tr>
<th>Summary statistic</th>
<th>Process</th>
<th>Parameter</th>
<th>4</th>
<th>6</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Precision</td>
<td></td>
<td>γ</td>
<td>1.13 (0.9 - 1.63)</td>
<td>0.91 (0.69 - 1.31)</td>
<td>0.77 (0.62 - 0.98)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α_{1,2}</td>
<td>2.02 (1.31 - 3.25)</td>
<td>1.58 (1.1 - 2.37)</td>
<td>1.34 (1.01 - 1.87)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α_{1,3}</td>
<td>1.89 (1.44 - 3.59)</td>
<td>1.48 (1.13 - 2.21)</td>
<td>1.23 (0.94 - 1.72)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α_{2,1}</td>
<td>1.43 (1.09 - 2.08)</td>
<td>1.18 (0.93 - 1.72)</td>
<td>1.02 (0.81 - 1.38)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α_{0,2}</td>
<td>1.05 (0.84 - 1.34)</td>
<td>0.86 (0.7 - 1.26)</td>
<td>0.72 (0.59 - 1.01)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α_{3,3}</td>
<td>1.44 (1.17 - 2.06)</td>
<td>1.19 (0.97 - 1.62)</td>
<td>0.98 (0.82 - 1.31)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α_{3,1}</td>
<td>2.57 (1.57 - 37.92)</td>
<td>2.08 (1.3 - 34.37)</td>
<td>1.81 (1.09 - 20.15)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α_{3,2}</td>
<td>1.79 (1.23 - 3.18)</td>
<td>1.53 (0.99 - 2.43)</td>
<td>1.29 (0.89 - 2.03)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>α_{0,3}</td>
<td>1.12 (0.9 - 1.67)</td>
<td>0.93 (0.7 - 1.27)</td>
<td>0.78 (0.61 - 1.07)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>covariate_{1}</td>
<td>0.89 (0.66 - 1.45)</td>
<td>0.71 (0.53 - 1.08)</td>
<td>0.6 (0.46 - 0.87)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>covariate_{2}</td>
<td>0.82 (0.58 - 1.27)</td>
<td>0.63 (0.47 - 0.92)</td>
<td>0.52 (0.41 - 0.81)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>covariate_{3}</td>
<td>0.99 (0.73 - 1.78)</td>
<td>0.79 (0.54 - 1.22)</td>
<td>0.65 (0.46 - 1.06)</td>
</tr>
<tr>
<td>Φ</td>
<td></td>
<td>d_{0,1}</td>
<td>0.9 (0.68 - 1.33)</td>
<td>0.71 (0.51 - 0.98)</td>
<td>0.62 (0.47 - 0.83)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d_{0,2}</td>
<td>0.79 (0.59 - 1.09)</td>
<td>0.62 (0.49 - 0.86)</td>
<td>0.53 (0.42 - 0.74)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>d_{0,3}</td>
<td>0.63 (0.55 - 0.86)</td>
<td>0.5 (0.45 - 0.65)</td>
<td>0.42 (0.38 - 0.51)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>covariate_{1}</td>
<td>1 (0.69 - 1.94)</td>
<td>0.78 (0.51 - 1.27)</td>
<td>0.66 (0.45 - 0.99)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>covariate_{2}</td>
<td>0.87 (0.55 - 1.37)</td>
<td>0.69 (0.48 - 1.14)</td>
<td>0.58 (0.42 - 0.89)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>covariate_{3}</td>
<td>0.69 (0.52 - 1.22)</td>
<td>0.55 (0.41 - 0.88)</td>
<td>0.46 (0.34 - 0.7)</td>
</tr>
<tr>
<td>Detection</td>
<td></td>
<td>p_{1}</td>
<td>0.027 (0.019 - 0.038)</td>
<td>0.022 (0.014 - 0.029)</td>
<td>0.019 (0.014 - 0.026)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>p_{2}</td>
<td>0.023 (0.017 - 0.031)</td>
<td>0.019 (0.013 - 0.024)</td>
<td>0.016 (0.012 - 0.021)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>p_{3}</td>
<td>0.02 (0.016 - 0.025)</td>
<td>0.016 (0.013 - 0.021)</td>
<td>0.014 (0.011 - 0.017)</td>
</tr>
</tbody>
</table>
### APPENDIX D

