UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

EPIDEMIOLOGY AND TRANSMISSION DYNAMICS OF ENVIRONMENTALLY MEDIATED INFECTIOUS DISEASES AT THE WILDLIFE-LIVESTOCK-HUMAN INTERFACE

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

HIMEL TALUKDER

Norman, Oklahoma

2026

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/342854

EPIDEMIOLOGY AND TRANSMISSION DYNAMICS OF ENVIRONMENTALLY MEDIATED INFECTIOUS DISEASES AT THE WILDLIFE-LIVESTOCK-HUMAN INTERFACE

A DISSERTATION APROVED FOR THE

DEPARTMENT OF GEOGRAPHY AND ENVIRONMENTAL SUSTAINABILITY

BY THE COMMITTEE CONSISTING OF

Dr. Anni Yang, Chair

Dr. Daniel Becker

Dr. Mike Wimberly

Dr. Nishan Bhattarai

© Copyright by HIMEL TALUKDER 2026

All Rights Reserved.

Acknowledgement

The completion of this dissertation marks the end of a remarkable journey, one shaped not only by intellectual curiosity and scientific inquiry, but also by the generosity, encouragement, and support of many individuals and institutions. While my name appears on the title page, this work is the product of countless contributions from those who guided, inspired, and believed in me along the way.

First and foremost, I offer my deepest gratitude to my advisor, Dr. Anni Yang. Her mentorship has been a steady light throughout my doctoral journey. She challenged me to think critically, encouraged me to pursue ambitious ideas, and provided unwavering support during both successes and setbacks. Her patience, wisdom, and confidence in my abilities have profoundly influenced my growth as a researcher and scholar. I am truly fortunate to have learned under her guidance.

I am sincerely grateful to my committee members, Dr. Daniel Becker, Dr. Michael Wimberly, and Dr. Nishan Bhattarai. Each of them contributed unique perspectives, thoughtful guidance, and constructive criticism that strengthened this dissertation in meaningful ways. Their willingness to share their expertise, ask difficult questions, and encourage deeper thinking enriched not only this work but also my broader understanding of science and research. It has been an honor to benefit from their mentorship.

I would also like to extend my heartfelt appreciation to Ryan Miller, Kimberly Pepin and Sarah Bevins of the United States Department of Agriculture. Their generosity in sharing data, knowledge, and expertise made several aspects of this research possible. I am also thanking Dr. Claudia Muñoz-Zanzi for her help in sharing the data with me. Scientific progress is built upon

collaboration, trust, and the open exchange of ideas, and I am deeply thankful for their support and partnership.

My sincere thanks go to the Department of Geography and Environmental Sustainability at the University of Oklahoma. The department provided not only the academic resources necessary for this work but also a community of scholars dedicated to advancing knowledge and addressing complex environmental and public health challenges. I am grateful for the opportunities, encouragement, and sense of belonging I found within this academic home.

Finally, and most importantly, I dedicate my deepest gratitude to my beloved wife, Nandita. There are no words sufficient to express what her presence has meant throughout this journey. Through long days of fieldwork, countless nights of writing, moments of uncertainty, and years spent pursuing distant goals, she remained my constant source of strength, patience, and love. She celebrated every success with me and carried me through every setback with unwavering faith. Her sacrifices often went unseen, yet they are woven into every page of this dissertation. When the path ahead seemed uncertain, she reminded me why I began. When the weight of the journey felt overwhelming, she gave me the courage to continue. Finally, to everyone who contributed, whether through guidance, collaboration, friendship, or encouragement, I offer my sincere gratitude.

Abstract

Infectious disease transmission is increasingly recognized as an ecological process shaped by interactions among hosts, environmental reservoirs, and landscapes. Many pathogens persist outside their hosts and are transmitted through shared environments, creating opportunities for disease spread within species, among species, and ultimately across wildlife–livestock–human systems. Understanding these interconnected transmission pathways is essential for improving disease surveillance, risk assessment, and prevention. This dissertation investigates the epidemiology and transmission dynamics of environmentally mediated infectious diseases across a continuum of ecological complexity, progressing from within-species transmission to cross-species transmission, and ultimately to transmission within a One Health framework.

Within-species transmission was examined using feral pigs (*Sus scrofa*) as a model system for environmentally mediated disease spread. High-resolution GPS telemetry data were used to quantify indirect contact opportunities under alternative assumptions regarding pathogen persistence, host behavior, and spatial use. Results demonstrated that transmission opportunities were highly heterogeneous and strongly influenced by behavioral and spatial constraints. Incorporating movement behavior into transmission models substantially altered contact network structure and epidemic projections, highlighting the importance of mechanistically defining indirect transmission pathways when evaluating disease dynamics in wildlife populations.

Cross-species transmission was investigated through avian influenza dynamics at the wild bird–feral pig interface across the contiguous United States. Using a Bayesian multi-species occupancy framework, I evaluated how environmental conditions, hydrological connectivity, wetland characteristics, and waterfowl communities influenced the probability of cross-species

avian influenza occupancy. Transmission probability was highest in environmentally connected wetland systems characterized by extensive wetlands, favorable moisture conditions, and abundant dabbling duck populations. Predicted hotspots were concentrated in Texas, the Gulf Coast, southern California, and portions of the Mississippi River Basin, demonstrating how landscape structure and environmental connectivity can facilitate pathogen exchange among host species.

The broader One Health transmission system was examined using leptospirosis as a model zoonosis linking wildlife reservoirs, domestic animals, environmental contamination, and human populations in southern Chile. Machine-learning analyses identified distinct drivers of human exposure across urban slum, semi-rural, and agricultural communities. Environmental conditions, animal reservoirs, contaminated water sources, and socio-demographic factors all contributed to disease risk, although their relative importance varied among landscapes. These findings emphasize that zoonotic disease transmission emerges from the interaction of ecological, environmental, and social processes operating simultaneously within coupled human–animal–environment systems.

Collectively, the findings of this dissertation demonstrate that environmentally mediated transmission operates across multiple ecological levels, from contacts among individuals within a host population, to pathogen exchange among species, and ultimately to human exposure within complex socio-ecological systems. Across all three disease systems, environmental conditions consistently shaped opportunities for pathogen persistence, transmission, and exposure. By integrating movement ecology, disease ecology, spatial epidemiology, and One Health approaches, this dissertation advances understanding of disease transmission at the wildlife–livestock–human interface and provides a framework for improving surveillance,

identifying transmission hotspots, and informing disease prevention strategies in an increasingly interconnected world.

Table of Contents

Acknowledgement ..... **iv**

Abstract ..... **vi**

List of Tables ..... **xii**

List of Figures ..... **xiii**

List of Abbreviations ..... **xvi**

Chapter 1 : Introduction **1**

1.1 Zoonotic Diseases 1

1.2 Environmentally Mediated Transmission 2

1.3 Cross-species Transmission and Wildlife Surveillance 4

1.4 Importance of One Health in Disease Management 5

1.5 Structure of the Dissertation 7

Chapter 2 : Estimating Environmental Transmission Risk from Host Movement Data **13**

2.1 Introduction 15

2.2 Methodology 20

2.2.1 Modeling Framework 20

2.2.2 Case Study 25

2.2.3 Model Evaluation and Analysis 31

2.3 Results 32

2.3.1 Contact Kernels 32

2.3.2 Network Structure and Matrices 34

2.3.3 Spatial Distribution 36

2.3.4 Disease Transmission 37

2.4 Discussion 39
2.4.1 Tradeoffs of Different Model Structures 39
2.4.2 Insights from the Case Study 42
2.4.3 Extensions, Future Directions, and Limitations 43
2.5 Conclusions 46
Chapter 3 : Predicting Cross-species Avian Influenza Transmission Between Migratory Waterfowl and Feral Pigs Through a Multi-species Occupancy Model to Inform Risk-based Surveillance in The United States 56
3.1 Introduction 58
3.2 Methodology 62
3.2.1 Study Areas and Spatial Unit 62
3.2.2 Surveillance Data 62
3.2.3 Covariates Data 64
3.2.4 Occupancy Models 68
3.3 Results 71
3.3.1 Infection Prevalence Across Host Species 71
3.3.2 Model Results 73
3.3.3 Predictions 76
3.4 Discussions 80
3.4.1 Importance of Integrated Multi-host Surveillance 80
3.4.2 Ecological and Environmental Drivers of Cross-species Transmission 81
3.4.3 Spatial Distribution and Surveillance Implications 83
3.4.4 Limitations and Future Directions 85
3.5 Conclusions 85
Chapter 4 : Drivers of Human Leptospirosis Exposure in Different Community Types of Southern Chile: A One Health Perspective 92

4.1 Introduction 94
4.2 Materials and Methods 97
4.2.1 Study area 97
4.2.2 Data collection 98
4.2.3 Extreme Gradient Boosting Model 103
4.3 Results 105
4.3.1 Seroprevalence of Leptospirosis 105
4.3.2 Urban Slum Community 108
4.3.3 Semi-rural Community 110
4.3.4 Farm Community 112
4.4 Discussions 114
4.5 Conclusions 120
Chapter 5 : Synthesis, Conclusions and Future Directions 126
5.1 Integrative Synthesis of Research Findings 126
5.2 Implications for Disease Ecology, Surveillance, and One Health 128
5.3 Limitations and Future Research Directions 131
5.4 Concluding Remarks 132
Supsplementary Files 136
Supplementary Files for Chapter 2 136
Supplementary files for Chapter 3 147

List of Tables

Table 2-1: Parameters used in the SEIR disease transmission model 31

Table 3-1: Description of covariates used in the study..... 65

Table 3-2: Spatial covariance functions, non-spatial, and approximation methods used in the occupancy model 69

Table 3-3: Performance comparison of the top-ranked non-spatial and spatial occupancy models across alternative covariance structures. Models were evaluated using average $\hat{R}$, effective sample size (ESS), and five-fold cross-validation deviance; results for all 123 candidate models are provided in Supplementary Table S2. 74

Table 4-1: Description and source of the variables included in the study. 101

Table 4-2: Parameters that were used to tune the models. 103

Table 4-3: Seroprevalence of Leptospirosis in different demographic and behavioral factors. 107

List of Figures

Figure 1-1: Structure of the dissertation 8

Figure 2-1: Methodological framework for capturing relevant environmentally mediated contact; (A) Quantifying potential exposure levels relating pathogen decay to the temporal lag in spatial overlap, (B) Estimating pathogen deposition levels based on time spent by infected individuals in a given location, (C) Adjusting exposure estimates by incorporating behavioral modes that influence pathogen deposition and/or acquisition and (D) integrating deposition-relevant space use and acquisition-relevant behaviors into a spatial representation of transmission risk. 19

Figure 2-2: Schematic Diagram of the environmentally mediated SEIR model with their respective PDEs. 29

Figure 2-3: Contact Kernel plots showing the total number of indirect contacts between each pair of wild pigs under four modeling for both 7-day and 30-day pathogen decay. The x- and y-axes indicate the unique animal IDs, with each cell representing the cumulative contact count between a pair. Darker colors correspond to higher number of contacts 34

Figure 2-4: Contact network structure under different contact models for 7-day and 30-day pathogen decay scenarios. Each node represents an individual wild pig, and each edge denotes amount of contacts. Edge thickness and color intensity correspond to the number of indirect contacts. 35

Figure 2-5: Spatial distribution of contact in landscape under different models; (A) pathogen decay only (7 days), (B) pathogen decay only (30days), (C) high use area (7 days), (D) high use area (7 days), (E) behavioral model (7 days), (F) behavioral model (7 days), (G) Integrated model (7 days), (H) Integrated model (30 days). Grid cells represent 30×30 m spatial units. Grey areas indicate zones with no contact, yellow indicates locations where contacts occurred, and red denotes high-intensity contact zones. 36

Figure 2-6: SEIR epidemic simulations for 7-day and 30-day pathogen decay durations under different contact modeling scenarios. For a specific contact model, each panel shows the mean trajectories (of 1000 runs) of the epidemiological compartments (Susceptible, Exposed, Infected, Recovered, and Environmental Load) over a period of 365 days. 38

Figure 3-1: Distribution of surveillance data collected from USDA between 2010-2015 for (a) Wild birds, (b) feral pigs. The data for 2012-2013 were missing in birds due to no surveillance records. 64

Figure 3-2: Infection prevalence among wild bird species and sex classes. (A) Species-specific prevalence estimates for bird species with a minimum sample size of 100 tested individuals. (B) Comparison of prevalence between female and male birds. 72

Figure 3-3: Infection prevalence among demographic groups of wild pigs. Sex-based prevalence estimates are shown for male and female pigs, while age-based prevalence estimates are presented for juvenile and adult pigs. 73

Figure 3-4: Occupancy, detection, and co-occurrence patterns of avian influenza in wild birds and feral pigs across HUC8 watershed–season sampling units. Posterior estimates from the Bayesian cross-species avian influenza occupancy model show host-specific occupancy and detection probabilities (A), the latent occupancy correlation between host groups (B), and the relationship between overall detection probability and sampling effort (C). 75

Figure 3-5: Posterior mean effect sizes and 95% credible intervals for environmental and ecological predictors of cross-species avian influenza occupancy from the best-performing model..... 77

Figure 3-6: Marginal effect curves for covariates included in the final cross-species avian influenza occupancy model predicting avian influenza occupancy probability. Black lines represent posterior mean occupancy probabilities across the observed range of each covariate while holding all other covariates constant at their median values, and gray shaded regions represent the corresponding 95% credible intervals. 78

Figure 3-7: Seasonal posterior predictions of cross-species avian influenza occupancy probability across HUC8 watersheds in the contiguous United States for (a) Fall, (b) Spring, (c) Summer, and (d) Winter. 79

Figure 3-8: Seasonal uncertainty (posterior standard deviation) associated with predicted cross-species avian influenza occupancy probabilities across HUC8 watersheds in the contiguous United States for (a) Fall, (b) Spring, (c) Summer, and (d) Winter. 79

Figure 4-1: Study area and distribution of leptospirosis seropositive (purple) and seronegative (blue) participants across the various community types indicated by shape (semi-rural: circle, farm: square, urban slums: triangle). 98

Figure 4-2: Leptospira positivity distribution in rodents (PCR kidney carriage positivity), most commonly sampled domestic animals (MAT seropositivity), water samples from the peri-domestic environment (PCR positivity for pathogenic species), and people (MAT seropositivity) (A) by the three community types and and (B) by the 12 communities, within each community type, Los Rios region, Chile Darker color in a bar represents the proportion of individuals with a MAT titer ≥1:400. 106

Figure 4-3: Variable Importance plot for the model results of the three community types, A) Urban slums, B) Semi-Rural, and C) Farm. Variables with more than a 2% importance frequency score are shown. The bars were color-coded for ease of interpretation of the overall trend in the predicted relationship between each variable and human exposure probability across the three different community types. Note that the relationship types

depicted here were derived visually from Figure 4-4, Figure 4-5 and Figure 4-6 and were simplified to overall positive or negative, but actual relationships are non-linear. 109

Figure 4-4: The partial dependency plots of covariates used in the model for urban slum communities. The gray shading area indicates the confidence intervals derived from the model iterations. The variable importance score is reported in the parentheses ((**A**–**N**): partial dependency of each variable in the model)..... 111

Figure 4-5: The partial dependency plots of covariates used in the model for semi-rural communities. The gray shading area indicates the confidence intervals derived from the model iterations. The variable importance score is reported in the parentheses ((**A**–**N**): partial dependency of each variable in the model)..... 112

Figure 4-6: The partial dependency plots of covariates used in the model for farm communities. The gray shading area indicates the confidence intervals derived from the model iterations. The variable importance score is reported in the parentheses ((**A**–**M**): partial dependency of each variable in the model)..... 113

List of Abbreviations

Abbreviation	Full Meaning
ABIR	Archbold Biological Station – Buck Island Ranch
APHIS	Animal and Plant Health Inspection Service
AUC	Area Under the Receiver Operating Characteristic Curve
CTMM	Continuous-Time Movement Modeling
CV	Cross Validation
COVID-19	Coronavirus Disease 2019
CWD	Chronic Wasting Disease
ESS	Effective Sample Size
GIS	Geographic Information System
GPS	Global Positioning System
GP	Gaussian Process
HMM	Hidden Markov Model
HPAI	Highly Pathogenic Avian Influenza
HUC	Hydrologic Unit Code
IAV	Influenza A Virus
JRC	Joint Research Centre (Global Surface Water Dataset)
LC	Land Cover
LSWI	Land Surface Water Index
MCMC	Markov Chain Monte Carlo
NDVI	Normalized Difference Vegetation Index
NFSP	National Feral Swine Program
NNGP	Nearest Neighbor Gaussian Process
NWDP	National Wildlife Disease Program
ROC	Receiver Operating Characteristic
SARS	Severe Acute Respiratory Syndrome
SD	Standard Deviation
SEIR	Susceptible–Exposed–Infectious–Recovered
USDA	United States Department of Agriculture
XGBoost	Extreme Gradient Boosting

Chapter 1 : Introduction

1.1 Zoonotic Diseases

Zoonotic infectious diseases, caused by pathogens that can transmit between animals and humans, represent a significant and growing global threat, with extensive consequences for public health, ecosystems, and environment (Bloom & Cadarette, 2019; Hauck, 2018; McArthur, 2019). These diseases, responsible for approximately 75% of all human infectious diseases, highlight the intricate interdependence between human, animal, and environmental health (Di Bari et al., 2023). Zoonotic pathogens, including viruses, bacteria, and parasites, can originate in wildlife reservoirs, domesticated animals, or the environment, often spilling over into human populations under favorable conditions (Rahman et al., 2020; Rohr et al., 2019; Zumla & Hui, 2019). Examples of such zoonotic diseases include COVID-19, severe acute respiratory syndrome (SARS), Ebola, Swine Flu (H1N1), Avian Influenza (H5N1), Anthrax, Leptospirosis, etc. (Hauck, 2018; Magouras et al., 2020; McArthur, 2019; Nii-Trebi, 2017).

The impacts of zoonotic diseases are not confined to health systems alone but extend to ecosystems and economies, disrupting multiple sectors. In ecosystems, diseases that spread among wildlife species can lead to population declines, loss of biodiversity, and altered ecological dynamics (Zumla & Hui, 2019). For instance, chytridiomycosis, caused by the Batrachochytrium fungus, has devastated amphibian populations worldwide, while rabies continues to impact mammalian wildlife, leading to cascading ecological consequences (Fisher et al., 2018; Fisher & Garner, 2020). Zoonotic diseases also threaten the stability of livestock

populations, causing production losses that affect food security and agricultural livelihoods. Livestock diseases such as avian influenza or foot-and-mouth disease require mass culling to prevent spread, further intensifying economic losses (Rohr et al., 2019).

In the human population, zoonotic outbreaks place a tremendous burden on healthcare systems, requiring significant investments in diagnostics, treatment, and preventive measures (Miller & Parent, 2012). The economic consequences are equally profound, as disruptions to tourism, trade, and agricultural markets arise due to quarantine measures, travel restrictions, and trade bans. Outbreaks like the COVID-19 pandemic have demonstrated the global scale of these disruptions, leading to severe declines in productivity, mass unemployment, and supply chain interruptions(Miller & Parent, 2012; Rahman et al., 2020). Beyond direct economic losses, zoonotic diseases exacerbate poverty, widen inequality, and undermine long-term development goals, particularly in vulnerable communities where access to healthcare and sanitation is limited (Magouras et al., 2020). To tackle the impact of infectious diseases, we might need some kind of interdisciplinary context that can help us to better understand the whole transmission process.

1.2 Environmentally Mediated Transmission

Pathogen transmission occurs through contact processes, which determine when, where, and how susceptible hosts encounter infectious agents (Craft, 2015; White et al., 2017). There are various ways this transmission can occur, including direct contact, where pathogens move directly between infected and susceptible hosts through physical contact or proximity (Yang et al., 2021), or transmitted among hosts via media such as invertebrates and environment like soil, water or fomites (i.e., indirect contacts). Although direct transmission and vector-borne transmission process has been extensively studied, environmentally driven transmission remains comparatively underexplored, even though many important zoonotic pathogens survive long in

the environment and create a persistent environmental reservoir of transmission. These wildlife reservoirs shape a spatially explicit landscape where exposure and pathogen transmission depend on a variety of complex processes such as pathogen biology, meteorological factors, host movement, host behavior, their habitat structure and temporal pattern of movement (Yang et al., 2023). A successful transmission, in this case, will be the interplay in a specific way that aligns all of the above factors.

Environmentally mediated transmission arises from the interplay among pathogen shedding, environmental persistence, and subsequent host exposure (Manlove et al., 2022; Wilber et al., 2022). In these systems, pathogens are released into shared resources of soil, water, or vegetation, where pathogen can sustain for an extended period of time, depending on environmental conditions such as temperature, rainfall, radiation etc. (Hellberg & Chu, 2016; Zhang et al., 2024). Thus, the risk of transmission is not simply a function of pathogen presence but also depends on the frequency of host encounters with contaminated sites over time. The emphasis is no longer on host-host interactions but on how the landscape itself mediates a pattern of exposure, with the environment as the mediator in the transmission process.

Pathogen acquisition and environmental contamination are linked by host movement in a continuous process (Daversa et al., 2017; Engering et al., 2013). Hosts contribute to spatially heterogeneous patterns of contamination as they traverse landscapes, and they are more likely to come into contact with infectious material when they travel through the same areas repeatedly (Daversa et al., 2017). Thus, even in low host densities, environmental elements that are frequently used, like water sources, feeding areas, or travel corridors, can serve as persistent transmission nodes (Daversa et al., 2017). Because these factors jointly determine where exposure risk accumulates and how localized transmission hotspots emerge, understanding

transmission in these systems would help us understand how host and pathogen both interact with the environment and ultimately determine the local hotspot and spatially explicit contact locations.

1.3 Cross-species Transmission and Wildlife Surveillance

Transmission of pathogens within different host species is termed as cross-species transmission (Choudhury et al., 2022; Parrish et al., 2008). Cross-species transmission of pathogens like avian influenza viruses is rarely caused by prolonged direct contact between reservoir and recipient hosts. Rather, it frequently results from indirect exposure pathways influenced by landscape structure, ecological context, and environmental persistence (Blagodatski et al., 2021). Wetlands, water bodies, agricultural lands, and other interface environments are examples of shared habitats that can aid in the spread of pathogens by serving as intermediaries where they are deposited, remain, and eventually come into contact with susceptible hosts (Blagodatski et al., 2021; Hénaux & Samuel, 2011). Because of this, cross-species transmission is often sporadic, spatially heterogeneous, and temporally decoupled from the presence of infectious reservoir hosts, making spillover events challenging to directly observe but extremely instructive regarding host adaptation and pathogen ecology.

These transmission pathways are largely shaped by environmental and landscape features, which have a dynamic influence. Climate variability modifies the environmental suitability for viral survival, while seasonal variations in water availability, flooding, and wetland extent can change where pathogens accumulate and how long they persist (Gass Jr et al., 2023; Morin et al., 2018). Exposure pathways change over time as a result of these temporal changes interacting with host movement and habitat use (Ferenczi et al., 2021). Therefore, areas where animal activity is concentrated, like wetlands, floodplains, or resource-rich agricultural

interfaces, may serve as temporary or recurrent focal points for interspecies indirect transmission.

A strong foundation for improving wildlife surveillance in such intricate, landscape-mediated transmission systems is provided by spatial data science. Surface water dynamics, vegetation phenology, land-use patterns, climatic variability, and other environmental features pertinent to avian influenza ecology can be continuously characterized through remote sensing, and GIS makes it easier to integrate these data with wildlife surveillance records over time and space. These methods enable the identification of high-risk landscapes, reconstruction of likely exposure pathways, and quantification of spillover risk by connecting spatially explicit representations of habitat use and environmental conditions with pathogen detection data from wild birds and feral pigs. Thus, spatially informed surveillance improves the interpretability and predictive value of wildlife monitoring initiatives and offers a vital basis for avian influenza risk assessment and spatial modeling.

1.4 Importance of One Health in Disease Management

Despite the inherently interconnected nature of zoonotic diseases, much of the historical research and policy framework has treated human health, animal health, and environmental processes as largely separate domains. This fragmented approach can miss important ecological routes through which pathogens may move from wildlife to domestic animals to people. As a result, interventions that developed from only one sector may be insufficient or unsustainable if the broader ecological systems are not taken into consideration.

The one health framework addresses these challenges by promoting a unified and interdisciplinary approach to understand and manage the process of zoonoses (Adisasmito et al., 2022; Hayman et al., 2023). This concept recognizes that transmission dynamics emergers as an

interactive process among human, animals, and the environment, and it emphasizes effective control of zoonotic diseases require integration of all the components simultaneously (Coker et al., 2011; Erkyihun & Alemayehu, 2022). By integrating and connecting concepts of ecology, epidemiology, veterinary medicine, environmental science, and public health, one health offers a comprehensive view of the factors that drive pathogen emergence, persistence and spread (Hayman et al., 2023; Lebov et al., 2017).

The one health framework has several advantages towards surveillance and intervention strategies (Bordier et al., 2020). Where the spatial information on host distribution, pathogen detection, and environmental conditions is integrated in a One Health approach, surveillance programs can be designed to focus on areas where risk of leptospirosis is most likely to emerge, such as wetlands close to human settlements, zones of high livestock density in agricultural use, or urban–rural interfaces where rodents, domestic animals, and humans commonly interact (Bordier et al., 2020). Data from these surveillance systems can feed into spatially explicit models that estimate the extent of environmental contamination, thereby identifying emerging hotspots and informing decisions about where to direct limited resources.

Similarly, the control strategies informed by the One Health approach can intervene in several places along this transmission cycle of leptospirosis (Bierque et al., 2020). Integrated approaches will address not only vaccination of livestock or treatment of human cases but also rodent reservoirs, environments contaminated after events of increased rainfall or flooding, and water and sanitation in a vulnerable community. These strategies reflect understanding that pathogen persistence, host behaviors, and environmental conditions together interact to shape overall risks.

1.5 Structure of the Dissertation

This dissertation is designed to fill the gap in linking geospatial science and One Health research. The overarching objective of this dissertation is to develop and apply spatially explicit, integrative modeling frameworks and incorporate real-time earth observations that quantify environmentally mediated zoonotic disease transmission across human, animal, and environmental systems to support disease management and surveillance decision-making. Chapter 1 builds a mechanistic model which formalizes how host movements and behavior generate environmental contamination, how contamination decays over time and how hosts re-encounter contaminated locations to estimate the spatially explicit environmental transmission risk. Chapter 2 presents how dynamic wetlands and environmental features can influence the transmission of avian influenza in a cross-species environment. Using dynamic environmental variables, this chapter will build a Bayesian hierarchical multi-species occupancy model to predict the cross-species risk surface to facilitate surveillance decisions. Chapter 3 integrates serological and surveillance data from human, animal and environment to find out the drivers of Leptospira exposures in urban slums, semi-rural communities and farm settings, showing the importance of One Health concept in disease control and intervention. The significance of this dissertation will be the integration of mechanistic, hierarchical, and machine learning models into a comprehensive one health modeling pipeline for environmentally transmitted and multi-host pathogens. On a conceptual level, it provides insight into how environmental spatial heterogeneity and host spatial behavior contribute to environmental transmission risk and how this risk then propagates in a multi-host community to humans. On a methodology level, it enhances modeling capabilities for mechanistic environmental risk modeling, multi-host species distribution modeling with uncertain detection, and One Health modeling of risk factors.

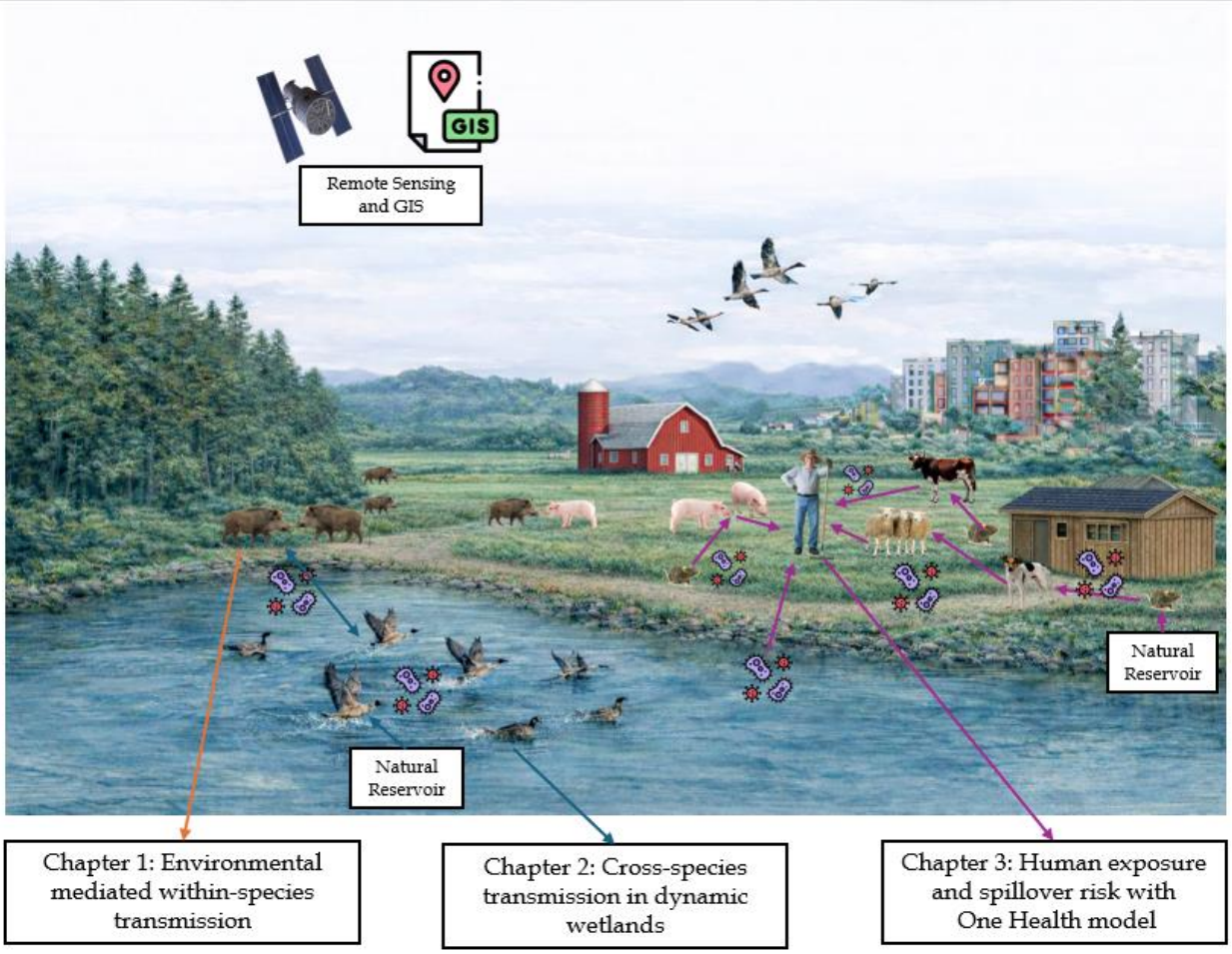


Figure 1-1: Structure of the dissertation

References

Adisasmito, W. B., Almuhairi, S., Behravesh, C. B., Bilivogui, P., Bukachi, S. A., Casas, N., Becerra, N. C., Charron, D. F., Chaudhary, A., & Zanella, J. R. C. (2022). One Health: A new definition for a sustainable and healthy future. *PLoS pathogens*, *18*(6), e1010537.

Alexander, A. D. (1960). The distribution of leptospirosis in Latin America. *Bull World Health Organ*, *23*(1), 113-125.

Bierque, E., Thibeaux, R., Girault, D., Soupé-Gilbert, M. E., & Goarant, C. (2020). A systematic review of Leptospira in water and soil environments. *PLoS One*, *15*(1), e0227055. https://doi.org/10.1371/journal.pone.0227055

Blagodatski, A., Trutneva, K., Glazova, O., Mityaeva, O., Shevkova, L., Kegeles, E., Onyanov, N., Fede, K., Maznina, A., Khavina, E., Yeo, S.-J., Park, H., & Volchkov, P. (2021). Avian Influenza in Wild Birds and Poultry: Dissemination Pathways, Monitoring Methods, and Virus Ecology. *Pathogens*, *10*(5), 630. https://www.mdpi.com/2076-0817/10/5/630

Bloom, D. E., & Cadarette, D. (2019). Infectious Disease Threats in the Twenty-First Century: Strengthening the Global Response [Review]. *Frontiers in Immunology*, *10*. https://doi.org/10.3389/fimmu.2019.00549

Bolin, C. A., & Koellner, P. (1988). Human-to-Human Transmission of Leptospira interrogans by Milk. *Journal of Infectious Diseases*, *158*(1), 246-247. https://doi.org/10.1093/infdis/158.1.246

Bordier, M., Uea-Anuwong, T., Binot, A., Hendrikx, P., & Goutard, F. L. (2020). Characteristics of One Health surveillance systems: A systematic literature review. *Preventive Veterinary Medicine*, *181*, 104560. https://doi.org/https://doi.org/10.1016/j.prevetmed.2018.10.005

Bradley, E. A., & Lockaby, G. (2023). Leptospirosis and the Environment: A Review and Future Directions. *Pathogens*, *12*(9), 1167. https://www.mdpi.com/2076-0817/12/9/1167
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10538202/pdf/pathogens-12-01167.pdf

Choudhury, P. R., Saha, T., Goel, S., Shah, J. M., & Ganjewala, D. (2022). Cross-species virus transmission and its pandemic potential. *Bulletin of the National Research Centre*, *46*(1), 18.

Coker, R., Rushton, J., Mounier-Jack, S., Karimuribo, E., Lutumba, P., Kambarage, D., Pfeiffer, D. U., Stärk, K., & Rweyemamu, M. (2011). Towards a conceptual framework to support one-health research for policy on emerging zoonoses. *The Lancet Infectious Diseases*, *11*(4), 326-331. https://doi.org/10.1016/S1473-3099(10)70312-1

Costa, F., Hagan, J. E., Calcagno, J., Kane, M., Torgerson, P., Martinez-Silveira, M. S., Stein, C., Abela-Ridder, B., & Ko, A. I. (2015). Global Morbidity and Mortality of Leptospirosis: A Systematic Review. *PLoS Negl Trop Dis*, *9*(9), e0003898. https://doi.org/10.1371/journal.pntd.0003898

Craft, M. E. (2015). Infectious disease transmission and contact networks in wildlife and livestock. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1669). https://doi.org/10.1098/rstb.2014.0107

Crecelius, E. M., & Burnett, M. W. (2020). Leptospirosis. *Journal of Special Operations Medicine*, *20*(4), 121. https://doi.org/10.55460/8ybj-0dlp

Daversa, D., Fenton, A., Dell, A., Garner, T., & Manica, A. (2017). Infections on the move: how transient phases of host movement influence disease spread. *Proceedings of the Royal Society B: Biological Sciences*, *284*(1869), 20171807.

Di Bari, C., Venkateswaran, N., Fastl, C., Gabriël, S., Grace, D., Havelaar, A. H., Huntington, B., Patterson, G. T., Rushton, J., & Speybroeck, N. (2023). The global burden of neglected zoonotic diseases: Current state of evidence. *One Health*, *17*, 100595.

Ellis, W. A. (2015). Animal leptospirosis. *Curr Top Microbiol Immunol*, *387*, 99-137. https://doi.org/10.1007/978-3-662-45059-8_6

Engering, A., Hogerwerf, L., & Slingenbergh, J. (2013). Pathogen–host–environment interplay and disease emergence. *Emerging microbes & infections*, *2*(1), 1-7.

Erkyihun, G. A., & Alemayehu, M. B. (2022). One Health approach for the control of zoonotic diseases. *Zoonoses*, *2*(1), 963.

Ferenczi, M., Beckmann, C., & Klaassen, M. (2021). Rainfall driven and wild-bird mediated avian influenza virus outbreaks in Australian poultry. *BMC Veterinary Research*, *17*(1), 306.

Fisher, C. R., Streicker, D. G., & Schnell, M. J. (2018). The spread and evolution of rabies virus: conquering new frontiers. *Nature Reviews Microbiology*, *16*(4), 241-255.

Fisher, M. C., & Garner, T. W. (2020). Chytrid fungi and global amphibian declines. *Nature Reviews Microbiology*, *18*(6), 332-343.

Gass Jr, J. D., Hill, N. J., Damodaran, L., Naumova, E. N., Nutter, F. B., & Runstadler, J. A. (2023). Ecogeographic drivers of the spatial spread of highly pathogenic avian influenza outbreaks in Europe and the United States, 2016–Early 2022. *International Journal of Environmental Research and Public Health*, *20*(11), 6030.

Guerra, M. A. (2013). Leptospirosis: Public health perspectives. *Biologicals*, *41*(5), 295-297. https://doi.org/10.1016/j.biologicals.2013.06.010

Haake, D. A., & Levett, P. N. (2015). Leptospirosis in Humans. In (pp. 65-97). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-45059-8_5

Harrison, N. A., & Fitzgerald, W. R. (1988). Leptospirosis--can it be a sexually transmitted disease? *Postgrad Med J*, *64*(748), 163-164. https://doi.org/10.1136/pgmj.64.748.163

Hauck, K. (2018). The Economics of Infectious Diseases. In: Oxford University Press.

Hayman, D. T. S., Adisasmito, W. B., Almuhairi, S., Behravesh, C. B., Bilivogui, P., Bukachi, S. A., Casas, N., Becerra, N. C., Charron, D. F., Chaudhary, A., Ciacci Zanella, J. R., Cunningham, A. A., Dar, O., Debnath, N., Dungu, B., Farag, E., Gao, G. F., Khaitsa, M., Machalaba, C.,…Koopmans, M. (2023). Developing One Health surveillance systems. *One Health*, *17*, 100617. https://doi.org/https://doi.org/10.1016/j.onehlt.2023.100617

Hellberg, R. S., & Chu, E. (2016). Effects of climate change on the persistence and dispersal of foodborne bacterial pathogens in the outdoor environment: A review. *Critical reviews in microbiology*, *42*(4), 548-572.

Hénaux, V., & Samuel, M. D. (2011). Avian influenza shedding patterns in waterfowl: implications for surveillance, environmental transmission, and disease spread. *The Journal of Wildlife Diseases*, *47*(3), 566-578.

Lebov, J., Grieger, K., Womack, D., Zaccaro, D., Whitehead, N., Kowalcyk, B., & MacDonald, P. D. M. (2017). A framework for One Health research. *One Health*, *3*, 44-50. https://doi.org/https://doi.org/10.1016/j.onehlt.2017.03.004

Lelu, M., Muñoz-Zanzi, C., Higgins, B., & Galloway, R. (2015). Seroepidemiology of leptospirosis in dogs from rural and slum communities of Los Rios Region, Chile. *BMC Veterinary Research*, *11*(1), 31. https://doi.org/10.1186/s12917-015-0341-9

Luna, J., Salgado, M., Tejeda, C., Moroni, M., & Monti, G. (2020). Assessment of Risk Factors in Synanthropic and Wild Rodents Infected by Pathogenic Leptospira spp. Captured in Southern Chile. *Animals*, *10*(11), 2133. https://www.mdpi.com/2076-2615/10/11/2133 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7697743/pdf/animals-10-02133.pdf

Magouras, I., Brookes, V. J., Jori, F., Martin, A., Pfeiffer, D. U., & Dürr, S. (2020). Emerging Zoonotic Diseases: Should We Rethink the Animal–Human Interface? [Opinion]. *Frontiers in Veterinary Science*, *7*. https://doi.org/10.3389/fvets.2020.582743

Manlove, K., Wilber, M., White, L., Bastille-Rousseau, G., Yang, A., Gilbertson, M. L. J., Craft, M. E., Cross, P. C., Wittemyer, G., & Pepin, K. M. (2022). Defining an epidemiological landscape that connects movement ecology to pathogen transmission and pace-of-life. *Ecology Letters*, *25*(8), 1760-1782. https://doi.org/https://doi.org/10.1111/ele.14032

McArthur, D. B. (2019). Emerging Infectious Diseases. *Nurs Clin North Am*, *54*(2), 297-311. https://doi.org/10.1016/j.cnur.2019.02.006

Miller, D. A., Wilson, M. A., & Beran, G. W. (1991). Relationships between prevalence of Leptospira interrogans in cattle, and regional, climatic, and seasonal factors. *Am J Vet Res*, *52*(11), 1766-1768.

Miller, G. Y., & Parent, K. (2012). The economic impact of high consequence zoonotic pathogens: why preparing for these is a wicked problem. *Journal of Reviews on Global Economics*, *1*, 47.

Montes, V., & Monti, G. (2021). Pathogenic Leptospira spp. Seroprevalence and Herd-Level Risk Factors Associated with Chilean Dairy Cattle. *Animals*, *11*(11), 3148. https://www.mdpi.com/2076-2615/11/11/3148 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8614305/pdf/animals-11-03148.pdf

Morin, C. W., Stoner-Duncan, B., Winker, K., Scotch, M., Hess, J. J., Meschke, J. S., Ebi, K. L., & Rabinowitz, P. M. (2018). Avian influenza virus ecology and evolution through a climatic lens. *Environment international*, *119*, 241-249.

Muñoz-Zanzi, C., Mason, M., Encina, C., Astroza, A., & Romero, A. (2014). Leptospira Contamination in Household and Environmental Water in Rural Communities in Southern Chile. *International Journal of Environmental Research and Public Health*, *11*(7), 6666-6680. https://doi.org/10.3390/ijerph110706666

Muñoz-Zanzi, C., Mason, M., Encina, C., Gonzalez, M., & Berg, S. (2014). Household characteristics associated with rodent presence and Leptospira infection in rural and urban communities from Southern Chile. *Am J Trop Med Hyg*, *90*(3), 497-506. https://doi.org/10.4269/ajtmh.13-0334

Mwachui, M. A., Crump, L., Hartskeerl, R., Zinsstag, J., & Hattendorf, J. (2015). Environmental and Behavioural Determinants of Leptospirosis Transmission: A Systematic Review. *PLOS Neglected Tropical Diseases*, *9*(9), e0003843. https://doi.org/10.1371/journal.pntd.0003843

Nii-Trebi, N. I. (2017). Emerging and Neglected Infectious Diseases: Insights, Advances, and Challenges. *BioMed Research International*, *2017*(1), 5245021. https://doi.org/https://doi.org/10.1155/2017/5245021

Pappas, G., Papadimitriou, P., Siozopoulou, V., Christou, L., & Akritidis, N. (2008). The globalization of leptospirosis: worldwide incidence trends. *Int J Infect Dis*, *12*(4), 351-357. https://doi.org/10.1016/j.ijid.2007.09.011

Parrish, C. R., Holmes, E. C., Morens, D. M., Park, E.-C., Burke, D. S., Calisher, C. H., Laughlin, C. A., Saif, L. J., & Daszak, P. (2008). Cross-species virus transmission and the emergence of new epidemic diseases. *Microbiology and Molecular Biology Reviews*, *72*(3), 457-470.

Rahman, M. T., Sobur, M. A., Islam, M. S., Ievy, S., Hossain, M. J., El Zowalaty, M. E., Rahman, A. T., & Ashour, H. M. (2020). Zoonotic Diseases: Etiology, Impact, and Control. *Microorganisms*, *8*(9), 1405. https://www.mdpi.com/2076-2607/8/9/1405

Rohr, J. R., Barrett, C. B., Civitello, D. J., Craft, M. E., Delius, B., DeLeo, G. A., Hudson, P. J., Jouanard, N., Nguyen, K. H., Ostfeld, R. S., Remais, J. V., Riveau, G., Sokolow, S. H., & Tilman, D. (2019). Emerging human infectious diseases and the links to global food production. *Nature Sustainability*, *2*(6), 445-456. https://doi.org/10.1038/s41893-019-0293-3

Romero, E. C., Bernardo, C. C., & Yasuda, P. H. (2003). Human leptospirosis: a twenty-nine-year serological study in São Paulo, Brazil. *Rev Inst Med Trop Sao Paulo*, *45*(5), 245-248. https://doi.org/10.1590/s0036-46652003000500002

White, L. A., Forester, J. D., & Craft, M. E. (2017). Using contact networks to explore mechanisms of parasite transmission in wildlife. *Biological Reviews*, *92*(1), 389-409. https://doi.org/https://doi.org/10.1111/brv.12236

Wilber, M. Q., Yang, A., Boughton, R., Manlove, K. R., Miller, R. S., Pepin, K. M., & Wittemyer, G. (2022). A model for leveraging animal movement to understand spatio-temporal disease dynamics. *Ecology Letters*, *25*(5), 1290-1304. https://doi.org/https://doi.org/10.1111/ele.13986

Yang, A., Boughton, R. K., Miller, R. S., Wight, B., Anderson, W. M., Beasley, J. C., VerCauteren, K. C., Pepin, K. M., & Wittemyer, G. (2021). Spatial variation in direct and indirect contact rates at the wildlife-livestock interface for informing disease management. *Preventive Veterinary Medicine*, *194*, 105423. https://doi.org/https://doi.org/10.1016/j.prevetmed.2021.105423

Yang, A., Wilber, M. Q., Manlove, K. R., Miller, R. S., Boughton, R., Beasley, J., Northrup, J., VerCauteren, K. C., Wittemyer, G., & Pepin, K. (2023). Deriving spatially explicit direct and indirect interaction networks from animal movement data. *Ecology and Evolution*, *13*(3), e9774. https://doi.org/https://doi.org/10.1002/ece3.9774

Zamora, J., Riedemann, S., Montecinos, M. I., & Cabezas, X. (1990). [Serological survey of human leptospirosis in a high risk population in Chile]. *Rev Med Chil*, *118*(3), 247-252. (Encuesta serológica de leptospirosis humana en ocupaciones de alto riesgo en Chile.)

Zhang, L., Lv, C., Guo, W., & Li, Z. (2024). Temperature and humidity as drivers for the transmission of zoonotic diseases. *Animal Research and One Health*, *2*(3), 323-336.

Zumla, A., & Hui, D. S. C. (2019). Emerging and Reemerging Infectious Diseases: Global Overview. *Infectious Disease Clinics*, *33*(4), xiii-xix. https://doi.org/10.1016/j.idc.2019.09.001

Chapter 2 : Estimating Environmental Transmission Risk from Host Movement Data

Keywords: environmental transmission, host behavior, host movement, indirect contact, pathogen persistence

Published in Wiley's Ecology and Evolution, on November 30, 2025

(DOI: https://doi.org/10.1002/ece3.72571)

Abstract

Environmentally mediated transmission is a critical pathway for disease spread, particularly for pathogens with environmental persistence. However, distinguishing which overlapping host movements result in transmission remains difficult, as most mobility models assume homogeneous exposure and overlook key spatial and behavioral factors. We extend previous models of indirect contact to account for variation in host movement behavior that could affect the probability of indirect transmission and examine these effects on transmission dynamics. We considered four models with different assumptions about the influence of movement behavior on indirect contact and transmission probability. The pathogen decay model includes all overlaps occurring within the pathogen's viability window. The high-use area model restricts transmission to frequently visited locations, assuming greater pathogen deposition where hosts repeatedly congregate. The behavioral model filters contacts based on host behaviors relevant to deposition or acquisition. The integrated model incorporates both spatial and behavioral constraints. We applied these models to GPS telemetry data from wild pigs in Florida and simulated disease dynamics using SEIR models with environmental transmission for Influenza A and *Brucella suis*, representing pathogens with short and long environmental persistence, respectively. Compared to pathogen decay model, total contact numbers decreased by 0.3% in high-use model, 70.3% in behavioral model, and 70.4% in integrated model under the short-lived condition. Edge density declined from 0.60 to 0.50, transitivity from 0.81 to 0.77, and assortativity from 0.43 to 0.17. R_0 declined from 2.42 ± 0.47 to 2.06 ± 0.39 (high-use), 1.56 ± 0.42 (behavioral), and 1.35 ± 0.40 (integrated), with a lag of approximately 21 days in peak incidence. The long-lived scenarios, were quantitively similar to those of short-lived scenarios. These results underscore that assumptions about how indirect contact from movement

data maps to pathogen transmission significantly influence network configuration and epidemic projections. Our findings highlight the importance of mechanistically defining indirect transmission in disease modeling.