#### TABLE XXIIv

VIDEOS WHOSE COMMENTS WERE COLLECTED ON YOUTUBE AND USED TO ANALYZE PUBLIC PERCEPTIONS TOWARDS COYOTE, OPOSSUM, AND RACCOON

<table>
<thead>
<tr>
<th>Species</th>
<th>Video Name</th>
<th>URL</th>
<th>Views at download</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coyote</td>
<td>Wolf vs Two Coyotes</td>
<td><a href="http://www.youtube.com/watch?v=rEB30KLpTM5s">http://www.youtube.com/watch?v=rEB30KLpTM5s</a></td>
<td>2684086</td>
</tr>
<tr>
<td></td>
<td>Wolves vs coyote [HD]</td>
<td><a href="http://www.youtube.com/watch?v=BXCvLzDNWz0">http://www.youtube.com/watch?v=BXCvLzDNWz0</a></td>
<td>2235133</td>
</tr>
<tr>
<td></td>
<td>Coyote 14 yards</td>
<td><a href="http://www.youtube.com/watch?v=Wa3V5tZXT7U">http://www.youtube.com/watch?v=Wa3V5tZXT7U</a></td>
<td>2210832</td>
</tr>
<tr>
<td></td>
<td>Coyote Caught on a Fence</td>
<td><a href="http://www.youtube.com/watch?v=vE007fFnei0">http://www.youtube.com/watch?v=vE007fFnei0</a></td>
<td>2170817</td>
</tr>
<tr>
<td></td>
<td>Bye Bye Coyote!</td>
<td><a href="http://www.youtube.com/watch?v=2P1uyeRqjBo">http://www.youtube.com/watch?v=2P1uyeRqjBo</a></td>
<td>1592183</td>
</tr>
<tr>
<td></td>
<td>Coyote Headshot</td>
<td><a href="http://www.youtube.com/watch?v=9fI7sMHNF2c">http://www.youtube.com/watch?v=9fI7sMHNF2c</a></td>
<td>1139144</td>
</tr>
<tr>
<td></td>
<td>coyote Hunting</td>
<td><a href="http://www.youtube.com/watch?v=LO0IVxXh16g">http://www.youtube.com/watch?v=LO0IVxXh16g</a></td>
<td>1321662</td>
</tr>
<tr>
<td></td>
<td>COYOTE KILL WITH 300 WSM BROWNING BALLISTIC TIP, ONE SHOT!!!</td>
<td><a href="http://www.youtube.com/watch?v=ox1bNqNE2ho">http://www.youtube.com/watch?v=ox1bNqNE2ho</a></td>
<td>1061225</td>
</tr>
<tr>
<td></td>
<td>COYOTE HUNTING</td>
<td><a href="http://www.youtube.com/watch?v=fzCHjNtDNU4">http://www.youtube.com/watch?v=fzCHjNtDNU4</a></td>
<td>993597</td>
</tr>
<tr>
<td>Opossum</td>
<td>Opossum and Cat fight</td>
<td><a href="http://www.youtube.com/watch?v=VShJC71GNzk">http://www.youtube.com/watch?v=VShJC71GNzk</a></td>
<td>3062951</td>
</tr>
<tr>
<td></td>
<td>Ratatouille the Snowboarding Opossum</td>
<td><a href="http://www.youtube.com/watch?v=21XBFaG5zGY">http://www.youtube.com/watch?v=21XBFaG5zGY</a></td>
<td>1228957</td>
</tr>
<tr>
<td></td>
<td>Pet Opossum Complaining</td>
<td><a href="http://www.youtube.com/watch?v=QqHyBzkYaFc">http://www.youtube.com/watch?v=QqHyBzkYaFc</a></td>
<td>851477</td>
</tr>
<tr>
<td></td>
<td>Das schielende Opossum Heidi</td>
<td><a href="http://www.youtube.com/watch?v=K_qlVI3zDSQ">http://www.youtube.com/watch?v=K_qlVI3zDSQ</a></td>
<td>798450</td>
</tr>
<tr>
<td></td>
<td>Raccoon vs Possum</td>
<td><a href="http://www.youtube.com/watch?v=QqHyBzkYaFc">http://www.youtube.com/watch?v=QqHyBzkYaFc</a></td>
<td>726480</td>
</tr>
<tr>
<td></td>
<td>Mary Cummins, Animal Advocates, Opossum eating strawberries</td>
<td><a href="http://www.youtube.com/watch?v=VWWnKlRGRnK">http://www.youtube.com/watch?v=VWWnKlRGRnK</a></td>
<td>628015</td>
</tr>
<tr>
<td></td>
<td>Baby opossum swimming, playing dead &amp; More!</td>
<td><a href="http://www.youtube.com/watch?v=Xh_zcaqAsY">http://www.youtube.com/watch?v=Xh_zcaqAsY</a></td>
<td>611645</td>
</tr>
<tr>
<td></td>
<td>WFW 21 - Awesome Possum</td>
<td><a href="http://www.youtube.com/watch?v=Jk4Sk1Ju4U">http://www.youtube.com/watch?v=Jk4Sk1Ju4U</a></td>
<td>592410</td>
</tr>
<tr>
<td></td>
<td>Mary Cummins, Animal Advocates, Baby opossum calling for mom</td>
<td><a href="http://www.youtube.com/watch?v=Z5SP4Mqe_vo">http://www.youtube.com/watch?v=Z5SP4Mqe_vo</a></td>
<td>564974</td>
</tr>
<tr>
<td></td>
<td>Feeding my possum</td>
<td><a href="http://www.youtube.com/watch?v=UBJRrR13eFl">http://www.youtube.com/watch?v=UBJRrR13eFl</a></td>
<td>530513</td>
</tr>
</tbody>
</table>
APPENDIX D (continued)