2.1 Introduction

Pathogens and parasites can spread through multiple routes of exposure, including contact among hosts (Craft, 2015; Parratt, 2016; Gwenzi et al., 2022). Direct contact transmission occurs when hosts engage in physical contact (White et al., 2017; VanderWaal & Ezwana, 2016). Pathogen transmission can also occur indirectly through vectors and fomites, mediated through environmental factors or mechanical vectors (Fofana & Hurford, 2017). Many pathogens, such as *Mycobacterium bovis, Brucella spp.,* and *influenza A viruses,* can survive in the environment, and be transmitted via host contact with contaminated surfaces, soil, water, or other environmental media (Pandey et al., 2024; Suh et al., 2024). This process, referred to as environmentally mediated transmission, can be a primary transmission pathway for those disease systems and plays an important role in facilitating pathogen spread among social groups of hosts. Despite its recognized importance in disease ecology, environmentally-mediated transmission remains less well-characterized than direct transmission and vector borne transmission, largely due to the challenges of identifying environmentally-mediated contacts (Collier et al., 2022).

Historically, contact networks have been quantified with a variety of direct and indirect observational and technological tools (Arthur et al., 2017). Camera trap data are useful for determining species composition, group size, and coarse contact patterns (Herraiz et al., 2024a; Smith et al., 2020; Kukielka et al., 2013). However, they are limited in terms of spatial and temporal continuity, and they might not capture the individual-level movements required to

model environmental exposure to pathogens (Triguero-Ocaña et al., 2021; Egan et al., 2023). Proximity loggers can track individuals' physical closeness at a continuous temporal scale (Yang et al., 2021; Pepin et al., 2021); however, they are limited to capturing environmentally-mediated contacts only via known point-sources when the stationary loggers were deployed at pre-defined locations (Kelt et al., 2019).

Wildlife telemetry offers rich spatiotemporal movement data and has been increasingly used to study individual–environment interactions (Noonan et al., 2021). The ubiquity of movement data in wildlife studies provides the opportunity to capture spatially-explicit direct and indirect contacts. Specifically, direct contact is often defined as colocation at the same time, while the environmentally-mediated contact as colocation at segregated time (Kenward et al., 1998; Long et al., 2022). Nevertheless, not all instances of colocation at segregated time are epidemiologically relevant (Dougherty et al., 2018). For instance, a host simply passing through a location previously visited by another host may rarely result in pathogen transmission. The success of environmentally-mediated transmission can be affected by various factors, including pathogen persistence in the environment, host-immune response, level of environmental contamination, and host acquisition behavior (Rees et al., 2021).

One of the principal challenges in defining environmentally-mediated contact is distinguishing between incidental spatial overlap and contacts that pose a meaningful risk for pathogen transmission (Manlove et al., 2022). Thus, identifying relevant environmentally-mediated contact requires the consideration of the factors that trigger and affect the transmission. For example, in environmentally persistent diseases like brucellosis, an infected host may deposit pathogens in a shared environment where they remain viable for days to weeks, allowing susceptible individuals to acquire the pathogen upon subsequent visitation (Salman &

Steneroden, 2022). However, if a susceptible host arrives at the site after the pathogen has decayed beyond infectious viability, the contact is no longer relevant (Daversa et al., 2017; Richardson & Gorochowski, 2015). Additionally, behavioral heterogeneity among hosts can add further complexity (Egan et al., 2023; Clontz et al., 2021). Certain behaviors, such as foraging, resting, drinking, or wallowing, may increase the likelihood of pathogen deposition and acquisition, while transient movement through contaminated areas may not provide sufficient exposure for transmission (Dougherty et al., 2022; Binning et al., 2017). Existing research (Yang et al., 2023; Wilber et al., 2022), tend to presume that all spatiotemporal overlaps correspond to equivalent transmission risk. This is a simplification of environmentally-mediated transmission that overlooks behavioral context and pathogen persistence. Our framework fills this gap by adding these dimensions and adjusting their effect on disease spread.

Previous research on indirect contact that relies on spatiotemporal overlap frequently makes the assumption that every instance of non-simultaneous co-location contributes equally to the risk of transmission (Dougherty et al., 2018; Yang et al., 2023). However, this method ignores epidemiological factors that influence pathogen deposition and acquisition likelihood, such as host movement behaviors and pathogen survival duration. These oversimplified overlap models may therefore greatly overestimate the risk of indirect transmission, particularly for pathogens that need particular behavioral or environmental factors to spread successfully. To address these limitations, an integrative approach is required to systematically quantify relevant environmentally-mediated contacts by incorporating host movement data, pathogen survival dynamics, and behavioral ecology. This modeling approach is a continuation of earlier foundational work that utilized movement data to construct spatio-temporally varied exposure networks (Wilber et al., 2022; Yang et al., 2023), which has recently been compared against

baseline proximity-based measures of risk in empirical disease networks (Herraiz et al., 2024b). They have illustrated how contact modeling frameworks can provide us with valuable information on transmission and force of infection based on pathogen decay and host movement data.

However, previous uses were limited in scope. For example, they often primarily focused on pathogen decay without explicitly parameterizing deposition or acquisition processes in empirical systems, or they lacked a systematic investigation of how varying mechanistic assumptions influence network structure or epidemic dynamics (Wilber et al., 2022; Yang et al., 2023, Herraiz et al., 2024b). In the present study, we expand on this framework to explore how mechanistic filters, such as behavioral state, pathogen decay dynamics, high-use area assignment, and spatial thresholds, interact to build exposure networks and epidemic impacts. Because of the critical role of pathogen ecology in translating indirect contact, this modular comparison is best suited to environmentally persistent pathogens, where contaminating and acquiring assumptions largely predominate transmission dynamics. The objective of this study is to develop a methodological framework to test hypotheses about the mechanisms that drive relevant indirect contact based on host movement data under different transmission scenarios. Specifically, we integrate host behaviors, pathogen survival times, and environmental contaminations to estimate dyadic exposure levels related to disease transmission risk. This framework will allow for more precise characterization of environmentally-mediated contact events distinguishing those that are likely to facilitate pathogens (e.g., viruses, bacteria, and fungi) from those that are not. Contrasting outputs from these different models allows identification of conserved or sensitive transmission properties subject to these different processes. By integrating environmental dynamics with classical compartmental transitions, this

framework not only enables retrospective consideration of environmental transmission but also comparative evaluation across host-pathogen systems with contrasting ecological features.

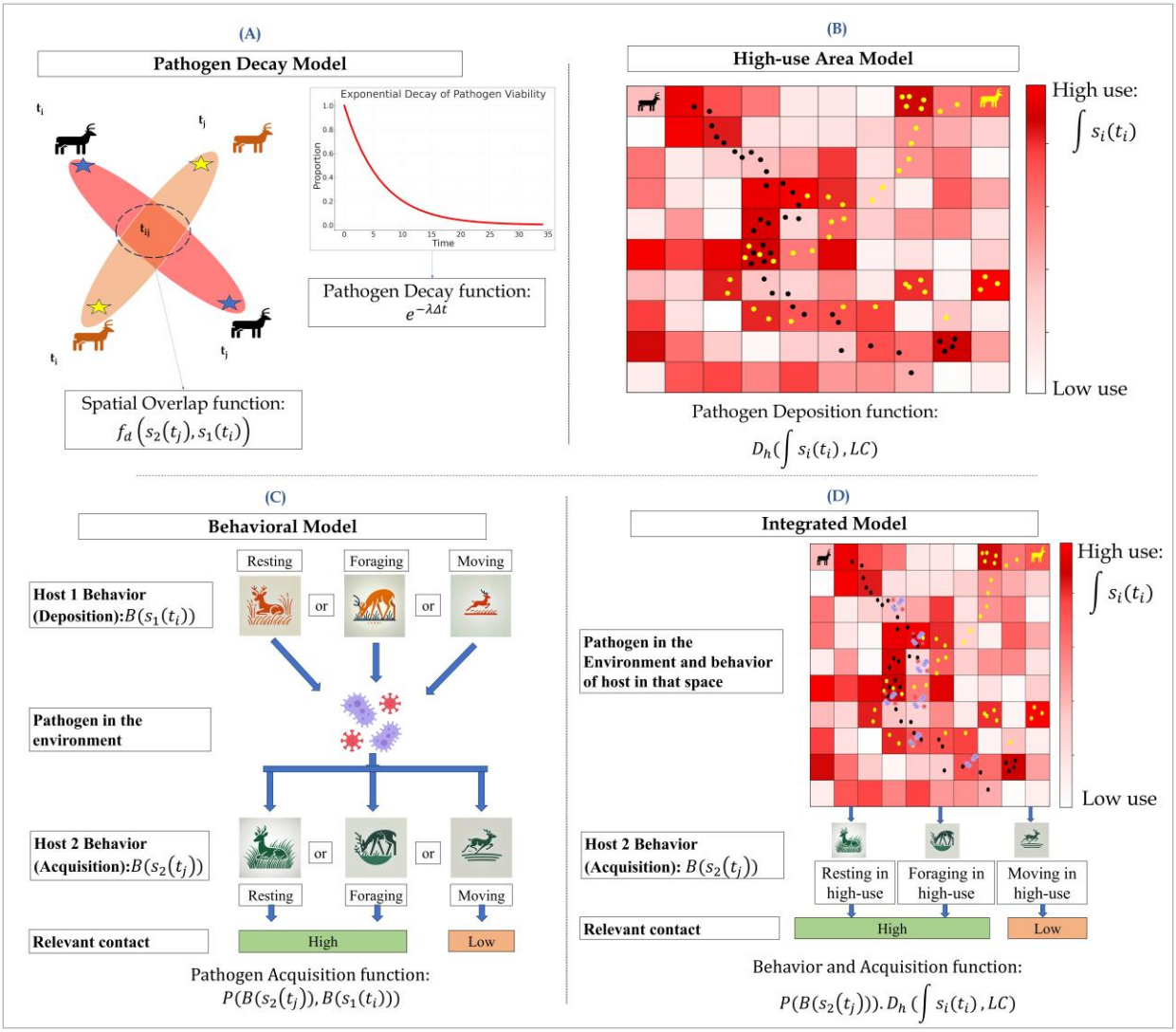


Figure 2-1: Methodological framework for capturing relevant environmentally mediated contact; (A) Quantifying potential exposure levels relating pathogen decay to the temporal lag in spatial overlap, (B) Estimating pathogen deposition levels based on time spent by infected individuals in a given location, (C) Adjusting exposure estimates by incorporating behavioral modes that influence pathogen deposition and/or acquisition and (D) integrating deposition-relevant space use and acquisition-relevant behaviors into a spatial representation of transmission risk.

2.2 Methodology

2.2.1 Modeling Framework

Building on the framework proposed by Wilber et al., (2022) and extended by Yang et al., (2023), we developed four distinct approaches to leveraging movement data to inform indirect transmission probability, i.e., pathogen decay model, high-use model, behavioral-state model, and integrated model (Figure 1).

According to Yang et al., (2023), the movement trajectory for the k^{th} host in the disease system is denoted as $s_k(t)$, where s is a two-dimensional vector of the host's spatial coordinates, and t represents time. The instantaneous weight of environmentally-mediated contact ($w_{i,j}^{1\to 2}$) between Host 1 at time t_i and Host 2 at time t_j is determined by the time difference (Δt) between t_i and t_j and the distance (d) and only j<i is considered, i.e., only past exposures are considered. This creates the constraint that indirect contact can occur only if there is a time lag between deposition and acquisition. Direct contact (i.e., j=i) is excluded by definition in this model. The equation is given by:

$$w_{i,j}^{1\to 2} = \underbrace{f_\delta\left(\Delta t = t_j - t_i\right)}_{\text{Temporal encounter}} . \underbrace{f_d\left(s_2(t_j), s_1(t_i)\right)}_{\text{Spatial encounter}}; t_i < t_j \quad \text{----------(1)}$$

Where f_δ is a temporal encounter function and $f_d(s_2(t_j), s_1(t_i))$ determines spatial distance threshold for contact, and that determines the temporal gap of the contact. The functions (f_d and f_δ) that relate time and distance between indirect contacts to risk of transmission can be defined by the user to reflect the nuances of the host-pathogen system. In Equation 1, we used binary step functions to express that there is a contact when the trajectories are within a specific time and distance, and not otherwise. In the subsequent models, this framework is modified or extended by incorporating terms for pathogen deposition, behavioral acquisition, or their contact.

A contact kernel $K^{1\to2}$ experienced by Host 2 from Host 1 consists of the instantaneous weight of contact across all timesteps:

$$\mathrm{K}^{1\to2} = \begin{bmatrix} 0 & w_{1,2}^{1\to2} & w_{1,3}^{1\to3} & \cdots & w_{1,n}^{1\to n} \\ 0 & 0 & w_{2,3}^{1\to2} & \cdots & w_{2,n}^{1\to2} \\ 0 & 0 & 0 & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & w_{n-1,n}^{1\to2} \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} \text{----------(2)}$$

2.2.1.1 Pathogen Decay Model

For highly transmissible pathogens that are transmitted via the environment (e.g., Influenza A virus;Leung, 2021), the inhalation or contact with virus-laden aerosols or respiratory droplets might trigger infection, regardless of host acquisition or deposition behaviors (Wang et al., 2021). In such scenarios, the transmission rate of the disease systems is primarily impacted by the pathogen decay rate in the environment, as pathogens may remain viable only for specific periods based on environmental conditions like temperature and humidity (Wimann et al., 2021). Several studies have suggested that the likelihood for many pathogens to remain infectious in the environment declines exponentially over time (Laggan et al., 2023), though we note exceptions to this are known (e.g., CWD). To estimate the environmentally-mediated contact weights based on the potential transmission risk influenced by the natural loss of pathogen viability, we parameterize the spatial encounter function as an exponential decay function (Figure 1):

$$w_{i,j}^{1\to2} = \underbrace{f_d\left(s_2(t_j), s_1(t_i)\right)}_{\text{Spatial encounter}} . \underbrace{e^{-\lambda \Delta t}}_{\text{Pathogen Decay}} \text{----------(3)}$$

where λ is a parameter that defines the rate at which the pathogen decays in the environment.

2.2.1.2 High-use Area Model

For some environmentally-mediated disease systems, such as those involving *Leptospira spp*, which primarily rely on exposure to contaminated environment, it is essential to identify the high-risk areas for potential contamination (Bradley et al., 2023). Additionally, in dry climates with limited water resources, animals may congregate around small water sources, such as livestock water supplies, leading to pathogen accumulation and increased disease transmission risk. Based on host movement patterns, it has been hypothesized that areas with higher visitation frequencies and longer durations of stay are more likely to be contaminated by hosts (Esposito et al., 2023). These areas increase transmission risk, as repeated contacts with the contaminated environment facilitate pathogen transfer between individuals (Plowright et al., 2017; Pepin et al., 2021). Building on the pathogen decay model, we incorporate an additional dimension by considering potential risk areas for pathogen deposition. Specifically, we define high-use areas, which are locations that animals visit frequently as potential contamination hotspots. Additionally, in some scenarios, certain point resources or specific habitats may be known where the shedding host is more likely to shed pathogens, such as the water ponds in avian influenza transmission. To better capture environmentally-mediated contact events relevant to transmission in such disease systems, we adjust the environmentally-mediated contact wights based on potential pathogen deposition. The equation for this high-use area model is defined as:

$$w_{i,j}^{1\to 2} = \underbrace{f_d\left(s_2(t_j), s_1(t_i)\right)}_{\text{Spatial encounter}} . \underbrace{e^{-\lambda \Delta t}}_{\text{Pathogen Decay}} . \underbrace{D_h \int s_i(t_i), LC)}_{\text{Pathogen Deposition}} \text{----------(4)}$$

where $D_h(\int s_i(t_i), LC)$ is a function of pathogen deposition based on the cumulative visitations of the host on the landscape ($\int s_i(t_i)$), i.e. the high-use areas. Landscape features (LC) are set up to act as a feature-presence-associated constant, enabling the introduction of

habitat-based deposition constraints where existing knowledge warrants those associations. Habitat-based variation in pathogen persistence or host behavior would add a substantial amount of uncertainty to our inferences. Here, LC functions as a spatial mask, allowing contacts to be evaluated only within a specified land-cover type.

2.2.1.3 Behavioral Model

The Behavioral Model highlights how specific behaviors impact disease transmission. Certain behaviors of Host 2 (the recipient), such as resting and foraging, can increase the chance of pathogen acquisition from the environment, making those indirect contacts relevant (Dougherty et al., 2022). In contrast, Host 2 with moving or walking behavior is considered irrelevant for transmission, as transient environmental interactions do not provide sufficient exposure. This distinction is particularly important in disease systems where transmission depends on behaviors like grazing or nesting (e.g., anthrax) (Binning et al., 2017; Yang et al., 2021). Additionally, some behaviors might contribute to more pathogen shedding in the environment than others. For example, in some fecal-oral transmission systems like chronic wasting disease, the foraging or licking behaviors by the shedders (Host 1) can cause environmental contamination (Otero et al., 2021). Host movement data can be used to probabilistically distinguish these behaviors, allowing for the better identification of relevant environmentally-mediated contact. The equation for this model is:

$$w_{i,j}^{1\to 2} = \underbrace{f_d\left(s_2(t_j), s_1(t_i)\right)}_{\text{Spatial encounter}} \cdot \underbrace{e^{-\lambda \Delta t}}_{\text{Pathogen Decay}} \cdot \underbrace{P(B(s_2(t_j)), B(s_1(t_i)))}_{\text{Pathogen Acquisition and Deposition}} \text{ ----------(5)}$$

where P($(B(s_2(t_j)), B(s_1(t_i)))$ is a probability function that defines the likelihood of transmission based on different acquisition and deposition behaviors, with a primary focus on the

behavior of Host 2 (the acquiring individual) at the time of environmental contact . $B(s_2(t_n))$ is a function that estimates the behavioral states of Host 1 and/or 2 based on its trajectory at that time. For example, P can be a piecewise function: when Host 1 and 2 are estimated as moving, the probability can be low, as this behavior is not related to disease transmission. Conversely, if Host 1 and 2 are estimated to be foraging or resting, we can assess a high probability, reflecting behaviors relevant to pathogen acquisition (Egan et al., 2021).

2.2.1.4 Integrated model

There are also some fecal-oral disease systems such as the bovine brucellosis, the major shedding route – contamination via body fluid and fetus during abortion – is hard to predict or capture by just using host movement data (Khurana et al., 2021). However, it is critical to consider pathogen deposition and acquisition to identify relevant environmentally-mediated contacts in such disease systems. Thus, we propose to integrate the high-use model and behavioral model to estimate the potential areas of high contamination and estimate the likelihood of hosts acquiring pathogens from the environment through specific behaviors, enhancing the refinement of relevant environmentally-mediated contacts. A contact is retained only if it satisfies both spatial and behavioral conditions—occurring in a high-use area and involving behavior combinations that elevate transmission risk. Compared to the earlier models reliant on single alterations to the base contact definition, the integrated model includes spatial and behavioral constraints, extending the function initially proposed in Equation 1. Our framework builds on prior work (Wilber et al., 2022; Yang et al., 2023), by combining behavior filtering and high-use environmental weighting into a modular structure, which allows these components to be toggled and compared independently to facilitate flexible, scenario-based

estimation of environmentally mediated transmission under different scenarios. The equation for this model is:

$$w_{i,j}^{1\rightarrow 2} = \underbrace{f_d\left(s_2(t_j), s_1(t_i)\right)}_{\text{Spatial encounter}} . \underbrace{e^{-\lambda \Delta t}}_{\text{Pathogen Decay}} . \underbrace{D_h \int s_i(t_i), LC)}_{\text{Pathogen Deposition}} . \underbrace{P(B(s_2(t_i)))}_{\text{Pathogen Acquisition}} \quad \text{----------(6)}$$

The equation is interpreted in a sequential process that governs environmentally mediated transmission. The term D_h indicates the deposition of pathogens by Host 1 upon shedding, and the behavioral probability function P(B) controls the chance of Host 2 acquisition upon subsequent environmental contact. Although the behavioral filter can theoretically include both hosts, our implementation targets conditioning to the behavior of the acquiring host because deposition is automatic when Host 1 visits the site, yet acquisition depends on the behavioral status of Host 2 upon contact with the contaminated environment.

2.2.2 Case Study

2.2.2.1 Data Description and study area

To demonstrate the applicability of our methodological framework, we applied it to determine transmission relevant environmentally-mediated contacts in the wild pig (*Sus scrofa*) population in Florida. Wild pigs exhibit a complex social structure, typically forming matriarchal groups consisting of adult females and their offspring, while adult males are often solitary or form small bachelor groups (Gabor et al., 1999; Kaminski et al., 2005; Clontz et al., 2023). This social behavior influences their movement patterns and contact dynamics, making them an ideal species for studying environmentally-mediated disease transmission (Yang et al., 2023; Podgorski et al., 2014). We captured 17 adult wild pigs (10 female, 7 male) from different social groups on the Archbold Biological Station - Buck Island Ranch (ABIR) in Florida, USA and

deployed GPS collars (Catlog GPS device and Lotek LMRT3 VHF Collars, Lotek ©, WA, US) to collect their movement trajectories from December 13, 2019 to September 13, 2020, with 16 days of missing information in between, thus having 259 days of data. Despite the fact that the subjects were not collared on the same day and tracking sessions differed slightly, we restricted our analysis to the overlap duration December 13, 2019 to September 13, 2020 when all 17 pigs were being tracked simultaneously so that we could have uniform temporal coverage for pairwise contact analysis. ABIR is a 42.3 km^2 commercial beef cow-calf operation managed at commercial production levels and maintains an average standing inventory of ~ 3,000 head of cattle. The details of the study site can be found at (Yang et al., 2023). The collars were programmed to record GPS fixes every 10 min with locational errors of 6–10 m on average. To test our models, we have also simulated pathogen transmission of swine influenza virus for short-lived (7 days) and brucellosis for long-lived (30 days).

2.2.2.2 Model Parameterization

For the parameterization of the pathogen decay model, we define the spatial encounter function, $f_d(s_2(t_j), s_1(t_i))$, as a piecewise function with a 10-m distance threshold (Yang et al., 2023). Specifically, we define f_d (d) = 1 if d ≤ 10 meters, and 0 otherwise; reflecting that contacts only occur within 10 meters, a threshold chosen to match the higher limit of GPS error range (6-10 meters). This accounts for location inaccuracies, ensuring identified co-locations are realistic and while accounting for locational uncertainty and reducing the likelihood of both false-positive and false-negative contact assignments. To consider the reduction in pathogen viability over time, we used two decay rates of λ_1=0.7 per day and λ_2=0.16 per day. These rates represent pathogens with different environmental persistence, allowing us to demonstrate and compare how environmentally-mediated contact kernels and distributions differ across disease

systems. The decay rates assume an exponential decline in pathogen viability, with λ_1=0.7 resulting in nearly complete decay of pathogens within seven days, while λ_2=0.16 reflecting a slower decay, leading to negligible pathogen presence after thirty days (Figure S1).

In the high-use model, we defined D_h (deposition of pathogens) as high-use areas, which were parameterized based on host visitation rates. The study area was divided into a 30×30-meter grids to match the spatial resolution of standard land cover products, such as the National Land Cover Database (NLCD) for facilitation of future fusing with environmental covariates. We used GPS positional uncertainty (~10 meters) as a guide for deriving the minimum practical grid size, but edge effects are still problematic at any resolution. However, Land cover was not explicitly incorporated in this case study to maintain focus on behavioral dynamics and due to limited habitat-specific shedding data, but following prior work (Yang et al., 2021; Wilber et al., 2022), incorporating land cover as a spatial weight in future models could enhance predictive accuracy by accounting for habitat-driven variation in host movement and environmental persistence. Land cover filtering in systems where pathogen deposition or persistence is understood to depend on particular landscape features is straightforward for the model to accommodate, but for this case study, we focused on only for clear interpretation of these effects on indirect contact. A visit is defined as an occurrence of an animal spending time in a specific location with at least one GPS point recorded, where successive visits to the same spatial unit are separated by a temporal gap of at least 10 minutes. Visitation rates were defined as the number of times hosts visited a given spatial zone (each grid cell), normalized by the total observation period. High-use areas were defined as locations within the 75th percentile of visitation rates, representing zones with elevated host activity. These areas are assumed to have higher levels of environmental contamination due to frequent deposition and acquisition events, making them

critical hotspots for pathogen transmission. Rather than weighting the deposition process towards higher or lower rates, these high-use zones act as spatial filters; a contact is only kept if it takes place inside one of these zones. We also tested the 50 and 90 percentiles to define the high-use area and how these thresholds influence the identification of relevant environmentally-mediated contacts (see supporting information, Figure S5 and Figure S6).