TABLE XXIII IV (continued)
VIDEOS WHOSE COMMENTS WERE COLLECTED ON YOUTUBE AND USED TO ANALYZE PUBLIC PERCEPTIONS TOWARDS COYOTE, OPOSSUM, AND RACCOON

<table>
<thead>
<tr>
<th>Species</th>
<th>Video Name</th>
<th>URL</th>
<th>Views at download</th>
</tr>
</thead>
<tbody>
<tr>
<td>Raccoon</td>
<td>Raccoon eating cats food (Original)</td>
<td><a href="http://www.youtube.com/watch?v=FTcjaqL0pE">http://www.youtube.com/watch?v=FTcjaqL0pE</a></td>
<td>2958914</td>
</tr>
<tr>
<td></td>
<td>Kitty vs Raccoon</td>
<td><a href="http://www.youtube.com/watch?v=aJZMunpgOtI">http://www.youtube.com/watch?v=aJZMunpgOtI</a></td>
<td>2659181</td>
</tr>
<tr>
<td></td>
<td>Cute raccoons begging for food</td>
<td><a href="http://www.youtube.com/watch?v=UKJYu2_5LBE">http://www.youtube.com/watch?v=UKJYu2_5LBE</a></td>
<td>2128244</td>
</tr>
<tr>
<td></td>
<td>Raccoon Steals a Carpet</td>
<td><a href="http://www.youtube.com/watch?v=S6u7ylr0zlg">http://www.youtube.com/watch?v=S6u7ylr0zlg</a></td>
<td>1958996</td>
</tr>
<tr>
<td></td>
<td>Oscar the Raccoon</td>
<td><a href="http://www.youtube.com/watch?v=dz_L9z09_s8">http://www.youtube.com/watch?v=dz_L9z09_s8</a></td>
<td>1737895</td>
</tr>
<tr>
<td></td>
<td>Raccoon Invades Kitchen</td>
<td><a href="http://www.youtube.com/watch?v=0DfGf4M3QZo">http://www.youtube.com/watch?v=0DfGf4M3QZo</a></td>
<td>1505765</td>
</tr>
<tr>
<td></td>
<td>Raccoon begging for food</td>
<td><a href="http://www.youtube.com/watch?v=q-EFXzHIFkk">http://www.youtube.com/watch?v=q-EFXzHIFkk</a></td>
<td>1280381</td>
</tr>
<tr>
<td></td>
<td>The Hillbilly Slide And One Mad Coon Starring Your</td>
<td><a href="http://www.youtube.com/watch?v=h8rnHXV7A8Q">http://www.youtube.com/watch?v=h8rnHXV7A8Q</a></td>
<td>1252841</td>
</tr>
<tr>
<td></td>
<td>Favorite Coon and Coonrippy!</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Raccoon and Cat Playing</td>
<td><a href="http://www.youtube.com/watch?v=KphBM5W6R8c">http://www.youtube.com/watch?v=KphBM5W6R8c</a></td>
<td>1149641</td>
</tr>
<tr>
<td></td>
<td>Baby Raccoon in the Tub</td>
<td><a href="http://www.youtube.com/watch?v=KphBM5W6R8c">http://www.youtube.com/watch?v=KphBM5W6R8c</a></td>
<td>844524</td>
</tr>
</tbody>
</table>

1. The URL to Raccoon vs Possum no longer works. The video can be seen at https://www.youtube.com/watch?v=IR6eij7foms&t=5s
APPENDIX E

Table XXIV
KEYWORDS ASSOCIATED TO THE SEVEN KELLERT CATEGORIES. ALTHOUGH SOME OF THESE KEYWORDS MAY APPEAR TO BE RELATED TO A SINGLE CATEGORY (E.G. THE KEYWORD ‘HUNT’ IS LIKELY RELATED TO THE DOMINIONISTIC KELLERT CATEGORY), THE WAY THAT IT IS USED WITHIN A COMMENT COULD DRASTICALLY ALTER HOW IT IS CATEGORIZED. THUS, THESE KEYWORDS ARE ONLY ORGANIZED ALPHABETICALLY BY ROW.

<table>
<thead>
<tr>
<th>adorable</th>
<th>aggressive</th>
<th>attack</th>
<th>attack</th>
<th>baby</th>
<th>beautiful</th>
<th>beauty</th>
<th>behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>bite</td>
<td>blood</td>
<td>conflict</td>
<td>cruel</td>
<td>cute</td>
<td>danger</td>
<td>dangerous</td>
<td>dead</td>
</tr>
<tr>
<td>disease</td>
<td>earth</td>
<td>eat</td>
<td>ecology</td>
<td>ecosystem</td>
<td>environment</td>
<td>environmental</td>
<td>fuck</td>
</tr>
<tr>
<td>fear</td>
<td>feed</td>
<td>food</td>
<td>free</td>
<td>fur</td>
<td>fuzzy</td>
<td>gross</td>
<td>gun</td>
</tr>
<tr>
<td>habitat</td>
<td>habitat</td>
<td>hate</td>
<td>hunt</td>
<td>hurt</td>
<td>kill</td>
<td>learn</td>
<td>love</td>
</tr>
<tr>
<td>mange</td>
<td>maul</td>
<td>meat</td>
<td>no</td>
<td>order</td>
<td>pet</td>
<td>PETA</td>
<td>predator</td>
</tr>
<tr>
<td>prey</td>
<td>rabies</td>
<td>scabies</td>
<td>scare</td>
<td>scum</td>
<td>service</td>
<td>shoot</td>
<td>shot</td>
</tr>
<tr>
<td>stupid</td>
<td>suffer</td>
<td>threat</td>
<td>trap</td>
<td>ugly</td>
<td>want</td>
<td>welfare</td>
<td>wild</td>
</tr>
</tbody>
</table>
This form confirms the *American Midland Naturalist* ("AMN"), a publication of the University of Notre Dame, hereby grants ___ Mason Fidino _________ the non-exclusive right to use the article listed below for personal, educational and/or academic purposes during the full term of the copyright.