In the behavioral model, the behavioral states based on host movement were classified using a Hidden Markov Model (HMM), which defines stages based on step length and turning angle derived from GPS movement data (Rabiner & Juang, 1986; Awad & Khanna, 2015). We parameterized the HMM using a Gaussian distribution with three behavioral states, incorporating step length and turning angle as observational features derived from normalized spatial and temporal GPS data. Each model was trained individually for each animal, utilizing iterative optimization with up to 200 iterations to ensure convergence by checking changes in log-likelihood. Sex of the animal was included as a covariate due to differences in sex-specific behavior and movement patterns in wild pigs (Clontz et.al, 2021), enabling the model to take into consideration variations in male and female step length, turning angle, and state transition probabilities. The Viterbi algorithm was employed to decode the most probable state sequence for each trajectory and then the states were assigned to each of the GPS points, facilitating biologically meaningful classification of movement behaviors (Lou, 1995; Clontz et.al, 2021). A schematic of the Hidden Markov Model structure, including transition probabilities, step length & turning angle distributions, is provided in Supporting Information Figure S7. We have used *hmmlearn* package in Python library for these analyses. For the probability function $P(B(s_2(t_j)))$, which defines the likelihood of transmission based on different acquisition and deposition behaviors, we used a simplified parameterization where the probability of

transmission is set to 1 for foraging and resting states and 0 for the moving state. We kept foraging and resting states as separate classes to demonstrate the flexibility of our approach for future applications using higher-resolution behavioral data, even though they were given equal transmission probabilities in this implementation. Finally, the integrated model combined the parameterizations of high-use and behavioral models. While continuous filters can be included in the framework, in this case study, the filters are binary and serve as filters to find pertinent contacts. This simplification makes it possible to interpret the relative influence on contact more clearly.

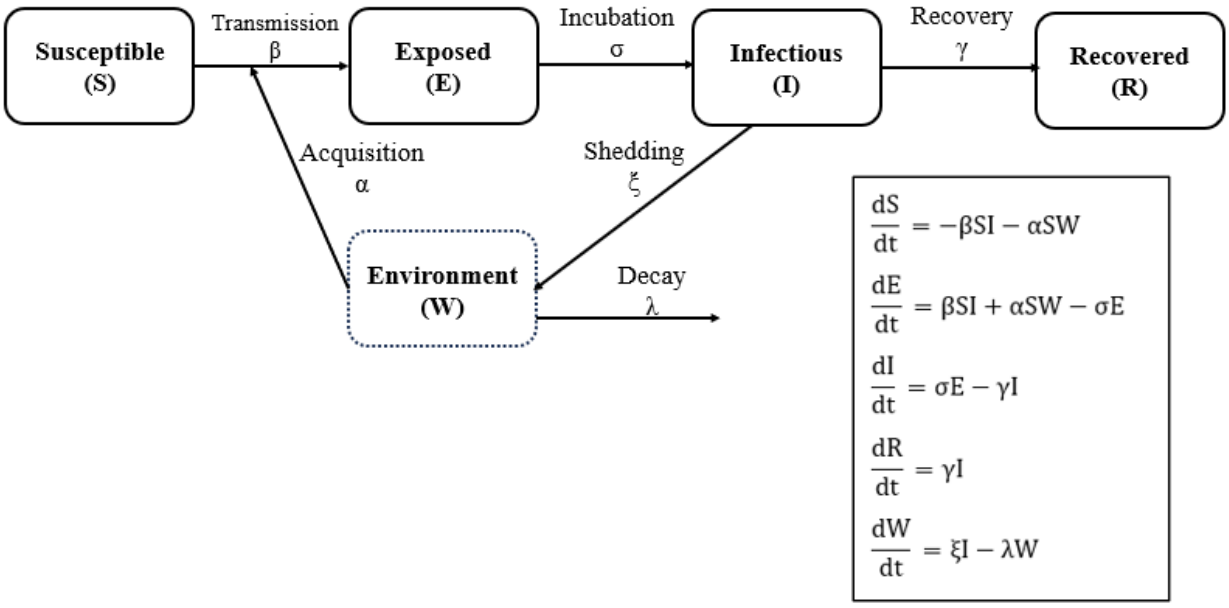


Figure 2-2: Schematic Diagram of the environmentally mediated SEIR model with their respective PDEs.

2.2.2.3 Disease Transmission Model

To illustrate the implications of our contact models for disease transmission, we implemented a stochastic SEIR (Susceptible–Exposed–Infectious–Recovered) (Figure S2) model augmented with an explicit environmental compartment to simulate outbreak dynamics for two disease systems with environmental transmission and contrasting life histories: Influenza A virus

and swine brucellosis (Brouwer et al., 2017; Ghosh et al., 2023). This enhanced model captures indirect transmission via environmental reservoirs, thereby allowing us to explore how differences in transmission mechanisms and network structures influence key epidemiological outcomes such as disease persistence, epidemic peaks, and the basic reproduction number (R_0) (Ogbunugafor et al., 2020). We parameterize the SEIR epidemic model to properly assess disease-specific transmission dynamics and provide realistic simulation of outbreak progression. Adequate parameterization enables the model to mimic biological processes including incubation periods, infectious periods, and contact rates. The environmental compartment explicitly models pathogen accumulation through shedding from infected individuals, its subsequent decay in the external environment, and eventual acquisition by susceptible hosts (Brouwer et al., 2017). Parameterization reflects the distinct biological properties of the pathogens: Influenza A virus exhibits rapid decay in external environments while *Brucella suis* shows prolonged environmental stability in contaminated water and soil (Buch et al., 2023). Stochastic transmission (β), incubation (σ), recovery (γ), shedding (ξ), and environmental acquisition (α) rates capture natural infection heterogeneity at host and environmental interfaces (Table S2), according to recent empirical research (Antolin, 2008; Buch et al., 2023). Deterministic pathogen decay rates (λ) capture differential rapid versus long-term environmental survival, as a function of exposure duration, according to pathogen longevity traits. Initial infection proportion (I_0) was initialized to simulate initial outbreak seeding in controlled environments. In our simulations, we conducted 1,000 stochastic iterations over a 365-day period per run, thereby capturing natural variability and enabling a robust assessment of how environmental persistence interacts with contact network characteristics to drive transmission dynamics (Dureau et al., 2013).

Key epidemiological metrics, including R_0 and time to peak incidence, were estimated from the simulation outcomes. By integrating these metrics, along with extinction probabilities where applicable, our model provides a comprehensive framework to evaluate how variations in contact structures and environmental transmission pathways influence the spread and persistence of disease within a population.

Table 2-1: Parameters used in the SEIR disease transmission model

Parameter	**Short-lived (7 days)**	**Long-lived (30 days)**	**Type**	**Reference**
Transmission Rate (β)	0.1 - 0.3	0.05 - 0.2	Stochastic	Antolin, 2008; Buch et al., 2023
Incubation Rate (σ)	0.13 - 0.30	0.05 - 0.1	Stochastic	Virlogeux et al., 2015; Elmonir et al., 2022
Recovery Rate (γ)	0.1 - 0.2	0.07-0.14	Stochastic	Mu et al., 2018; Elmonir et al., 2022
Pathogen Decay Rate (λ)	0.7	0.16	Deterministic	User defined
Environmental Shedding Rate (ξ)	0.3 - 0.6	0.2 - 0.4	Stochastic	Bhattarai et al., 2011; Rebollada-Merino et al., 2022
Environmental Acquisition Rate(α)	0.2 - 0.5	0.2 - 0.4	Stochastic	Brouwer et al., 2022; Rebollada-Merino et al., 2022
Initial Proportion Infected (I_o)	0.01 - 0.03	0.01 - 0.03	Deterministic	User defined

2.2.3 Model Evaluation and Analysis

To contrast the effect of different assumptions about environmentally mediated contact on transmission dynamics, we used a structured assessment framework to approximate the total dyadic contacts and count of unique host pairs identified by each contact model for each pathogen persistence scenarios (7-day and 30-day windows). We also monitored the maximum

frequency of contact between each pair of individuals throughout the whole 259-day observation period to enable heterogeneity in exposure intensity.

Second, we constructed weighted, directed contact networks from the estimated contact matrices of each model. Each pair is unidirectional and unique, i.e. A→B is not equal to B→A. To summarize the structural characteristics of such networks, we calculated four widely used measures with epidemiological relevance like edge density, which is the fraction of observed contacts relative to all possible dyads; transitivity or global clustering coefficient, which explains the tendency of nodes to get connected with densely connected triads that may sustain local outbreaks; modularity, which measures the degree to which networks are divided into distinct communities that may act as transmission bottlenecks; and degree assortativity, which measures the correlation in node connectivity and may influence epidemic spread velocity and amplitude. We also generated spatial risk surfaces for each of these models by rastering contact intensity onto the 30×30 m grid cells used by the behavioral and high-use models. This allowed us to see and compare the spatial pattern of transmission-relevant contact hotspots under each of the definitions of contacts.

2.3 Results

2.3.1 Contact Kernels

The pathogen decay model recorded the highest number of unique pairs with 162 for a 7-day pathogen persistence scenario and 176 for a 30-day scenario out of a possible 256 unique pairs, alongside maximum contact counts of 882 and 2,075 across 259 days, respectively (Figure 2). In the high-use area model, while the maximum contact counts remained similar (882 for 7-day scenario and 2,072 for 30-day scenario), the unique pairs slightly decreased to 159 for 7-day

scenario and 172 for 30-day scenario. As expected, the behavioral model showed a marked reduction in both contact counts and unique pairs (because some of the movement behaviors had 0 probability of transmission as defined through the step function), with maximum counts

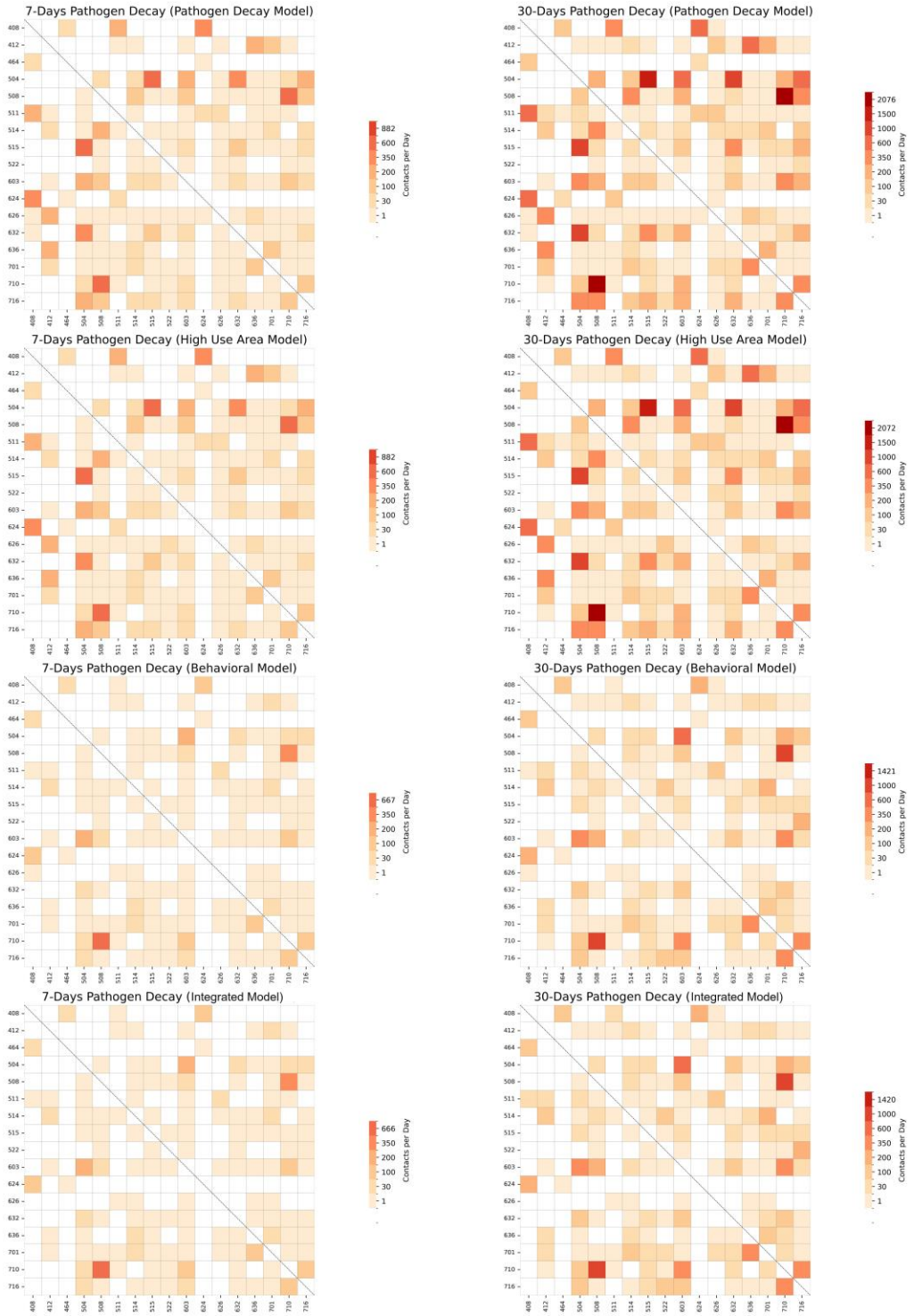


Figure 2-3: Contact Kernel plots showing the total number of indirect contacts between each pair of wild pigs under four modeling for both 7-day and 30-day pathogen decay. The x- and y-axes indicate the unique animal IDs, with each cell representing the cumulative contact count between a pair. Darker colors correspond to higher number of contacts

decreasing to 667 for 7 days and 1,421 for 30 days, and unique pairs reducing to 136 for 7 days and 149 for 30 days. The integrated model had the lowest unique pairs, with 135 for 7 days and 148 for 30 days, and maximum contact counts of 666 for 7 days and 1,420 for 30 days. These results illustrate a consistent reduction in unique pairs and contact counts across models as stricter criteria were applied. For the 7-day decay model, a shift from the pathogen decay model to the high-use model led to a moderate 0.30% decline in total contacts. Conversely, there were marked declines against the behavioral (70.34%) and the integrated (70.43%) models. As indicated, such tendencies were replicated under the 30-day decay model, showing consistent model-driven variation in the reduction of contact (Figure S3).

2.3.2 Network Structure and Matrices

The network structure and corresponding matrices (Figure 3 and Table S1) reveal differences in clustering and connectivity across models. The pathogen decay model shows the highest edge density (0.60 for 7 days, 0.65 for 30 days) and strong local clustering with transitivity values of 0.81 and 0.84. Modularity is moderately high (0.50 for 7 days, 0.44 for 30 days), while assortativity is low (0.29 and 0.17), indicating a network with widespread connectivity and weak degree assortative patterns. The high-use area model increases modularity slightly (0.52 for 7 days, 0.46 for 30 days) and achieves the highest assortativity (0.43 for 7 days, 0.30 for 30 days), highlighting more structured clustering. The behavioral and integrated models both show the same edge density (0.50) for 7 days and slightly different for 30 days, with same modularity 0.42 for 7 days and 0.37 for 30 days in both models. Transitivity stays at 0.77–0.79,

and assortativity declines to its lowest values (0.26 for 7 days, 0.30 for 30 days), reflecting sparse, selectively filtered networks.

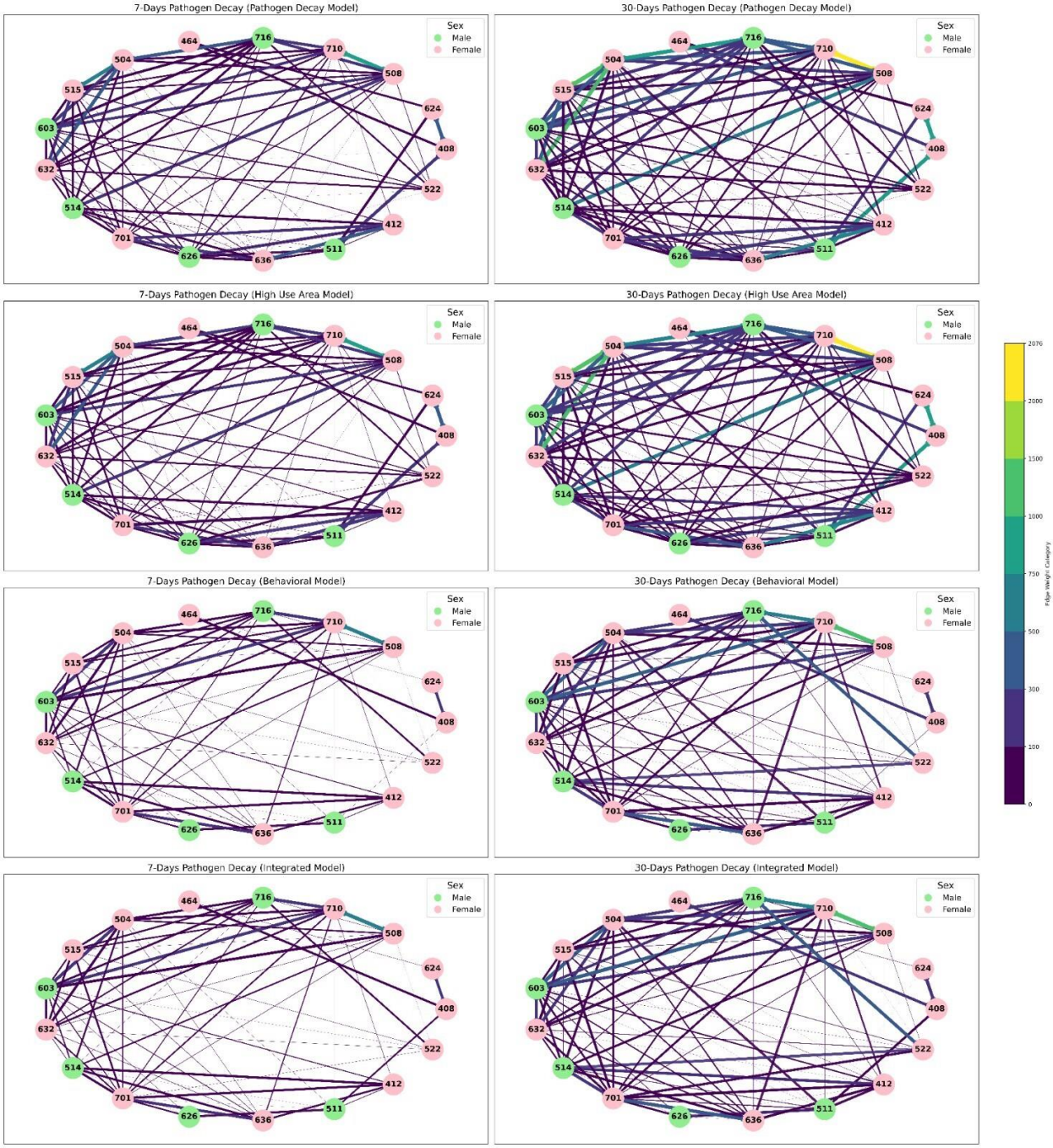


Figure 2-4: Contact network structure under different contact models for 7-day and 30-day pathogen decay scenarios. Each node represents an individual wild pig, and each edge denotes amount of contacts. Edge thickness and color intensity correspond to the number of indirect contacts.

2.3.3 Spatial Distribution

We define 'contact zones' as areas where contact events occur, and 'high-intensity areas' as contact zones with the highest contact frequencies. The pathogen decay model displays

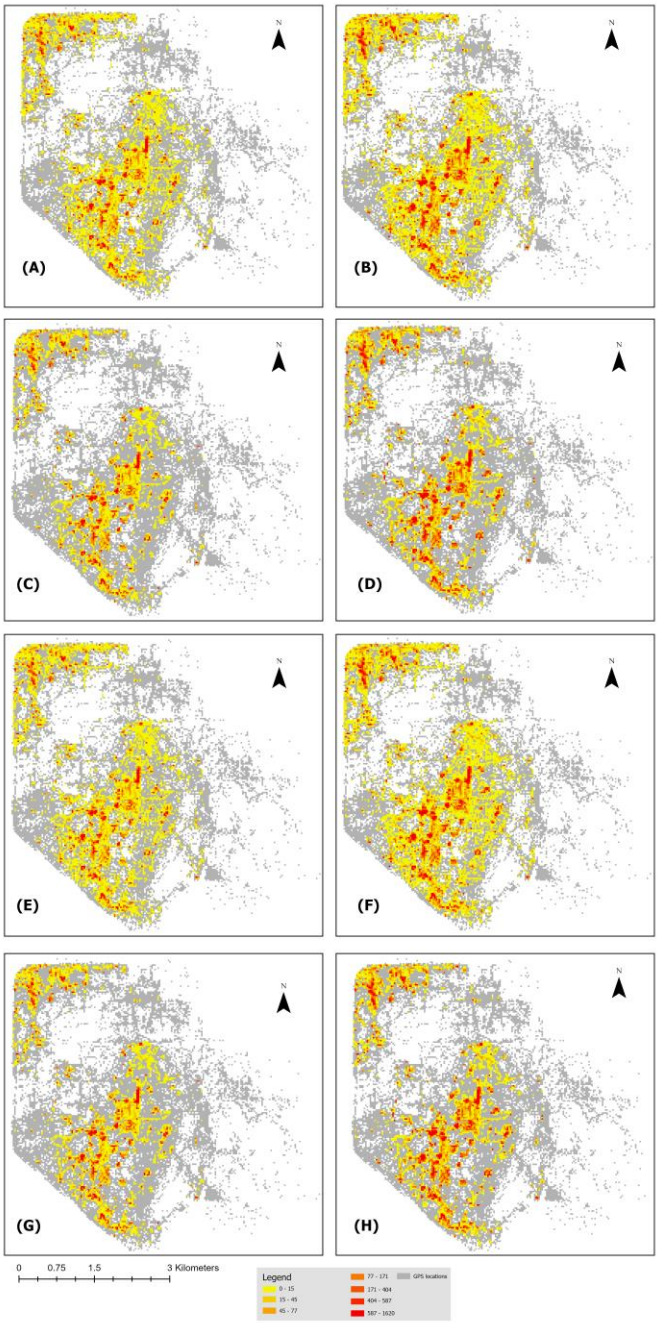


Figure 2-5: Spatial distribution of contact in landscape under different models; (A) pathogen decay only (7 days), (B) pathogen decay only (30days), (C) high use area (7 days), (D) high use area (7 days), (E) behavioral model (7 days), (F) behavioral model (7 days), (G) Integrated model (7 days), (H) Integrated model (30 days). Grid

cells represent 30×30 m spatial units. Grey areas indicate zones with no contact, yellow indicates locations where contacts occurred, and red denotes high-intensity contact zones.

widespread contact zones (yellow colored zone in Figure-4), with high-intensity areas (red colored zone in Figure-4) concentrated in central and peripheral regions for both 7- and 30-day durations (Figure 4). The high-use area model reduces the extent of these zones, concentrating high-intensity areas in revisited central patches, while peripheral regions become less prominent. In contrast, the behavioral model reveals a pattern similar to the pathogen decay model, with widespread but slightly more localized contact zones, as transient activities are filtered. The integrated model produces the most selective pattern, confining high-intensity areas to a few small, central patches where spatial and behavioral criteria overlap. The distribution narrows progressively in the high-use and integrated models, while pathogen decay and behavioral models retain broader spatial patterns. Pairwise difference and spatial distribution of contacts among different model has been shown in supporting information (Figure S8 and Figure S9)

2.3.4 Disease Transmission

The SEIR model outputs illustrate how extending pathogen persistence from 7 days to 30 days fundamentally alters the epidemic curve and the probability of transmission (Figure S4) in each environmentally-mediated contact models. For the 7-day scenarios, the Pathogen Decay model carries the highest R_0 of 2.42±0.47 (mean ± standard deviation) and peaks in incidence at 95.06±7.80 days, while the High-use, Behavioral, and Integrated models have successively lower R_0 values (2.06±0.39, 1.56±0.42, and 1.35±0.50, respectively) with peak incidence at 82.21±9.01-74.73±6.67 days (Figure 5). Under 30-day persistence, R_0 rises steeply in the Pathogen Decay model to 6.82±2.09 with delay in peak incidence to 120.50±2.88 days, whereas the other models are of the same trend with R_0 5.12±1.57 (Behavioral) to 4.62±1.60 (Integrated) with time to peak incidence decreasing to 89.87±3.37 days.