**Name of article:** Habitat Dynamics of the Virginia Opossum in a Highly Urban Landscape

**Author(s):** Fidino M, Lehrer EW, and Magle SB

**Volume and date of first publication in AMN:** volume 175 in 2016

**How article will be used:** In Mason Fidino's PhD dissertation

The right granted herein is non-exclusive, non-assignable and limited in scope to use of the Article in hard copy, other print format or pdf, e-book and languages worldwide. This does not include the right to use the Article solely for financial gain. The University of Notre Dame warrants that it is the sole copyright owner of the publication, that it has complete authority and capacity to grant the rights contained herein, and that no other releases or permissions are necessary.

By granting this non-exclusive right to limited use of the Article, neither the University of Notre Dame or AMN release or revoke their copyright ownership or other interest of any kind or nature whatsoever in the Article.

**American Midland Naturalist**

By ____________________________

Date 1/1/17
VITA

Mason Fidino

Ecological Analyst, Urban Wildlife Institute
Director of Analytics, Urban Wildlife Information Network
Department of Conservation and Science, Lincoln Park Zoo
2001 N. Clark Street, Chicago, IL 60614
E-MAIL: mfidino@lpzoo.org
PHONE: 312-742-1670
FAX: 312-742-7220

EDUCATION:

Dissertation: Urban Wildlife Through Space and Time
Advisor: Dr. Joel Brown, Department of Biological Sciences

B.S.  Environmental Science, Terrestrial Ecology Emphasis.
Western Washington University, 2009.

RESEARCH EXPERIENCE:

PUBLICATIONS:


PROFESSIONAL EXPERIENCE:

**Ecological Analyst.** April 2015-Present. Urban Wildlife Institute, Lincoln Park Zoo.

In this role I work to develop techniques to better manage and analyze ecological data and assist the other science centers across the Lincoln Park Zoo with statistical analyses, data management, and computer programming. Furthermore, I have also developed statistics and computer programming workshops for science staff, which were widely attended. Finally, I helped initiate a large-scale citizen science project, Chicago Wildlife Watch (www.chicagowildlifewatch.org) and have written extensive software to verify and upload data to the project. Finally, I took the role of Director of Analytics of the Urban Wildlife Information Network (UWIN), which is the world’s first systematic multi-city urban wildlife monitoring network.

**Coordinator of Wildlife Management.** April 2012-April 2015, Urban Wildlife Institute, Lincoln Park Zoo.

Led all of the research of wildlife on zoo grounds, which included daily avian point counts, tracking relocated turtles in a newly restored pond habitat, on grounds rabbit management, arthropod surveys, and monitoring a nesting colony of state-endangered Black-crowned Night Herons (Nycticorax nycticorax). To do all of this work, I managed and trained teams of interns each year, mentored them on urban ecology, and helped them present their work to their peers at the Lincoln Park Zoo. Additionally, I collaborated extensively with the Lincoln Park Zoo’s education department on numerous projects to educate zoo visitors on wildlife conservation efforts throughout Chicago.

**Research Intern.** July 2011-April 2012, Urban Wildlife Institute, Lincoln Park Zoo.

Aided with field work and data entry for the Urban Wildlife Institute’s biodiversity monitoring survey, helped with research on zoo grounds, and created year-end permit reports.

SELECTED SCIENTIFIC PRESENTATIONS:


Fidino, M., Simonis, J.S.: Mesocarnivore dynamics in a highly fragmented, yet highly permeable urban landscape. ESA. Baltimore, Maryland. August, 2015.

GRANTS AND AWARDS:

North American Bluebird Society – Assessing the nesting success of urban cavity nesting birds - $600.00

SELECTED MEDIA COVERAGE:


ACADEMIC ORGANIZATIONS:

Society for Conservation Biology (2016 – Present)

Ecological Society of America (2015-Present)