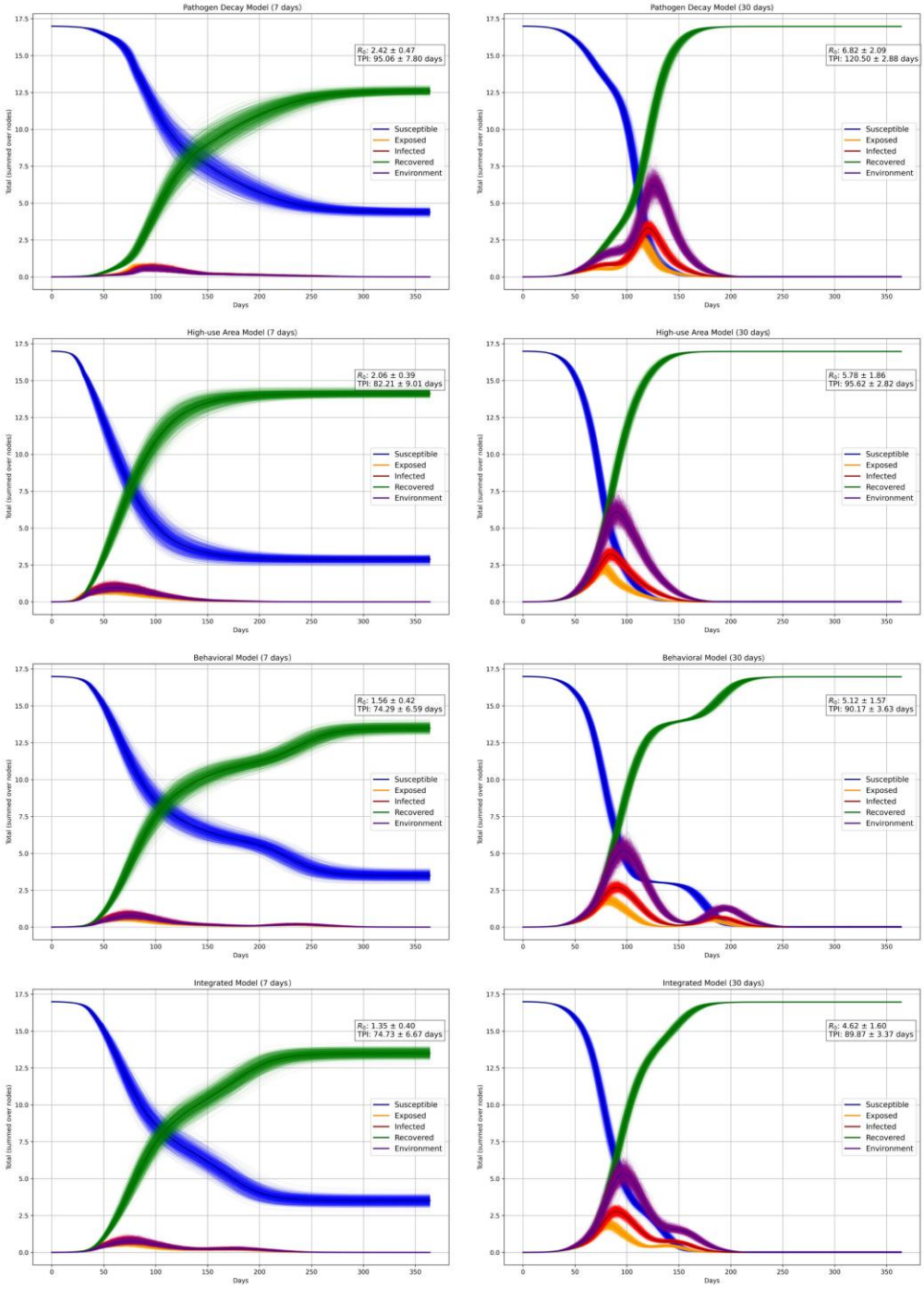


Figure 2-6: SEIR epidemic simulations for 7-day and 30-day pathogen decay durations under different contact modeling scenarios. For a specific contact model, each panel shows the mean trajectories (of 1000 runs) of the epidemiological compartments (Susceptible, Exposed, Infected, Recovered, and Environmental Load) over a period of 365 days.

2.4 Discussion

Environmentally-mediated contacts play a critical role in disease transmission dynamics for environmentally resilient pathogens such as *Bacillus anthracis, brucella spp.*, and influenza A viruses by enabling pathogens to persist in the environment and facilitating indirect transmission even when hosts are present simultaneously. Previously, researchers often defined those contacts based on colocation of hosts within a segregated timeframe, which can overestimate the environmentally-mediated contacts, as not all instances are epidemiologically relevant. Building on previous framework, this paper offers a refined framework for quantitatively defining relevant indirect contacts by considering pathogen decay, host movement, and behavior. We applied this framework to wild pigs in Florida to demonstrate how pathogen persistence, high-use areas, and host behavioral states shape the environmentally-mediated contact networks. We found that while prolonged pathogen survival enhances connectivity, the contamination levels and host behaviors restrict transmission risk in specific ways. Our model framework captures environmentally-mediated contacts via environment at different levels of epidemiological relevance, providing a nuanced understanding of disease transmission.

2.4.1 Tradeoffs of Different Model Structures

Our framework presents four alternative structures for characterizing environmentally mediated contacts, each of which embodies different assumptions about host behavior, space use, and pathogen viability. By articulating these models and comparing them, we illustrate how transmission estimates and the contact networks that support them, are sensitive to the definition of epidemiologically relevant exposure. This comparative framework is one of the strengths of our approach: it allows us to identify which assumptions have the greatest effect on transmission risk and which are scenario-consistent.

Although pathogen decay model is straightforward to apply, it is often oversimplified in many real-world disease systems, since pathogen decay is not the only factor determining transmission. Pathogen deposition and acquisition both influence transmission, and ideally, we would consider behaviors related to these processes in the identification of relevant environmentally-mediated contacts. Some behavioral states, such as resting and foraging states, which have long duration and high visitation, respectively, are often considered as high potential of pathogen deposition and acquisition. However, these states do not necessarily represent specific behaviors or actions (Pohle et al., 2017). The detection of certain deposition and acquisition behaviors, such as urination and defecation, cannot rely on host movement data only and often requires additional monitoring methods, such as camera traps (Egan et al., 2023). Camera traps and video footage from individuals wearing video devices can supply useful validation information for behavioral state assignments in movement models, although they only record a very limited temporal window of animal activity (Lavelle et al., 2012; Lavelle et al., 2014). Camera trap detections that overlap with GPS-collared animals could be used to estimate within-state rates of behavioral activity and test state-specific activity (e.g., resting or foraging), which can improve interpretation and parameterization of state-dependent contact models. The high-use area model in our framework, which considers the pathogen deposition risk as high visitation areas, can be used to capture the proxy of deposition/acquisition, as it identifies where repeated host revisitation increases exposure risk. However, this approach may not capture the full extent of transmission that is occurring outside of the pre-defined highly concentrated areas. Some new technologies combining GPS collars with camera are developing and can be used to better understand movement and behavior in the context of these models (Lavelle et al., 2012; Lavelle et la., 2014).

Although the behavioral model excludes contact during movement states to avoid overestimation of exposure risk, animals may still deposit pathogens during movement. Excluding this behavior can lead to underestimation of contamination risk for frequently traveled movement corridors. Many species repeatedly use the same corridors to transit between habitats, increasing the risk of contamination and pathogen exposure along those routes. The presented integrated model, which both accounts for acquisition behaviors and incorporates the high-use areas as deposition sites, can retain such corridor-based contacts. This is particularly important for species with high site fidelity, where repeated movement patterns maximize the likelihood of exposure. Our results indicate that behaviorally bounded contact networks can generate slower or delayed epidemic peaks, capturing dynamics that previous methods were unable to resolve.

Our models can also be applied to forecast hotspots of transmission through the integration of predictive movement modeling and environmental data. Future trajectories of individual animals can be predicted using predictive movement models, such as step selection functions (Thurfjell et al., 2014, Potts et al. 2022, Potts et al. 2023) or continuous-time movement models (CTMM; Codling et al., 2008), parameterized on the basis of historical movement behavior and environmental covariates. The predicted trajectories can be used to project the spatiotemporal overlap of contaminated patches and host presence when combined with pathogen environmental persistence models. In those systems where upcoming host-pathogen overlap would tend to be foreseeable, i.e., in those with migratory behavior, site fidelity, memory or seasonal movement patterns, this approach is particularly well-suited. The applicability of the everyday movement model and presence of representative training data, however, determine how reliable such projections are.

2.4.2 Insights from the Case Study

Wild pigs are a group-living species with a complex social structure, typically forming matriarchal groups consisting of females and their offspring. This social organization results in high contact rates within social groups, while contact rates between groups are relatively low. However, between-group contacts are often the primary drivers of disease spread. In our case study, which involved collaring wild pigs from different groups, our models indicate that even between-group contacts play a disproportionately large role in transmission, particularly those contacts filtered by space and behavior. While less common, intergroup contacts are important bridging ties that link highly connected within-group clusters, improving overall network connectivity and allowing for wider transmission throughout the population. In particular, high weighted indirect contact pairs seem to often indicate transient group merging or range overlap, and fission–fusion processes therefore influence not only contact structure but also potential outbreak pathways (Yang et al., 2021).

When implemented to contrasting pathogens (IAV and Brucella suis), our method reveals that assumptions about contact relevance make significant impacts on epidemic projections. For instance, including all spatial intersections without behavioral or temporal filtering leads to overestimation of R_0, especially for very long-lived pathogens like Brucella suis (Pinn-Woodcock et al., 2023; Pepin et al., 2023). Conversely, behavioral filter use leads to lower R_0 values and epidemics with later peaks that better reflect empirical ecological data (Hoyt et al., 2023; Pepin et al., 2021). Because acquisition of pathogens can only occur during particular host activities (i.e., foraging and resting in our application), the behavioral model delayed and flattened SEIR curves through reducing opportunity for successful transmission after deposition had taken place. Our comparative findings reflect how host behavior, visitation patterns, and

pathogen breakdown influence each other to produce indirect contact risk, which is extremely context-sensitive (Dougherty et al., 2022; Rees et al., 2021). Traditional models frequently obscure this variability, highlighting the value of integrative, system-specific frameworks for comprehending the transmission of diseases mediated by the environment. Our method has roots in earlier modeling of indirect contact (i.e., Yang et al., 2023; Wilber et al., 2022; Dougherty et al., 2018), which customarily assigns equal epidemiological relevancy to all the indirect overlaps. In contrast, we illustrate how behavior state and visitation intensity alter the expected risk landscape in a way consistent with the experienced pathogen biology in systems such as brucellosis or IAV.

2.4.3 Extensions, Future Directions, and Limitations

2.4.3.1 Extensions and future directions of the models

To illustrate the implementation of our modeling framework for environmentally-mediated contact, we employed a single species system as a case study. Similar concepts have been shown in multi-host systems (Herraiz et al., 2024b), but our approach is novel in its modular design, explicitly separating and contrasting mechanistic processes like environmental degradation, state-of-behavior filtering, and high-usage area identification to estimate how each affects indirect contact networks and epidemic outcomes. Our models can be extended to include more sophisticated ecological settings, such as environmentally persistent or vector-borne multi-host pathogens. Multi-host disease systems involve species with varying movement patterns, habitat use, and susceptibility, all of which influence indirect transmission risk via environment (Webster et al., 2017; Lambin et al., 2010). Unlike earlier models that treat all indirect contact as epidemiologically equivalent, our model allows researchers to specify species- or behavior-specific causes of environmental contamination and risk (Murray et al., 2016; Titcomb et al.,

2021). For example, in the brucellosis system with elk–cattle interface, our framework can help identify the core shedders and the acquirers based on their behaviors or space-use patterns as well as their transmission hotspots (Berry et al., 2018). Expanding the framework to a multi-species context would also involve integrating species-specific movement and behavioral data as well as pathogen compatibility across host species. Further studies have shown that species interactions, for example through competitive exclusion or habitat partitioning may modify environmentally-mediated contact structures that affect disease transmission dynamics (Baishya et al., 2021). Incorporating these variations would enhance the predictive power of environmentally-mediated contact models and provide more accurate estimates of interspecies transmission risk.

Additionally, although our framework is designed explicitly for quantifying environmentally-mediated contact, it also has the potential to be extended to a vector-borne system, which is considered challenging from an epidemiological perspective because vectors act as both pathogen reservoirs and transmission agents. Vector-borne diseases are distinct from environmentally-mediated transmission because pathogen decay occurs throughout both environmental transmission and the vector itself (Khong et al., 2023) and most vectors can move (thus the point source changes location). For example, in tick-borne disease systems, a tick can pick up a pathogen from one host and then remain infected for an extended period before transmitting the pathogen to another host during its subsequent feeding (Boulanger et al., 2019). However, not all tick-host interactions are followed by transmission since vector competence alongside pathogen replication within the vector and host immune responses determine the probability of transmission success (Rocha et al., 2022). The integration of vector distribution, their biting rate, and habitat use into the framework will provide better estimates of when and

where transmission events are most likely to occur. Furthermore, environmental factors including temperature and humidity which influence both tick activity and pathogen viability need to be included to enhance transmission risk predictions (Rocha et al., 2022). In this context, vector-related processes can be integrated through the host-contact and environmental persistence modules, which allow for the explicit definition of parameters like habitat preference, biting rate, and vector density. Furthermore, to better capture the dynamics of vector-mediated transmission, environmental modifiers (such as temperature and humidity) that affect vector activity and pathogen viability can be included in the decay and exposure functions. This enables evaluation of where and when spillovers are most likely, particularly under climate-sensitive circumstances that change vector distribution and activity windows.

The inclusion of temporally and spatially variable environmental factors that are known to affect pathogen persistence, such as temperature, moisture, pH, soil type, and solar radiation, is an important extension of this framework. Even though our current implementation assumes simple exponential decay, persistence is frequently nonlinear under various environmental conditions. For instance, depending on soil chemistry and organic matter content, the prions of chronic wasting disease can remain infectious for months or years (Otero et al., 2021). Future models would be more representative of real-world heterogeneity if covariates like decay modifiers were included (Wißmann et al., 2021; Cohen et al., 2020). This would also enhance forecasts for pathogens like influenza, whose stability varies seasonally with climate. Building on earlier frameworks (Wilber et al., 2022), this refinement would enable mechanistic testing of how environmental variability alters indirect transmission risk across host–pathogen systems.

2.4.3.2 Limitations of the Case Study

There are also some limitations within the case study. First, the model presumes equal host susceptibility, but immune response variability or prior exposure likely affects acquisition risk (Siva-Jothy & Vale, 2021). Future versions should consider immunodynamics and heterogeneity to improve transmission estimates. Second, accurately parameterizing shedding and acquisition within environmental compartments remain challenging due to limited empirical data on infectious doses and host behaviors (Pepin et al., 2013; Brouwer et al., 2017). While the behavioral module herein uses a simple binary filter, this modular framework supports the addition of probabilistic weighting of network edges, accelerometer-informed classification of contacts, and habitat-based refinements to more fully represent behavioral effects on transmission (Egan et al., 2023; Triguero-Ocaña et al., 2021). Finally, while we aim for high-use areas as a deposition surrogate, landscape heterogeneity (e.g., vegetation cover, water, resource distribution) likely influences both host movement and pathogen survival. The incorporation of such environmental factors in future models will enhance the predictive mapping of spatiotemporal risk (Bradley & Lockaby, 2023; Esposito et al., 2023).

2.5 Conclusions

This study offers a flexible framework for modeling environmentally mediated transmission in a variety of ecological contexts by incorporating behavioral state, intensity of space use, and environmental viability within a shared framework. We illustrated the applicability of the model through case studies utilizing GPS tracking data from feral pigs. Our case study results found that contact network topology for indirect contact and the predicted pathogen spread pattern are sensitive to assumptions about host behavior, spatial overlap, and pathogen persistence in the environment. Inflated R_0 and distorted epidemic dynamics can lead

to models that disregard behavioral context or overextend the duration of exposure relevance. These results highlight the need to create epidemiologically relevant definitions of indirect contact that are consistent with both host ecology and pathogen biology. By operationalizing deposition, persistence, and acquisition processes explicitly, this framework offers a clear tool for making mechanistic hypotheses about indirect transmission. Furthermore, the modular design facilitates future extensions to multi-host and vector-borne systems, incorporation of environmental covariates such as land cover or climatic variables, and refinement using higher-resolution behavioral data (e.g., accelerometry). Consequently, this framework not only improves the interpretability of system-specific dynamics but also helps address the larger objective of building generalizable, comparative insights into the role of environmental processes in transmission across host–pathogen systems.

References

Antolin, M. F. (2008). Unpacking β: within-host dynamics and the evolutionary ecology of pathogen transmission. *Annual Review of Ecology, Evolution, and Systematics*, 39(1), 415-437. https://doi.org/10.1146/annurev.ecolsys.37.091305.110119

Arthur, R. F., Gurley, E. S., Salje, H., Bloomfield, L. S., & Jones, J. H. (2017). Contact structure, mobility, environmental impact and behaviour: the importance of social forces to infectious disease dynamics and disease ecology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1719), 20160454. https://doi.org/10.1098/rstb.2016.0454

Awad, M., Khanna, R., Awad, M., & Khanna, R. (2015). Hidden Markov model. In *Efficient Learning Machines* (pp. 81–104). https://doi.org/10.1007/978-1-4302-5990-9_5

Baishya, J., Bisht, K., Rimbey, J. N., Yihunie, K. D., Islam, S., Al Mahmud, H., Waller, J. E., & Wakeman, C. A. (2021). The Impact of Intraspecies and Interspecies Bacterial Interactions on Disease Outcome. *Pathogens*, *10*(2), 96. https://doi.org/10.3390/pathogens10020096

Barrios-Garcia, M. N., & Ballari, S. A. (2012). Impact of wild boar (*Sus scrofa*) in its introduced and native range: a review. *Biological Invasions*, 14, 2283-2300.https://doi.org/10.1007/s10530-012-0229-6

Berry, E. D., & Wells, J. E. (2018). Reducing foodborne pathogen persistence and transmission in animal production environments: challenges and opportunities. *Preharvest Food Safety*, 177-203. https://doi.org/10.1128/microbiolspec.PFS-0006-2014

Bhattarai, A., Villanueva, J., Palekar, R. S., Fagan, R., Sessions, W., Winter, J., & Pennsylvania Working Group. (2011). Viral shedding duration of pandemic influenza A H1N1 virus during an elementary school outbreak—Pennsylvania, May–June 2009. *Clinical Infectious Diseases*, 52(suppl_1), S102–S108. https://doi.org/10.1093/cid/ciq026

Binning, S. A., Shaw, A. K., & Roche, D. G. (2017). Parasites and host performance: incorporating infection into our understanding of animal movement. *Integrative and Comparative Biology*, 57(2), 267–280. https://doi.org/10.1093/icb/icx024

Bradley, E. A., & Lockaby, G. (2023). Leptospirosis and the environment: A review and future directions. *Pathogens*, 12(9), 1167. https://doi.org/10.3390/pathogens12091167

Brouwer, A. F., Eisenberg, M. C., Shulman, L. M., Famulare, M., Koopman, J. S., Kroiss, S. J., ... & Eisenberg, J. N. (2022). The role of time-varying viral shedding in modelling environmental surveillance for public health: revisiting the 2013 poliovirus outbreak in Israel. *Journal of the Royal Society Interface*, 19(190), 20220006. https://doi.org/10.1098/rsif.2022.0006

Brouwer, A. F., Weir, M. H., Eisenberg, M. C., Meza, R., & Eisenberg, J. N. (2017). Dose-response relationships for environmentally mediated infectious disease transmission

models. *PLOS Computational Biology*, 13(4), e1005481. https://doi.org/10.1371/journal.pcbi.1005481

Buch, D. A., Johndrow, J. E., & Dunson, D. B. (2023). Explaining transmission rate variations and forecasting epidemic spread in multiple regions with a semiparametric mixed effects SIR model. *Biometrics*, 79(4), 2987-2997. https://doi.org/10.1111/biom.13901

Clontz, L. M., Pepin, K. M., VerCauteren, K. C., & Beasley, J. C. (2021). Behavioral state resource selection in invasive wild pigs in the Southeastern United States. *Scientific Reports*, 11(1), 6924. https://doi.org/10.1038/s41598-021-86363-3

Clontz, L. M., Yang, A., Chinn, S. M., Pepin, K. M., VerCauteren, K. C., Wittemyer, G., ... & Beasley, J. C. (2023). Role of social structure in establishment of an invasive large mammal after translocation. *Pest Management Science*, 79(10), 3819–3829. https://doi.org/10.1002/ps.7567

Codling, E.A., Plank, M.J. & Benhamou, S. (2008) Random walk models in biology. *Journal of the Royal Society Interface*, **5**, 813–834.

Cohen, J. M., Sauer, E. L., Santiago, O., Spencer, S., & Rohr, J. R. (2020). Divergent impacts of warming weather on wildlife disease risk across climates. *Science*, 370(6519), eabb1702. https://doi.org/10.1126/science.abb1702

Collier, M., Albery, G. F., McDonald, G. C., & Bansal, S. (2022). Pathogen transmission modes determine contact network structure, altering other pathogen characteristics. *Proceedings of the Royal Society B*, 289(1989), 20221389. https://doi.org/10.1098/rspb.2022.1389

Craft, M. E. (2015). Infectious disease transmission and contact networks in wildlife and livestock. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1669), 20140107. https://doi.org/10.1098/rstb.2014.0107

Daversa, D. R., Fenton, A., Dell, A. I., Garner, T. W. J., & Manica, A. (2017). Infections on the move: how transient phases of host movement influence disease spread. *Proceedings of the Royal Society B: Biological Sciences*, 284(1869), 20171807. https://doi.org/10.1098/rspb.2017.1807

Dougherty, E. R., Seidel, D. P., Blackburn, J. K., Turner, W. C., & Getz, W. M. (2022). A framework for integrating inferred movement behavior into disease risk models. *Movement Ecology*, 10(1). https://doi.org/10.1186/s40462-022-00331-8

Dougherty, E. R., Seidel, D. P., Carlson, C. J., Spiegel, O., & Getz, W. M. (2018). Going through the motions: incorporating movement analyses into disease research. *Ecology Letters*, 21(4), 588–604. https://doi.org/10.1111/ele.12917

Dureau, J., Kalogeropoulos, K., & Baguelin, M. (2013). Capturing the time-varying drivers of an epidemic using stochastic dynamical systems. *Biostatistics*, 14(3), 541–555. https://doi.org/10.1093/biostatistics/kxs052

Egan, M. E., Pepin, K. M., Fischer, J. W., Hygnstrom, S. E., VerCauteren, K. C., & Bastille-Rousseau, G. (2023). Social network analysis of white-tailed deer scraping behavior: Implications for disease transmission. *Ecosphere*, 14(2), e4434. https://doi.org/10.1002/ecs2.4434

Elmonir, W., Abdel-Hamid, N. H., Hamdy, M. E., Beleta, E. I., El-Diasty, M., Melzer, F., Wareth, G., & Neubauer, H. (2022). Isolation and molecular confirmation of *Brucella suis* biovar 2 from slaughtered pigs: an unanticipated biovar from domestic pigs in Egypt. *BMC Veterinary Research*, 18(1), 224. https://doi.org/10.1186/s12917-022-03332-2

Esposito, M. M., Turku, S., Lehrfield, L., & Shoman, A. (2023). The Impact of Human Activities on Zoonotic Infection Transmissions. *Animals*, 13(10), 1646. https://doi.org/10.3390/ani13101646

Fofana, A. M., & Hurford, A. (2017). Mechanistic movement models to understand epidemic spread. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1719), 20160086. https://doi.org/10.1098/rstb.2016.0086

Gabor, T. M., Hellgren, E. C., Van Den Bussche, R. A., & Silvy, N. J. (1999). Demography, sociospatial behaviour and genetics of feral pigs (*Sus scrofa*) in a semi-arid environment. *Journal of Zoology*, 247(3), 311–322. https://doi.org/10.1111/j.1469-7998.1999.tb00994.x

Ghosh, S., Birrell, P. J., & De Angelis, D. (2023). An approximate diffusion process for environmental stochasticity in infectious disease transmission modelling. *PLOS Computational Biology*, 19(5), e1011088. https://doi.org/10.1371/journal.pcbi.1011088

Gwenzi, W., Skirmuntt, E. C., Musvuugwa, T., Teta, C., Halabowski, D., & Rzymski, P. (2022). Grappling with (re)-emerging infectious zoonoses: Risk assessment, mitigation framework, and future directions. *International Journal of Disaster Risk Reduction*, 82, 103350. https://doi.org/10.1016/j.ijdrr.2022.103350

Herraiz, C., Ferrer-Ferrando, D., Vicente, J., & Acevedo, P. (2024a). Camera trapping and telemetry for detecting and quantifying animal interactions: Not anything goes. *Ecological Indicators*, *160*, 111877.

Herraiz, C., Triguero-Ocaña, R., Laguna, E., Jiménez-Ruiz, S., Peralbo-Moreno, A., Martínez-López, B., García-Bocanegra, I., Risalde, M. Á., Vicente, J., & Acevedo, P. (2024b). Movement-driven modelling reveals new patterns in disease transmission networks. *Journal of Animal Ecology*, 93, 1275–1287. https://doi.org/10.1111/1365-2656.14142

Hoyt, J. R., Parise, K. L., DePue, J. E., Kaarakka, H. M., Redell, J. A., Scullon, W. H., O'Reskie, R., Foster, J. T., Kilpatrick, A. M., Langwig, K. E., & White, J. P. (2023). Reducing environmentally mediated transmission to moderate impacts of an emerging wildlife disease. *Journal of Applied Ecology*, 60(5), 923–933. https://doi.org/10.1111/1365-2664.14371

Kaminski, G., Brandt, S., Baubet, E., & Baudoin, C. (2005). Life-history patterns in female wild boars (*Sus scrofa*): Mother-daughter postweaning associations. *Canadian Journal of Zoology*, 83(3), 474–480. https://doi.org/10.1139/z05-019

Kelt, D. A., Heske, E. J., Lambin, X., Oli, M. K., Orrock, J. L., Ozgul, A., & Sommer, S. (2019). Advances in population ecology and species interactions in mammals. *Journal of Mammalogy*, 100(3), 965–1007. https://doi.org/10.1093/jmammal/gyz017

Kenward, R., Hodder, K. H., Rose, R., Walls, C., Parish, T., Holm, J., & Doyle, F. (1998). Comparative demography of red squirrels (*Sciurus vulgaris*) and grey squirrels (*Sciurus carolinensis*) in deciduous and conifer woodland. *Journal of Zoology*, 244(1), 7–21. https://doi.org/10.1111/j.1469-7998.1998.tb00002.x

Khong, V. H., Carmona, P., & Gandon, S. (2023). Seasonality and the persistence of vector-borne pathogens. *Journal of the Royal Society Interface*, 20(209), 20230470. https://doi.org/10.1098/rsif.2023.0470

Khurana, S. K., Sehrawat, A., Tiwari, R., Prasad, M., Gulati, B., Shabbir, M. Z., & Chaicumpa, W. (2021). Bovine brucellosis–a comprehensive review. *Veterinary Quarterly*, 41(1), 61–88. https://doi.org/10.1080/01652176.2020.1868616

Kukielka, E., Barasona, J. A., Cowie, C. E., Drewe, J. A., Gortazar, C., Cotarelo, I., & Vicente, J. (2013). Spatial and temporal interactions between livestock and wildlife in South Central Spain assessed by camera traps. *Preventive veterinary medicine*, *112*(3-4), 213–221. https://doi.org/10.1016/j.prevetmed.2013.08.008

Laggan, N. A., Parise, K. L., White, J. P., Kaarakka, H. M., Redell, J. A., DePue, J. E.,& Hoyt, J. R. (2023). Host infection and disease-induced mortality modify species contributions to the environmental reservoir. *Ecology*, 104(10), e4147. https://doi.org/10.1002/ecy.4147

Lambin, E. F., Tran, A., Vanwambeke, S. O., Linard, C., & Soti, V. (2010). Pathogenic landscapes: interactions between land, people, disease vectors, and their animal hosts. *International Journal of Health Geographics*, 9, 1–13. https://doi.org/10.1186/1476-072X-9-54

Lavelle, Michael J., et al. "Utility of improvised video-camera collars for collecting contact data from white-tailed deer: Possibilities in disease transmission studies." *Wildlife Society Bulletin* 36.4 (2012): 828-834.

Lavelle, Michael J., et al. "Assessing risk of disease transmission: direct implications for an indirect science." *BioScience* 64.6 (2014): 524-530.

Leung, N. H. (2021). Transmissibility and transmission of respiratory viruses. *Nature Reviews Microbiology*, 19(8), 528–545. https://doi.org/10.1038/s41579-021-00535-6

Long, J. A., Webb, S. L., Harju, S. M., & Gee, K. L. (2022). Analyzing contacts and behavior from high frequency tracking data using the wildlifeDI R package. *Geographical Analysis*, 54(3), 648–663. https://doi.org/10.1111/gean.12303

Lou, H. L. (1995). Implementing the Viterbi algorithm. *IEEE Signal Processing Magazine*, 12(5), 42–52. 10.1109/79.410439

Manlove, K., Wilber, M., White, L., Bastille-Rousseau, G., Yang, A., Gilbertson, M. L., Craft, M. E., Cross, P. C., Wittemyer, G., & Pepin, K. M. (2022). Defining an epidemiological landscape that connects movement ecology to pathogen transmission and pace-of-life. *Ecology Letters*, 25(8), 1760–1782. 10.1111/ele.14032

Miguel, E., Grosbois, V., Caron, A., Pople, D., Roche, B., & Donnelly, C. A. (2020). A systemic approach to assess the potential and risks of wildlife culling for infectious disease control. *Communications Biology*, 3(1), 353. https://doi.org/10.1038/s42003-020-1032-z

Murray, M. H., Becker, D. J., Hall, R. J., & Hernandez, S. M. (2016). Wildlife health and supplemental feeding: a review and management recommendations. *Biological Conservation*, 204, 163–174. https://doi.org/10.1016/j.biocon.2016.10.034

Noonan, M. J., Martinez-Garcia, R., Davis, G. H., Crofoot, M. C., Kays, R., Hirsch, B. T., & Sinn, D. L. (2021). Estimating encounter location distributions from animal tracking data. *Methods in Ecology and Evolution*, 12(7), 1158–1173.https://doi.org/10.1111/2041-210X.13597

Ogbunugafor, C. B., Miller-Dickson, M. D., Meszaros, V. A., Gomez, L. M., Murillo, A. L., & Scarpino, S. V. (2020). Variation in microparasite free-living survival and indirect transmission can modulate the intensity of emerging outbreaks. *Scientific Reports*, 10(1), 20786. https://doi.org/10.1038/s41598-020-77048-4

Otero, A., Velásquez, C. D., Aiken, J., & McKenzie, D. (2021). Chronic wasting disease: a cervid prion infection looming to spillover. *Veterinary Research*, 52(1), 115. https://doi.org/10.1186/s13567-021-00986-y

Pandey, A., Wojan, C., Feuka, A., Craft, M. E., Manlove, K., & Pepin, K. M. (2024). The influence of social and spatial processes on the epidemiology of environmentally transmitted pathogens in wildlife: implications for management. *Philosophical Transactions B*, 379(1912), 20220532. https://doi.org/10.1098/rstb.2022.0532

Parratt, S. R., Numminen, E., & Laine, A. L. (2016). Infectious disease dynamics in heterogeneous landscapes. *Annual Review of Ecology, Evolution, and Systematics*, 47(1), 283–306. https://doi.org/10.1146/annurev-ecolsys-121415-032321

Pepin, K. M., Golnar, A., & Podgórski, T. (2021). Social structure defines spatial transmission of African swine fever in wild boar. *Journal of the Royal Society Interface*, 18(174), 20200761. https://doi.org/10.1098/rsif.2020.0761

Pepin, K. M., Leach, C. B., Barrett, N. L., Ellis, J. W., VanDalen, K. K., Webb, C. T., & Shriner, S. A. (2023). Environmental transmission of influenza A virus in mallards. *mBio*, 14(5), e00862-23. https://doi.org/10.1128/mbio.00862-23

Pinn-Woodcock, T., Frye, E., Guarino, C., Franklin-Guild, R., Newman, A. P., Bennett, J., & Goodrich, E. L. (2023). A one-health review on brucellosis in the United States. *Journal of the American Veterinary Medical Association*, 261(4), 451–462. https://doi.org/10.2460/javma.23.01.0033

Plowright, R. K., Parrish, C. R., McCallum, H., Hudson, P. J., Ko, A. I., Graham, A. L., & Lloyd-Smith, J. O. (2017). Pathways to zoonotic spillover. *Nature Reviews Microbiology*, 15(8), 502–510. https://doi.org/10.1038/nrmicro.2017.45

Podgórski, T., Apollonio, M., & Keuling, O. (2018). Contact rates in wild boar populations: Implications for disease transmission. *The Journal of Wildlife Management*, 82(6), 1210–1218. https://doi.org/10.1038/nrmicro.2017.45

Podgórski, T., Lusseau, D., Scandura, M., Sonnichsen, L., & Jędrzejewska, B. (2014). Long-lasting, kin-directed female interactions in a spatially structured wild boar social network. *PLOS ONE*, 9, e99875. https://doi.org/10.1371/journal.pone.0099875

Pohle, J., Langrock, R., Van Beest, F. M., & Schmidt, N. M. (2017). Selecting the number of states in hidden Markov models: pragmatic solutions illustrated using animal movement. *Journal of Agricultural, Biological and Environmental Statistics*, 22, 270–293. https://doi.org/10.1007/s13253-017-0283-8

Potts, Jonathan R., et al. "Assessing the predictive power of step selection functions: How social and environmental interactions affect animal space use." *Methods in Ecology and Evolution,* 13.8 (2022): 1805-1818.

Potts, Jonathan R., and Luca Börger. "How to scale up from animal movement decisions to spatiotemporal patterns: An approach via step selection." *Journal of Animal Ecology*, 92.1 (2023): 16-29.

Rabiner, L., & Juang, B. (1986). An introduction to hidden Markov models. *IEEE ASSP Magazine*, 3(1), 4–16. 10.1109/MASSP.1986.1165342

Rebollada-Merino, A., Pérez-Sancho, M., Rodríguez-Bertos, A., García, N., Martínez, I., Navarro, A., & García-Seco, T. (2022). Environment and offspring surveillance in porcine brucellosis. *Frontiers in Veterinary Science*, 9, 915692. https://doi.org/10.3389/fvets.2022.915692

Rees, E. M., Minter, A., Edmunds, W. J., Lau, C. L., Kucharski, A. J., & Lowe, R. (2021). Transmission modelling of environmentally persistent zoonotic diseases: a systematic review. *The Lancet Planetary Health*, 5(7), e466–e478. https://doi.org/10.1016/S2542-5196(21)00137-6

Richardson, T. O., & Gorochowski, T. E. (2015). Beyond contact-based transmission networks: the role of spatial coincidence. *Journal of the Royal Society Interface*, 12(111), 20150705. https://doi.org/10.1098/rsif.2015.0705

Rocha, S. C., Velásquez, C. V., Aquib, A., Al-Nazal, A., & Parveen, N. (2022). Transmission cycle of tick-borne infections and co-infections, animal models and diseases. *Pathogens*, 11(11), 1309. https://doi.org/10.3390/pathogens11111309

Salman, M., & Steneroden, K. (2022). Important Zoonotic Diseases of Cattle and Their Prevention Measures. In *Zoonoses: Infections Affecting Humans and Animals* (pp. 1–22). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-85877-3_1-1

Silk, M. J., Drewe, J. A., Delahay, R. J., Weber, N., Steward, L. C., Wilson-Aggarwal, J., ... & McDonald, R. A. (2018). Quantifying direct and indirect contacts for the potential transmission of infection between species using a multilayer contact network. *Behaviour*, 155(7–9), 731–757. https://doi.org/10.1163/1568539X-00003493

Siva-Jothy, J. A., & Vale, P. F. (2021). Dissecting genetic and sex-specific sources of host heterogeneity in pathogen shedding and spread. *PLOS Pathogens*, 17(1), e1009196. https://doi.org/10.1371/journal.ppat.1009196

Smith JA, Suraci JP, Hunter JS, et al. Zooming in on mechanistic predator–prey ecology: Integrating camera traps with experimental methods to reveal the drivers of ecological interactions. *Journal of Animal Ecology*. 2020; 89: 1997–2012. https://doi.org/10.1111/1365-2656.13264

Smith, O. M., Snyder, W. E., & Owen, J. P. (2020). Are we overestimating risk of enteric pathogen spillover from wild birds to humans? *Biological Reviews*, 95(3), 652–679. https://doi.org/10.1111/brv.12581

Suh, D. C., Lance, S. L., & Park, A. W. (2024). Abiotic and biotic factors jointly influence the contact and environmental transmission of a generalist pathogen. *Ecology and Evolution*, 14(8), e70167. https://doi.org/10.1002/ece3.70167

Thurfjell, H., Ciuti, S., & Boyce, M. S. (2014). Applications of step-selection functions in ecology and conservation. *Movement Ecology*, 2(1), 4.

Titcomb, G., Mantas, J. N., Hulke, J., Rodriguez, I., Branch, D., & Young, H. (2021). Water sources aggregate parasites with increasing effects in more arid conditions. *Nature Communications*, 12(1), 7066. https://doi.org/10.1038/s41467-021-27352-y

Triguero-Ocaña, R., Vicente, J., Lavelle, M., & Acevedo, P. (2021). Collecting data to assess the interactions between livestock and wildlife. In *Diseases at the Wildlife–Livestock Interface: Research and Perspectives in a Changing World*, 307–338. https://doi.org/10.1007/978-3-030-65365-1_10

VanderWaal, K. L., & Ezenwa, V. O. (2016). Heterogeneity in pathogen transmission. *Functional Ecology*, 30(10), 1606–1622. https://doi.org/10.1111/1365-2435.12645

Virlogeux, V., Li, M., Tsang, T. K., Feng, L., Fang, V. J., Jiang, H., & Cowling, B. J. (2015). Estimating the distribution of the incubation periods of human avian influenza A (H7N9)

virus infections. *American Journal of Epidemiology*, 182(8), 723–729. https://doi.org/10.1093/aje/kwv115

Wang, C. C., Prather, K. A., Sznitman, J., Jimenez, J. L., Lakdawala, S. S., Tufekci, Z., & Marr, L. C. (2021). Airborne transmission of respiratory viruses. *Science*, 373(6558), eabd9149. https://doi.org/10.1126/science.abd9149

Webster, J. P., Borlase, A., & Rudge, J. W. (2017). Who acquires infection from whom and how? Disentangling multi-host and multi-mode transmission dynamics in the 'elimination' era. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 372(1719), 20160091. https://doi.org/10.1098/rstb.2016.0091

White, L. A., Forester, J. D., & Craft, M. E. (2017). Using contact networks to explore mechanisms of parasite transmission in wildlife. *Biological Reviews*, 92(1), 389–409. https://doi.org/10.1111/brv.12236

Wilber, M. Q., Yang, A., Boughton, R., Manlove, K. R., Miller, R. S., Pepin, K. M., & Wittemyer, G. (2022). A model for leveraging animal movement to understand spatio-temporal disease dynamics. *Ecology Letters*, 25(5), 1290–1304. https://doi.org/10.1111/ele.13986

Wißmann, J. E., Kirchhoff, L., Brüggemann, Y., Todt, D., Steinmann, J., & Steinmann, E. (2021). Persistence of pathogens on inanimate surfaces: a narrative review. *Microorganisms*, 9(2), 343. https://doi.org/10.3390/microorganisms9020343

Yang, A., Proffitt, K. M., Asher, V., Ryan, S. J., & Blackburn, J. K. (2021). Sex-specific elk resource selection during the anthrax risk period. *Journal of Wildlife Management*, 85, 145–155. https://doi.org/10.1002/jwmg.21952

Yang, A., Wilber, M. Q., Manlove, K. R., Miller, R. S., Boughton, R., Beasley, J., ... & Pepin, K. M. (2023). Deriving spatially explicit direct and indirect interaction networks from animal movement data. *Ecology and Evolution*, 13(3), e9774. https://doi.org/10.1002/ece3.9774

Yang, A., Yang, A., Schlichting, P. E., Wight, B. R., Anderson, W., Chinn, S. M., ... & Pepin, K. M. (2021). Effects of social structure and management on risk of disease establishment in wild pigs. *Journal of Animal Ecology*, 90(4), 820–833. https://doi.org/10.1111/1365-2656.13412

Chapter 3 : Predicting Cross-species Avian Influenza Transmission Between Migratory Waterfowl and Feral Pigs Through a Multi-species Occupancy Model to Inform Risk-based Surveillance in The United States

Keywords: Avian influenza, Cross-species transmission, Feral pigs, Multi-species occupancy model, Wildlife disease surveillance

Prepared for Wiley's Ecological Applications

Abstract

Avian influenza continues to expand across North America, increasing opportunities for transmission among wildlife, livestock, and other susceptible hosts. Although migratory waterfowl are recognized reservoirs of avian influenza viruses, the role of feral pigs within transmission systems remains poorly understood despite their widespread distribution, growing population, and extensive overlap with waterfowl habitats throughout the United States. Because feral pigs frequently utilize wetlands, floodplains, and riparian habitats occupied by migratory birds, they may facilitate cross-species transmission through repeated exposure to environmentally contaminated aquatic systems. Identifying the environmental and ecological conditions associated with bird- pig transmission is therefore critical for improving surveillance at the wildlife–livestock interface. We developed a Bayesian multi-species occupancy model to quantify cross-species avian influenza transmission between migratory waterfowl and feral pigs across the contiguous United States. Laboratory-confirmed avian influenza surveillance data from wild birds and feral pigs collected by the United States Department of Agriculture between 2010 and 2015 were analyzed at watershed scale. Dynamic environmental variables, hydrological connectivity metrics, wetland characteristics, and water-associated avian host assemblages were incorporated to estimate transmission probability and generate seasonal prediction surfaces. Our findings suggest that cross-species occupancy probability increased with wetland area, hydrological connectivity, soil moisture, vapor pressure, and dabbling duck abundance, indicating that transmission is most likely within environmentally connected wetland systems where host overlap and environmental persistence are maximized. Feral pig-associated variables were among the strongest predictors of occupancy probability, highlighting the

potential epidemiological importance of feral pigs within avian influenza transmission landscapes. Predicted transmission hotspots were concentrated in Texas, the Gulf Coast, southern California, and portions of the Mississippi River Basin, where extensive wetlands, major migratory flyways, and established feral pig populations coincide. Our findings suggest that feral pigs represent an important but underutilized component of avian influenza surveillance. By identifying high-risk watersheds for targeted surveillance, this study provides a practical framework for improving early detection of avian influenza transmission at the wildlife–livestock interface.

3.1 Introduction

Avian influenza can infect a wide range of avian and mammalian hosts, including migratory waterfowl, shorebirds, poultry, swine, carnivores, marine mammals, and humans, creating complex transmission networks across wildlife, livestock, and human-associated environments (Runstadler & Puryear, 2024). It has become one of the most significant transboundary wildlife diseases, which threatens the domestic poultry industry, affects ecosystem function, and potentially plays a role in the emergence of human influenza (Verhagen, Fouchier et al. 2021, Ramey, Hill et al. 2022, Runstadler and Puryear 2024). Single, low pathogenic avian influenza can cause a minimum of $131 million estimated loss in domestic poultry (Capua and Alexander, 2004), while high pathogenic avian influenza can cause more than $1.15 billion in response and indemnity costs and up to $3.3 billion in trade-related losses associated with poultry outbreaks and disease control measures (U.S. Department of Agriculture 2016). The recent highly pathogenic avian influenza (HPAI) H5N1 clade 2.3.4.4b outbreaks in North America have shown the remarkable capacity of the virus for geographic expansion via migratory birds' movement and environmental persistence (Caliendo, Lewis et al. 2022, Harvey,

Mullinax et al. 2023, Renaud, Osborn et al. 2023, Brüssow 2024). Following its re-introduction into North America in late 2021, the virus was detected across numerous U.S. states and Canadian provinces within months, resulting in unprecedented mortality in wild birds and substantial impacts on poultry production systems (Caliendo et al., 2022; Renaud et al., 2023). The continued expansion of HPAI across wildlife hosts has intensified concern regarding the ecological mechanisms that facilitate long-distance viral dispersal, environmental persistence, and cross-species spillover within environmentally connected landscapes (Mena, von Fricken et al. 2025).

Wild aquatic birds, particularly migratory waterfowl, are recognized as the primary natural reservoirs of avian influenza and play a central role in continental-scale viral dissemination (Fourment, Darling et al. 2017, Blagodatski, Trutneva et al. 2021, Ramey, Hill et al. 2022). To monitor the dynamics of avian influenza virus in these wild birds, the United States has maintained national avian influenza surveillance programs in wild birds since 2006 through the USDA National Wildlife Disease Program (NWDP) (Bevins et al., 2014). These surveillance efforts focus heavily on migratory waterfowl because of their large population sizes, broad geographic distributions, and long-distance seasonal migrations that connect distant ecosystems and flyways (U.S. Department of Agriculture, 2013; U.S. Department of Agriculture, 2015; U.S. Department of Agriculture, 2022). However, surveillance coverage remains spatially uneven due to logistical and financial limitations that restrict annual sampling intensity across the country. As a result, surveillance programs must prioritize high-risk regions and allocate limited resources strategically to maximize detection efficiency. Previous studies have demonstrated that environmental conditions such as wetland availability, precipitation patterns, temperature variability, and waterfowl abundance strongly influence avian influenza occurrence and

environmental persistence (Ferenczi, Beckmann et al. 2016, Wu, Perrings et al. 2020, Gorsich, Webb et al. 2021). Seasonal migrations of birds along continental flyways result in birds concentrating in common stopover areas where high concentrations of birds and repeated environmental contact can facilitate viral transmission and long-distance dispersal (Blagodatski, Trutneva et al. 2021). These movements link geographically remote ecosystems and enable the spread of avian influenza viruses over wide geographic areas over relatively short time scales (Fourment, Darling et al. 2017, Gorsich, Webb et al. 2021). Thus, it is crucial to understand how environmental structure and dynamics influence transmission risk to identify high-risk transmission landscapes for improving wildlife surveillance.

Among the wildlife hosts of avian influenza, feral pigs (*Sus scrofa*) represent a potentially important but understudied component of its transmission system. Feral pigs are susceptible to both avian and mammalian influenza viruses because they have compatible cellular receptors, which has raised concerns about their role as potential bridge hosts in cross-species transmission of avian influenza (Nelli, Kuchipudi et al. 2010, Martin, Sun et al. 2017, Bourret 2018). Feral pig populations have grown substantially in the United States, their population exceeds approximately 6.9 million animals distributed across at least 38 states in recent decades and are found in diverse ecological habitats, such as wetlands, agricultural areas, riparian zones, and forested areas, often in close proximity to migratory bird habitats (Brown, Marlow et al. 2019, Lewis, Corn et al. 2019). The shared habitat provides potential opportunities for indirect contact between feral pigs and migratory waterfowl, especially in the wetland associated environments where avian influenza viruses have been found to survive for 12-20 days in water, depending on the temperature and contaminated substrates (Domanska-Blicharz et al., 2010). Although active influenza surveillance has been conducted in the past in domestic swine production systems,

surveillance of influenza viruses in feral swine remains limited and inconsistent. Programs such as the National Feral Swine Damage Management Program (NFSP) and NWDP collected biological samples during feral swine management and removal activities, but avian influenza virus testing was conducted sporadically relative to surveillance in poultry and wild birds. Previous studies estimated an overall influenza A virus seroprevalence of approximately 4.9% in U.S. feral swine populations during 2010–2013, suggesting widespread but incompletely characterized exposure (Martin et al., 2017). Despite increasing evidence of exposure, the ecological role of feral swine in avian influenza transmission remains poorly understood. Most previous studies have been largely centered on single host systems, especially wild birds, although a smaller number have also examined avian influenza occurrence in mammalian hosts (Martin et al., 2017; Caserta et al., 2024). However, no studies, to our best knowledge, have explicitly evaluated the potential role of feral swine within cross-species avian influenza transmission network across shared environmental landscapes. Thus, there is an urgent need to understand this overlooked component in the avian influenza transmission system.

To address this gap, we developed a Bayesian multi-species occupancy modeling framework to investigate cross-species avian influenza occurrence in wild birds and feral pigs throughout the contiguous United States. We integrated wildlife surveillance data, dynamic climatic variables, wetland distribution, hydrological connectivity metrics and bird community abundance estimates to quantify cross-species transmission potential between migratory wild birds and feral pigs, and evaluating the ecological and environmental drivers shaping infection risk across host communities and assessing how dynamic environmental conditions influence seasonal avian influenza occurrence across heterogeneous landscapes. By integrating surveillance data and environmental variability into a single hierarchical framework, this study

offers a large-scale ecological assessment of avian influenza dynamics relevant to control strategies, multi-host surveillance, and environmental management.

3.2 Methodology

3.2.1 Study Areas and Spatial Unit

This study focuses on watersheds within the contiguous US. The spatial unit Hydrologic Unit Code 8 (HUC-8) watersheds, which offer ecologically significant units that reflect hydrological connectivity, wetland distribution, and landscape heterogeneity that are important to host movement and environmental persistence of pathogens. This spatial framework enables the study to model wetland systems as interrelated ecological units in which indirect interactions between migratory birds and feral pigs can take place. Temporal variation was incorporated by arranging the dataset into seasonal intervals across study period, enabling the model to capture intra-seasonal fluctuations in climate, hydrology, and host aggregation patterns that can impact environmentally mediated transmission.

3.2.2 ***Surveillance Data***

The 2010 – 2015 avian influenza surveillance data, georeferenced at Hydrologic Unit Code 8 (HUC-8) watershed-level were sourced from the United States Department of Agriculture (USDA), which included laboratory-confirmed infection data for both wild birds and feral pigs. The surveillance records consisted of individual animal-level diagnostic observations, including laboratory-confirmed test results for avian influenza virus infection. Wild bird surveillance included samples collected through hunter-harvest surveillance, targeted field sampling, morbidity and mortality investigations, and opportunistic surveillance efforts, following USDA national surveillance protocols and reporting frameworks (USDA NWDP, 2015). Feral swine

surveillance data were obtained from targeted monitoring and diagnostic testing programs in which swab or serum samples were collected and tested using standard laboratory procedures. Infection status for both host groups was determined by using laboratory-based diagnostic assays with high specificity for avian influenza virus detection. The study period was limited to 2010–2015 because this represented the period of overlap between avian influenza surveillance efforts targeting wild birds and feral pigs, enabling concurrent assessment of infection occurrence across both host groups.

The surveillance data was then screened and filtered to remove incomplete observations including missing diagnostic test results, records that have unspecified or missing species identification and duplicate entries. Then we kept only those surveillance records for bird species that have at least 100 samples tested. Species represented by fewer than 100 surveillance records were excluded to reduce uncertainty associated with sparse detection histories and improve the reliability of occupancy parameter estimation. This threshold was selected based on the highly skewed distribution of species-specific sample sizes (Figure S3-1) and evidence that occupancy-based inference becomes increasingly unreliable under limited sample sizes (Cowans et al., 2025). We have further organized the individual level data by HUC-8 and season to construct detection histories and sampling effort per HUC-8. This approach allowed individual surveillance observations collected through heterogeneous sampling programs to be integrated within a consistent watershed-scale hierarchical modeling framework while preserving the observation process required for occupancy analyses.

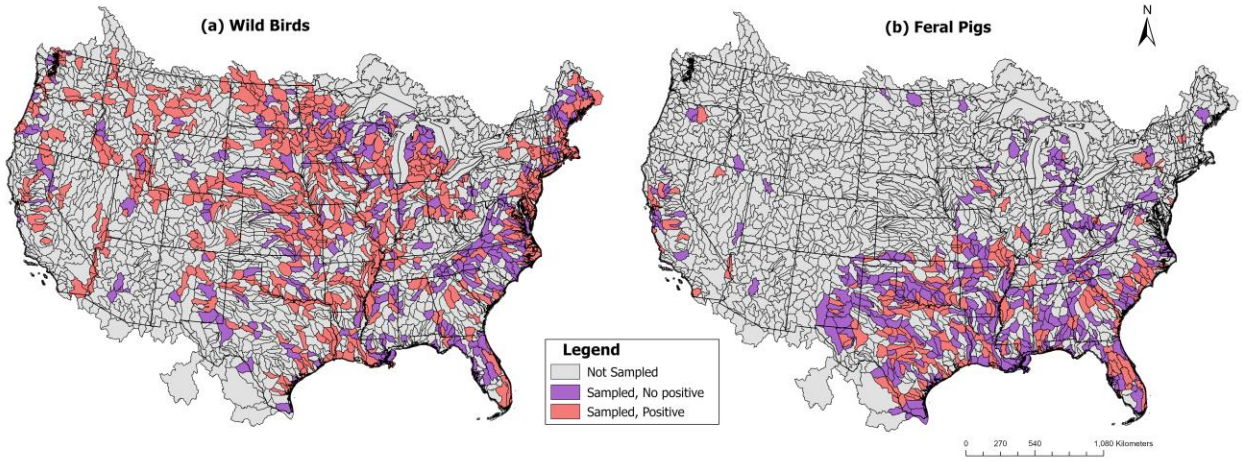


Figure 3-1: Distribution of surveillance data collected from USDA between 2010-2015 for (a) Wild birds, (b) feral pigs. The data for 2012-2013 were missing in birds due to no surveillance records.

3.2.3 Covariates Data

To explore the environmental and ecological drivers that affect avian influenza transmission, we compiled a suite of covariates that represent climate, landscape structure, hydrology, connectivity, and host community (Table 3-1). Climate variables were obtained from the Daymet gridded data product and included precipitation-related variables, temperature variability, relative humidity variability, solar radiation, and vapor pressure (Thornton, Shrestha et al. 2022). These climatic factors can influence avian influenza virus survival, seasonal wetland availability, hydrological connectivity, and migratory waterfowl movement and aggregation patterns (Gorsich et al., 2021; Scolamacchia et al., 2021; Ferenczi et al., 2016). We also considered soil-related variables sourced from SoilGrids, including soil moisture, soil pH, and soil organic matter (Poggio, De Sousa et al. 2021). Soil and surface characteristics can influence viral persistence in wetland sediments, flood-prone landscapes, and contaminated terrestrial environments where migratory birds and feral swine frequently interact (Mateus-Anzola et al., 2024; Gorsich et al., 2021; Ferenczi et al., 2016).

To describe the dynamics of wetlands and the state of surface water, we used remote sensing products of the JRC Global Surface Water dataset (Hao, Cai et al. 2024). Pixels with greater than 50% wetland cover were classified as wetlands. The fraction of wetland pixels within one spatial unit was subsequently used to calculate total wetland area, proportional wetland cover, water occurrence frequency, and inundation variability. We also derived landscape configuration measures, including patch density, edge density, and mean patch area, based on those seasonal JRC-derived wetland layers. These metrics characterize the extent, persistence, fragmentation, and hydrological connectivity of wetland habitats that influence waterfowl aggregation patterns, and indirect exposure opportunities for feral swine within shared aquatic landscapes (Hubbard et al., 2023; Gilbert & Pfeiffer, 2012). We also used satellite-derived indices, Normalized Difference Vegetation Index (NDVI) and Land Surface Water Index (LSWI), calculated using MODIS products (MOD13A1 and MOD09A1, respectively) at 500 m resolution. NDVI represents vegetation quality and productivity associated with waterfowl habitat use, while LSWI captures surface wetness conditions associated with environmental persistence of avian influenza viruses (Hogerwerf et al., 2010).

Table 3-1: Description of covariates used in the study

Variable	Source	Definition / Metric	Ecological Relevance for HPAI
Wetland Area		Total wetland area (ha) within HUC	Determines habitat availability for waterfowl reservoirs and environmental viral deposition
% Wetland Cover		(Wetland area / HUC area) × 100	Reflects landscape-level aggregation potential of wild birds
Water Frequency (Hydroperiod)		Mean seasonal DSWE water presence	Influences viral persistence and host congregation duration
Inundation Variability		SD of water frequency across years	Captures hydrologic instability affecting exposure consistency
Mean Patch Area		Average wetland patch size	Larger patches support higher bird density and mixing
Patch Density		Number of wetland patches per area	Indicates habitat fragmentation and movement between sites

Edge Density	Total wetland edge length / HUC area	Represents contact interfaces between wetlands and agriculture
eBird abundance	Mean and relative abundance of bird groups	Capture host community composition and density
Major Crop Type	Dominant agricultural crop class within HUC8	Influences wild bird attraction, foraging behavior, and indirect contact with poultry systems
Wetland Connectivity	Graph/distance-based connectivity metric	Facilitates pathogen spread across wetland networks
Temperature (SD, Range)	Variability and range per season	Affects viral survival and host physiology
Precipitation (# Wet Weeks, SPI)	Frequency and intensity anomalies	Drives flooding, water persistence, and contamination spread
Relative Humidity (SD, THI)	Variability in humidity and thermal stress	Influences aerosol stability and poultry susceptibility
Vapor Pressure (SD)	Atmospheric moisture variability	Linked to viral persistence in air and environment
Solar Radiation (SD)	Variability in radiation	UV exposure impacts viral inactivation rates
Wind Speed (SD)	Variability in wind	Supports airborne dispersal between farms
NDVI (Mean)	Average vegetation greenness	Proxy for habitat suitability and bird use intensity
LSWI (Mean)	Surface water–vegetation moisture index	Indicates wet surface conditions favorable for persistence
Poultry Density	Birds per unit area	Determines amplification potential and outbreak magnitude
Feral Pig Density	Pigs per unit area	Host availability and contact intensity
Network Centrality Metrics	In-/Out-degree, betweenness, bridging index, flow centrality	Identifies movement corridors and high-risk transmission nodes
Soil Moisture	Mean surface volumetric water content (0–10 cm)	Controls environmental persistence and surface contamination of virus
Soil Organic Matter	% organic carbon in topsoil	Enhances viral adsorption and protection from UV degradation
Soil Salinity	Electrical conductivity (dS/m)	Influences viral stability and wetland habitat suitability
Soil pH	Soil acidity/alkalinity (pH units)	Viral survival sensitive to extreme pH conditions

To capture large-scale connectivity, we built movement networks using Bird Banding Laboratory data (Smith 2013), with edges representing movements between HUC-8 watersheds. Given that

the original movement data were sparse, to get a general trend of birds' seasonal movement, we used a five-year rolling window for each season to calculate birds' movement connectivity. For each year and season, we combined movements from the previous two years and the following two years, and summed edges by origin-destination pairs to create weighted networks and computed several network metrics, such as in-degree, out-degree, betweenness centrality, eigenvector centrality, flow centrality, and bridging index (Bodin & Norberg, 2007; Freeman et al., 1991; Minor et al., 2008).

Finally, we used bird abundance data from raw eBird data to represent host community structure (Sullivan, Aycrigg et al. 2014). These data were grouped by functional group (dabbling ducks, diving ducks, geese and swans, gulls, pelagic seabirds, and shorebirds) as defined by the most important species group for avian influenza surveillance by USDA (USDA, 2015). These groupings also reduced sparsity among individual species observations while retaining ecologically meaningful differences in host assemblages. We also included the density of feral pigs as a host-related covariate, using the predicted density data (Lewis, Corn et al. 2019). To capture the domestic transmission interface, we also included annual poultry density, derived from agricultural census data (USDA NASS, 2017).

All environmental and ecological variables were aggregated at HUC-8 watershed level and grouped by seasons to align with the spatial and temporal resolution of the surveillance data. All covariates were standardized, and collinearity among variables were evaluated using pairwise correlation test prior to model fitting. For variables that were highly correlated ($|r| > 0.70$), the predictor with higher variance inflation factor (VIF) was removed to avoid redundancy (Dormann et al., 2012; Zuur et al., 2010) and ensure that the remaining predictors captured different aspects of the system.

3.2.4 Occupancy Models

We developed a hierarchical Bayesian cross-species occupancy framework, based on the principles of multi-species occupancy modeling, to quantify the occurrence of avian influenza and explicitly test the dynamics of cross-species transmission probability between migratory wild birds and feral pigs in dynamic environmental conditions (Devarajan, Morelli et al. 2020). This framework isolates the ecological process that controls the actual presence of pathogens and the observation process that is linked with imperfect detection, which enables sound inference of wildlife surveillance data that are incomplete and heterogeneous by nature (Doser, Finley et al. 2022). Here, the actual infection state is a latent (unobserved) variable, and the observed detections are conditionally dependent on this latent process (Doser, Finley et al. 2023). This is a critical requirement of wildlife disease systems, in which non-detection does not always mean actual absence and in which sampling effort is uneven across space and time (Doser, Finley et al. 2022). This part of the analysis were conducted in R (version 4.4.1) (R Core Team., 2024) using the *spOccupancy* package, which provides efficient Bayesian estimation for both spatial and non-spatial multi-species occupancy models (Doser, Finley et al. 2022).

For each species s, site i, and time t, the true infection state is represented by a latent variable:

$$z_{i,t,s} \sim Bernoulli(\psi_{i,t,s})$$

where $\psi_{i,t,s}$ denotes the probability of avian influenza occurrence. This probability is modeled as a function of environmental covariates using a logit link:

$$logit\left(\psi_{i,t,s}\right) = \beta_{0,s} + \sum_{k=1}^{K} \beta_{k,s} X_{i,t,k} + \omega_i$$

where $X_{i,t,k}$ represents the set of standardized environmental and ecological covariates, $\beta_{k,s}$ are species-specific coefficients, and ω_i is a spatial random effect capturing residual spatial dependence among sites.

The observation process links the latent infection state to the observed detection data:

$$y_{i,t,j,s} \sim Bernoulli(p_{i,t,j,s} \,.\, z_{i,t,s})$$

$$logit\,(p_{i,t,j,s}) = \gamma_{0,s}$$

where $y_{i,t,j,s}$ is the detection/non-detection outcome for replicate j, and $p_{i,t,j,s}$ is the detection probability. Detection was modeled with species-specific intercepts, reflecting differences in surveillance processes across taxa.

Table 3-2: Spatial covariance functions, non-spatial, and approximation methods used in the occupancy model

Component	Type	Description	Interpretation
Model Structure	Non-spatial	No spatial random effect included	Assumes independence among sites
	Spatial (GP)	Full Gaussian Process using all pairwise distances	Captures continuous spatial dependence across all sites
	Spatial (NNGP)	Nearest Neighbor Gaussian Process	Approximates GP using local neighbors for computational efficiency
Covariance Function	Gaussian	Smooth decay of correlation with distance (Datta, Banerjee et al. 2016, Finley, Datta et al. 2019)	Strong long-range spatial influence
	Exponential	Rapid decay with distance (Datta, Banerjee et al. 2016, Finley, Datta et al. 2019)	Local spatial dependence dominates
	Spherical	Finite range with cutoff distance (Finley, Datta et al. 2019)	No correlation beyond a threshold
	Matérn	Flexible covariance with smoothness parameter (ν) (Lindgren, Rue et al. 2011, Fuglstad, Simpson et al. 2019)	Balances local and global spatial structure
NNGP Settings	Neighbor size (k = 5, 10, 15)	Number of nearest neighbors used in approximation	Controls trade-off between accuracy and computation

Spatial dependence was modelled using Gaussian process (GP) random effects, via the spatial extension of the model in *spOccupancy*. We tested four covariance functions, including Gaussian, exponential, spherical, and Matérn, that represent different spatial dependence structures (Table 2) (Lindgren, Rue et al. 2011, Fuglstad, Simpson et al. 2019). The models were fitted using both full Gaussian process models and Nearest Neighbor Gaussian Process (NNGP) models to reduce computational time for large spatial datasets (Datta, Banerjee et al. 2016, Finley, Datta et al. 2019). The full Gaussian process (GP) models spatial dependence using a covariance matrix among all sampling locations, whereas NNGP approximates spatial correlation with a fixed number of nearest neighbors (5, 10 and 15 neighbors were tested) (Finley, Datta et al. 2019). Models were fitted using the *spMsPGOcc* function for spatial multi-species occupancy models and *msPGOcc* for non-spatial models. Posterior distributions were sampled using the Polya-Gamma data augmentation approach in *spOccupancy* and Markov Chain Monte Carlo (MCMC). Bayesian inference was performed using Markov Chain Monte Carlo (MCMC) with 40,000 posterior samples, a burn-in of 5,000, thinning interval of 10, and three chains. Initial values and prior distributions were specified for all model parameters, and for full details of parameterization and prior settings, see Table S1 in the Supplementary Material. We fitted 123 candidate models, 110 of which were spatial and 10 non-spatial, and 3 additional non-spatial models that incorporated neighboring positive detections as covariates (all model specifications are provided in supplementary file Table S3-2). Models varied in spatial covariance functions, spatial approximation (full GP vs. NNGP) and neighborhood sizes. For each candidate model, backward variable elimination was performed by iteratively removing non-significant covariates while retaining predictors that remained statistically supported in the model. This enabled us to assess the influence of assumptions about spatial dependence and

ecological processes on model fit. We have selected the best model based on the K-fold deviance score, convergence and effective sample size (ESS) (Broms et al., 2016).

The best fitted model was used to produce posterior predictions of the occurrence of avian influenza for each HUC-8 watershed and season. In particular, posterior samples of occupancy probability (ψ) were drawn from the model and summarized as mean estimates and credible intervals for each site–season combination. To account for the actual cross-species transmission risk, these seasonal predictions were also adjusted for feral pig presence, with predicted probabilities weighted by the presence of feral pigs within a particular HUC-8. This allowed for areas without feral pigs to not be included in estimates of cross-species transmission risk. The outputs were spatially explicit, seasonally varying predictions of avian influenza occurrence, limited by both environmental suitability and host co-occurrence, and were used to create risk surfaces for analysis and visualization.

3.3 Results

3.3.1 Infection Prevalence Across Host Species

Infection prevalence was highly variable among bird species with at least 100 individuals tested (Figure 3-2). The highest prevalence was found in cinnamon teal (32.8%), followed by mallards (13.7%) and northern shovelers (11.1%). Intermediate prevalence was seen in blue-winged teal (7.4%) and green-winged teal (7.2%) while lower prevalence was noted in American wigeon (5.9%), northern pintail (5.9%) and American coot (5.4%). There was a slightly higher prevalence for females than males across sex classes (8.2% vs 7.3%).

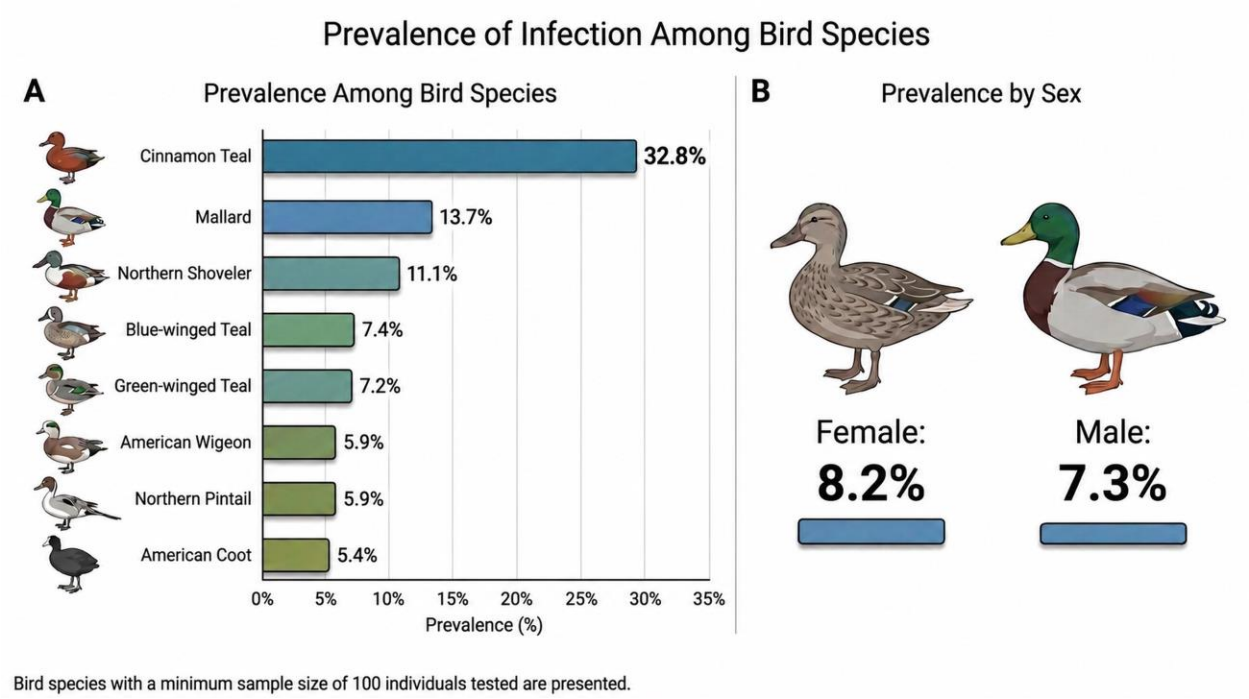


Figure 3-2: Infection prevalence among wild bird species and sex classes. (A) Species-specific prevalence estimates for bird species with a minimum sample size of 100 tested individuals. (B) Comparison of prevalence between female and male birds.

Prevalence estimates were similar among the demographic groups in wild pigs, although there was a slight variation by sex and age class (Figure 3-3). The prevalence of females (5.0%) was higher than that of males (4.5%). The prevalence was higher in adults (5.9%) than in juveniles (3.3%). Overall, prevalence patterns suggested moderate demographic variation with comparatively high prevalence of infection in adult and female wild pigs.

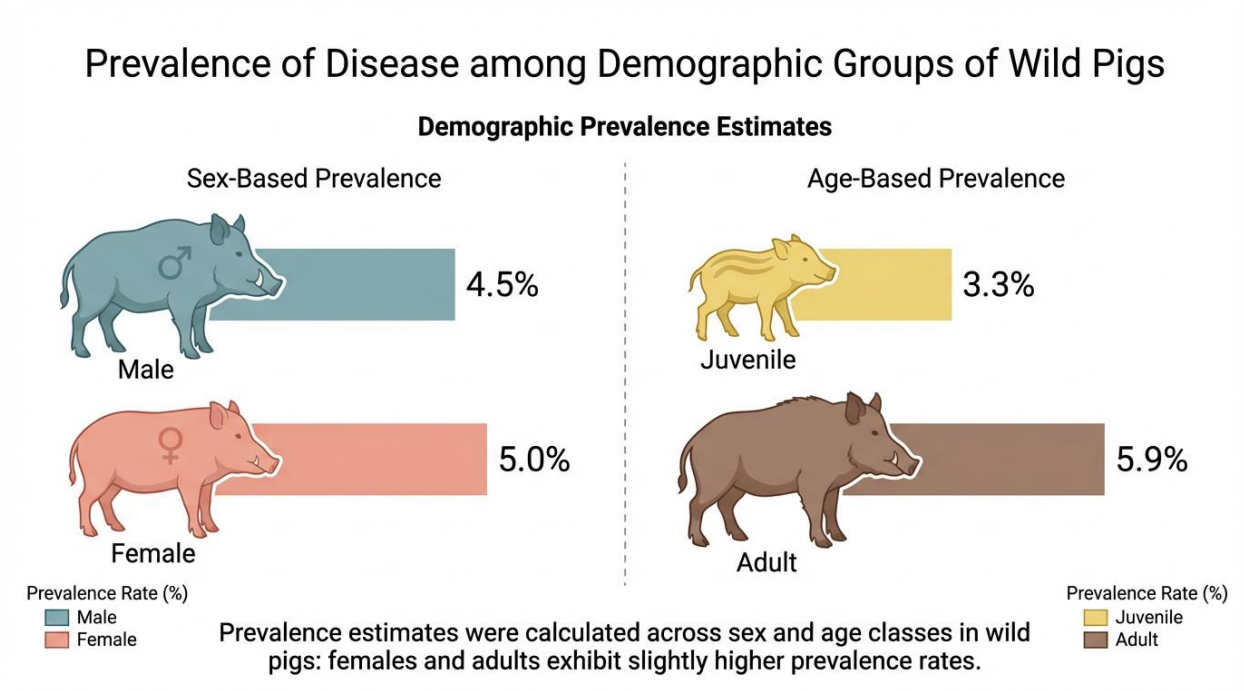


Figure 3-3: Infection prevalence among demographic groups of wild pigs. Sex-based prevalence estimates are shown for male and female pigs, while age-based prevalence estimates are presented for juvenile and adult pigs.

3.3.2 Model Results

The non-spatial model with log-transformed pig positive counts had the lowest K-fold cross-validation deviance (3301.32) of all the candidate models and thus performed best overall (Table 3-3). This model also had very good convergence diagnostics, with an average convergence value of 0.9999, and the ESS was high (ESS = 8118), which shows a good stability of the model and efficient posterior sampling. Several spatial formulations exhibited similar convergence and predictive ability, but none outperformed the non-spatial pig-count model in terms of cross-validation accuracy. The estimated occupancy probabilities were significantly higher than detection probabilities for both host groups (Figure 3-4). Pathogen occurrence was widespread in sampled HUC8–season units, with mean occupancy probability of 72.9% for wild birds and 68.7% for feral pigs. Detection probabilities were significantly lower, however, at 0.1% for wild birds and 10.4% for feral pigs, indicating that pathogens were often not detected in surveillance,

despite relatively high occupancy. The posterior latent occupancy states showed a weak positive association between wild bird and feral pig occurrence, with an estimated probability of positive co-occurrence of 59.2% and a probability of avoidance of 40.8%. Detection probability decreased with sampling effort, and then increased and peaked at about 8–10% at intermediate sampling intensities (~150–250 samples per site), and then slowly decreased.

Table 3-3: Performance comparison of the top-ranked non-spatial and spatial occupancy models across alternative covariance structures. Models were evaluated using average $\hat{R}$, effective sample size (ESS), and five-fold cross-validation deviance; results for all 123 candidate models are provided in Supplementary Table S2.

Model Structure	Model Formula	R-hat	ESS	K-fold CV (deviance score)
Non-spatial with log-pig positive count	log_neighbour_pig_cases + wetland_area_ha + ndvi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks + mean_gulls	0.9999	8118	3301.32
Non-spatial with log positive count (bird and pig)	log_neighbour_positive_cases +wetland_area_ha + ndvi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks + mean_gulls	1.0002	7967	3303.23
Non-spatial with log-bird positive count	log_neighbour_bird_cases + wetland_area_ha + ndvi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks + mean_gulls	1.0002	8033	3304.04
Non-spatial	wetland_area_ha + ndvi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks + mean_gulls	1.0006	7346	3307.27
Exponential - 15 Neighbors	wetland_area_ha + ndvi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks + mean_gulls	1.0011	7077	3318.69
Spherical - Pig HR	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks	1.0016	4368	3336.96
Spherical - Bird Local HR	wetland_area_ha + lswi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_geese_swans	1.002	7356	3357.11
Spherical - HUC	pct_wetland + edge_density + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_shorebirds	1.0005	8168	3406.61
Exponential - 10 Neighbors	pct_wetland + lswi + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + rel_dabbling_ducks + mean_diving_ducks + mean_geese_swans	1.0003	5745	3409.36
Exponential - 5 Neighbors	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + Out_Degree + Eigenvector + Bridging_Index + mean_geese_swans + mean_shorebirds	1.0007	7786	3409.98
Matern	pct_wetland + edge_density + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_shorebirds	1.0014	7231	3416.48
Gaussian - 10 Neighbors	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + In_Degree + Betweenness + mean_dabbling_ducks + mean_diving_ducks	1.0008	8073	3419.65

Gaussian - 15 Neighbors	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + In_Degree + Betweenness + mean_dabbling_ducks + mean_diving_ducks	1.0005	7965	3421.73
Gaussian - 5 Neighbors	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + In_Degree + Betweenness + mean_dabbling_ducks + mean_diving_ducks	1.0006	7865	3423.97
Gaussian - Full GP	pct_wetland + patch_density + area_mn + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + mean_diving_ducks + mean_gulls	1.0008	7868	3433.27

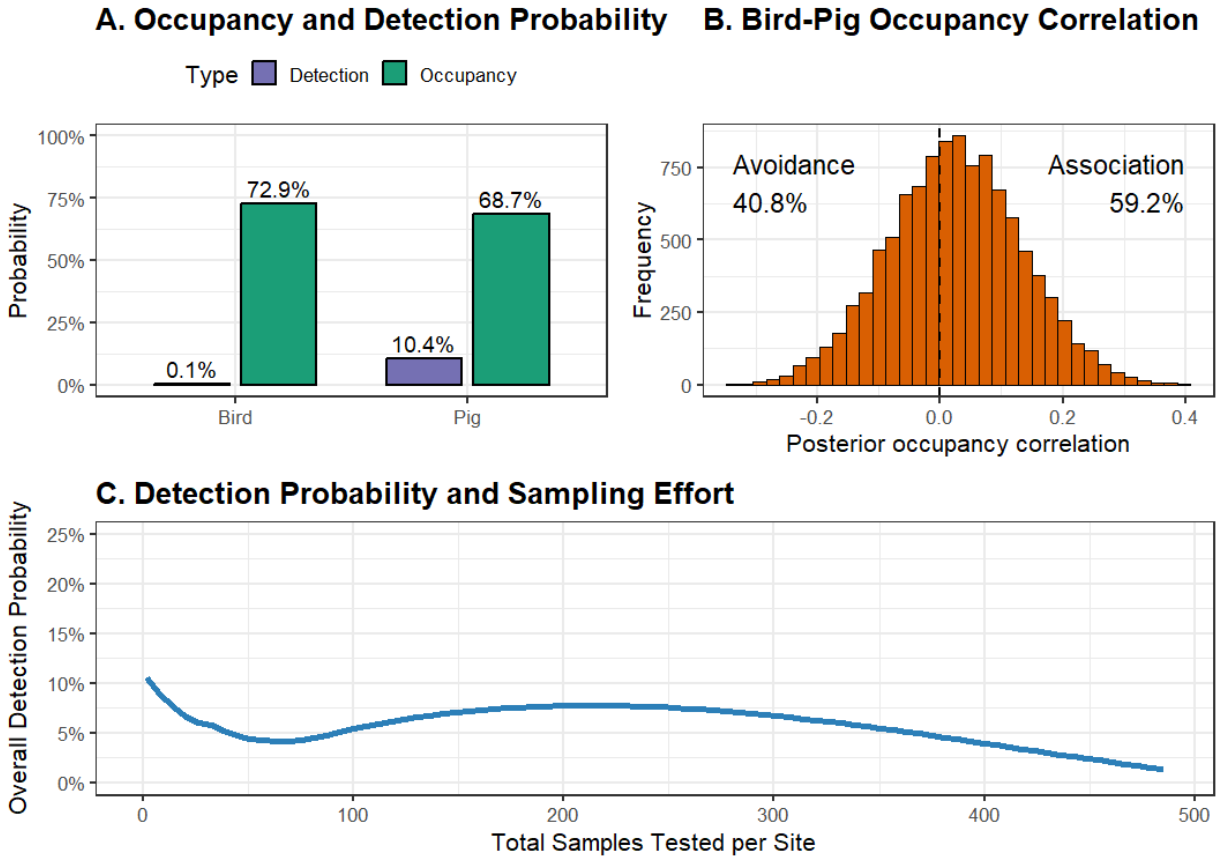


Figure 3-4: Occupancy, detection, and co-occurrence patterns of avian influenza in wild birds and feral pigs across HUC8 watershed–season sampling units. Posterior estimates from the Bayesian cross-species avian influenza occupancy model show host-specific occupancy and detection probabilities (A), the latent occupancy correlation between host groups (B), and the relationship between overall detection probability and sampling effort (C).

Several environmental and ecological factors were identified as important predictors of cross-species avian influenza occupancy from posterior estimates of the best performing model (Figure 3-5). Wetland area (ha), Vapor pressure, Soil moisture, Flow centrality, Mean Dabbling

Ducks Abundance, Mean Diving Ducks Abundance, Mean Gulls Abundance, and Log neighbor pig cases were associated with positive occupancy responses, although the functional forms of these relationships varied across predictors (Figure 3-6). Wetland area (ha) and Mean Dabbling Ducks Abundance had relatively high positive effects, with occupancy probability rising quickly from lower to intermediate values of the covariate before leveling off at higher values (Figure 3-6). Vapor pressure, Soil moisture, and Flow centrality showed moderate nonlinear positive responses with reducing marginal gains as the covariate levels increased (Figure 3-5). NDVI and Temperature range, on the other hand, had a negative relationship with occupancy probability, with NDVI exhibiting the steepest decrease over its observed range. The number of precipitation weeks showed a weak nonlinear response with relatively stable occupancy probabilities at lower numbers of weeks and a gradual decrease at higher numbers of precipitation weeks. The other host related covariates showed relatively strong positive effects and narrower uncertainty intervals than Mean Diving Ducks Abundance and Mean Gulls Abundance. The negative relationship between NDVI and occupancy probability was also strongest in Fall, Spring and Summer, while the interaction between NDVI and Winter had a weak positive effect size near zero on the logit scale.

3.3.3 Predictions

Cross-species avian influenza occupancy probabilities were highly heterogeneous in space and season throughout the contiguous United States (Figure 3-7). Occupancy probabilities were generally high in the southern part of the state, with the highest probabilities in Texas, the Gulf Coast, southern California, and parts of the lower Mississippi River Basin, and low in the northern Great Plains, upper Midwest, and northeastern United States. High-probability

watersheds were identified as being clustered in distinct areas along major hydrological corridors and in areas of overlap with known feral pig distributions.

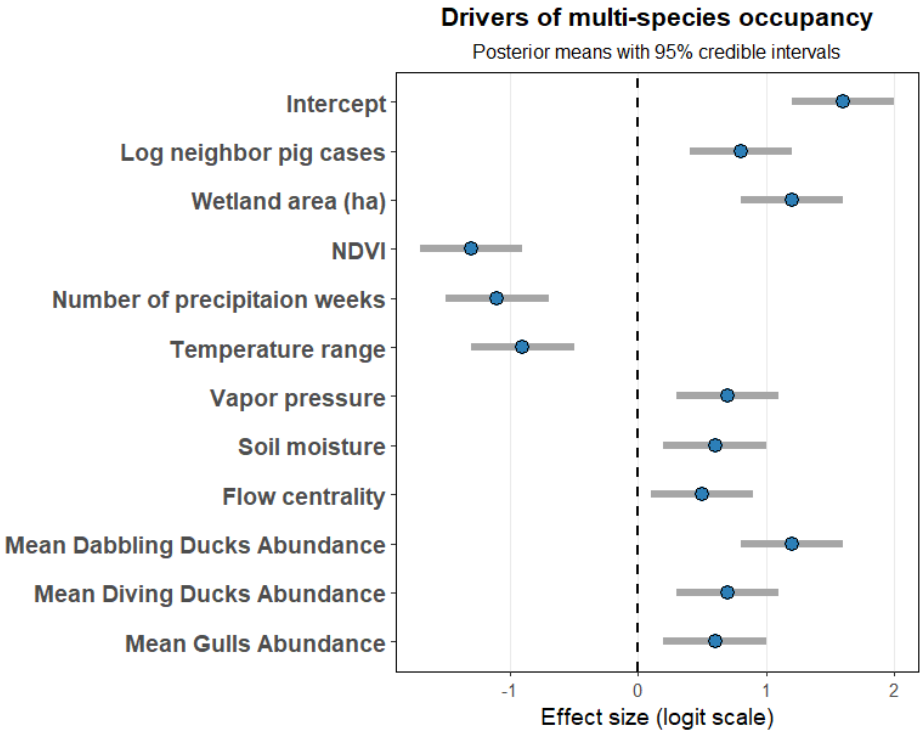


Figure 3-5: Posterior mean effect sizes and 95% credible intervals for environmental and ecological predictors of cross-species avian influenza occupancy from the best-performing model.

The seasonal predictions showed changing spatial patterns throughout the year. Fall moderate to high occupancy probabilities were found throughout Texas, the Gulf Coast, southern California and parts of the Mississippi River Basin. Moderate high occupancy areas were predicted to spread in parts of the Mississippi watershed systems in the central and southeastern Mississippi. For the summer, there were large contiguous high probability areas across Texas and surrounding southwestern watersheds and moderate occupancy areas were present throughout the southeastern United States. The area with high occupancy probabilities was the widest in winter, and the high-risk watersheds were concentrated in Texas, the Gulf Coast, southern California, and much of the Mississippi River Basin. The highest prediction uncertainty occurred across the

major hotspot areas in Texas, the Gulf Coast, southern California, and parts of the Mississippi River Basin, while areas with high uncertainty were found along the edges of these hotspot areas and in isolated watersheds outside of the main areas of predicted occupancy.

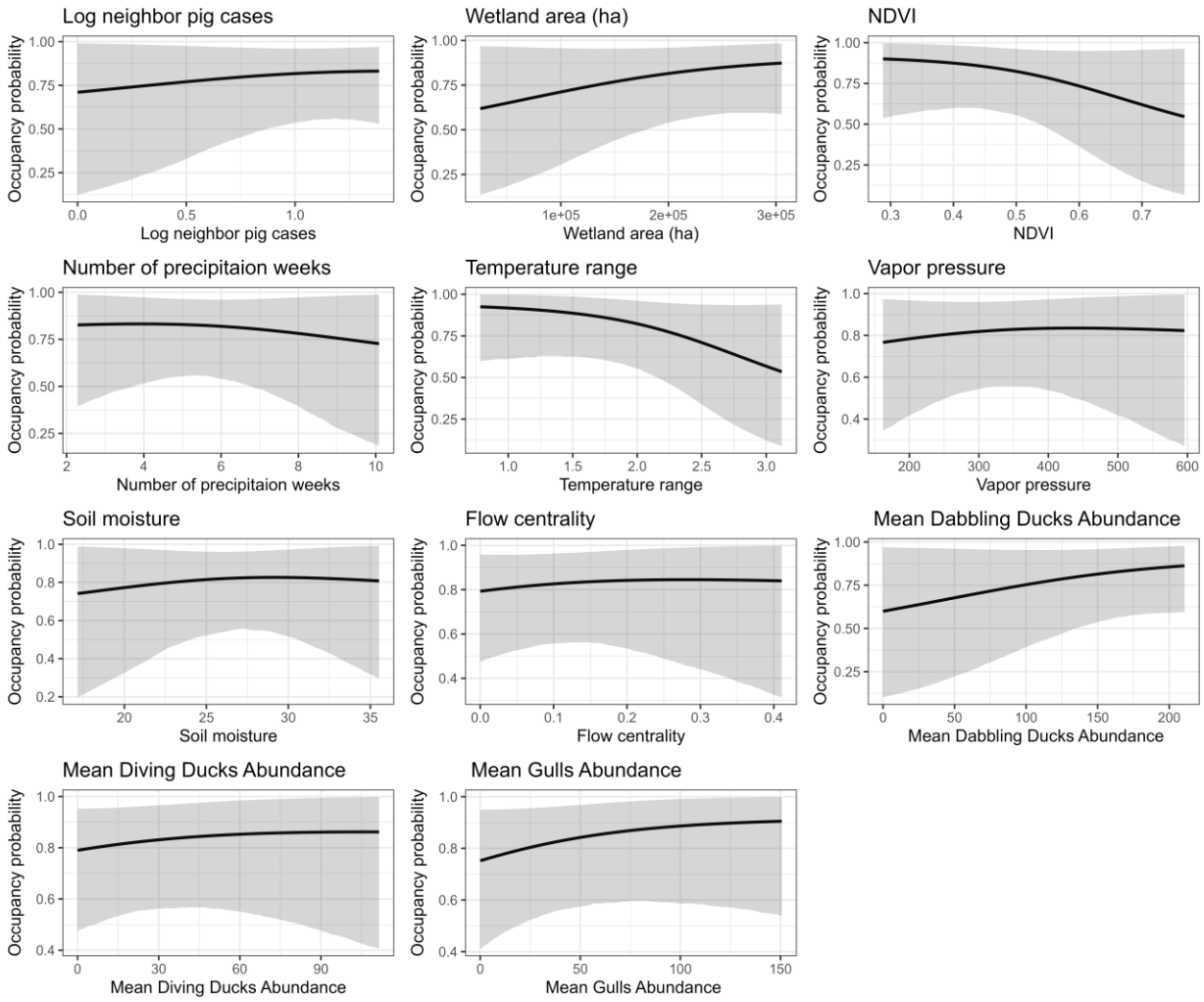


Figure 3-6: Marginal effect curves for covariates included in the final cross-species avian influenza occupancy model predicting avian influenza occupancy probability. Black lines represent posterior mean occupancy probabilities across the observed range of each covariate while holding all other covariates constant at their median values, and gray shaded regions represent the corresponding 95% credible intervals.

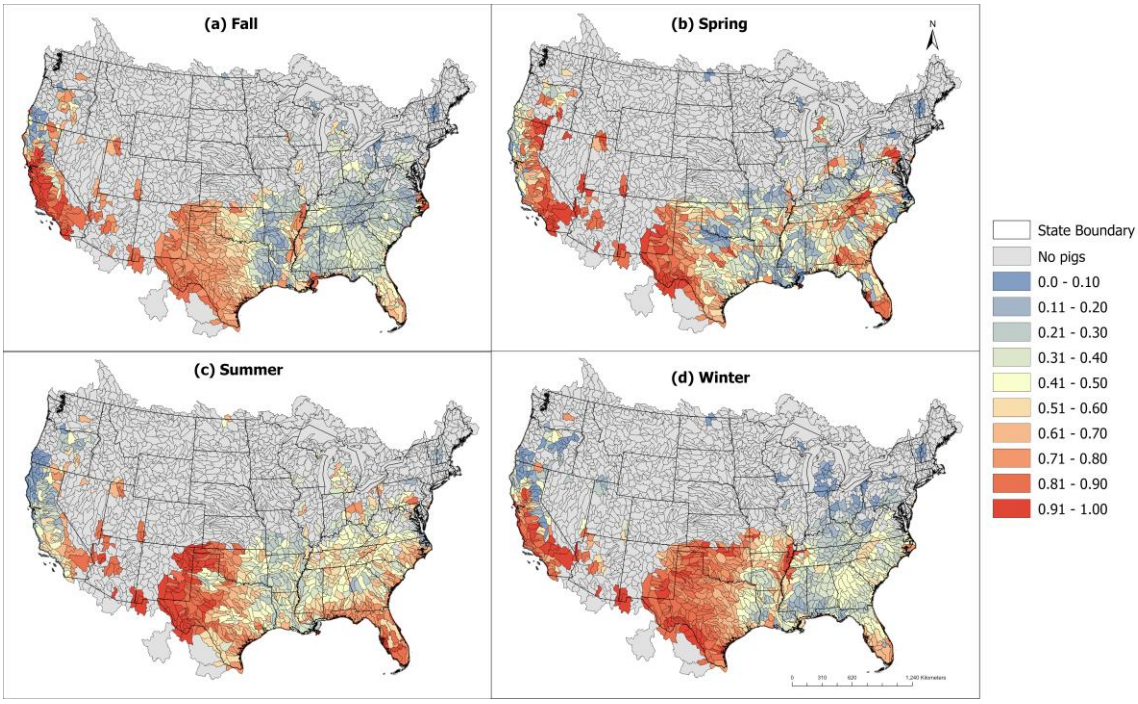


Figure 3-7: Seasonal posterior predictions of cross-species avian influenza occupancy probability across HUC8 watersheds in the contiguous United States for (a) Fall, (b) Spring, (c) Summer, and (d) Winter.

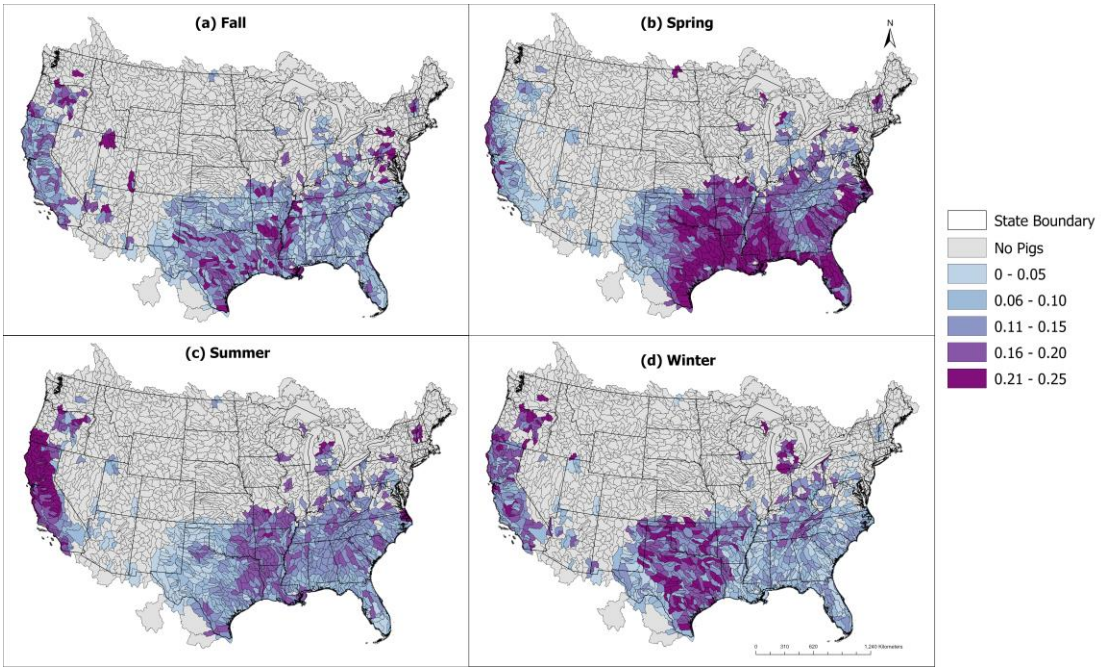


Figure 3-8: Seasonal uncertainty (posterior standard deviation) associated with predicted cross-species avian influenza occupancy probabilities across HUC8 watersheds in the contiguous United States for (a) Fall, (b) Spring, (c) Summer, and (d) Winter.

3.4 Discussions

Understanding the ecological processes that facilitate avian influenza transmission among wildlife hosts is essential for developing effective surveillance and early detection strategies, particularly as highly pathogenic avian influenza continues to expand across North America. This study quantified the probability of cross-species avian influenza transmission between wild birds and feral pigs across the contiguous United States at the watershed scale. Using a multi-species occupancy modeling framework that accounted for imperfect detection in surveillance data, we identified several environmental, ecological, and host-related drivers associated with transmission. Our findings demonstrate that avian influenza occupancy is strongly influenced by wetland structure, hydrological connectivity, climatic conditions, and the abundance of migratory water-associated bird species. These findings provide important insights for developing ecologically informed surveillance strategies that incorporate feral pigs as potential bridge hosts within integrated avian influenza monitoring frameworks.

3.4.1 Importance of Integrated Multi-host Surveillance

Our results suggest that occupancy probabilities were estimated at around 73% for wild birds and 69% for feral pigs, but detection probabilities were significantly lower, meaning that the presence of pathogens may often be missed in routine surveillance. From a surveillance perspective, the large discrepancy between estimated occupancy and detection probabilities indicates that non-detection should not be interpreted as pathogen absence. The positive association of detection probability with sampling effort is further evidence that more intensive surveillance increases the probability of detecting the presence of a pathogen, which will help to minimize the probability of false absences. The lack of an increase in detection probability at higher sampling levels, however, suggests diminishing returns at higher levels of effort and that

surveillance programs may be more efficient if they are able to optimize sampling intensity with a wider geographic distribution. A weak occupancy association between wild birds and feral pigs also indicates that a surveillance program targeting only one host species may not fully represent the occurrence of pathogens in the other host species, highlighting the importance of integrated multi-host surveillance approaches.

The incorporation of feral pig surveillance into existing avian influenza monitoring programs may further improve understanding of transmission dynamics because feral pigs provide information on environmental exposure at temporal scales that differ from migratory waterfowl. As resident hosts, feral pigs remain within the same watershed throughout the year and may therefore serve as indicators of sustained viral circulation within shared wetland systems after migratory birds have departed. Integrating avian influenza testing into ongoing USDA APHIS feral pig monitoring and removal programs could provide a cost-effective mechanism for identifying watersheds where bird-pig transmission is occurring and for evaluating the persistence of environmental transmission pathways. Such information could guide targeted environmental sampling, support risk assessments for nearby poultry operations, and help identify locations where management actions aimed at reducing contact between feral pigs and waterfowl may be most effective. Furthermore, uncertainty estimates generated by the model can inform adaptive surveillance by identifying watersheds where additional sampling is needed to better resolve potential bird–pig transmission pathways and improve future predictions of transmission risk (Miller et al., 2022).

3.4.2 Ecological and Environmental Drivers of Cross-species Transmission

The environmental drivers identified in this study suggest that cross-species avian influenza transmission between migratory waterfowl and feral pigs is governed by the interaction

of habitat availability, seasonal host dynamics, and hydrological processes. The positive association with wetland area highlights the importance of wetlands as ecological interfaces where migratory waterfowl and feral pigs repeatedly utilize shared resources. These habitats support large aggregations of water-associated birds while simultaneously providing drinking, foraging, and wallowing sites for feral pigs, thereby increasing opportunities for indirect exposure through shared environmental substrates (Elledge et al., 2013; Clontz et al., 2021; Yang et al., 2021). The negative relationship with NDVI was particularly notable and likely reflects seasonal vegetation phenology rather than vegetation structure itself. Lower NDVI values occur during autumn and winter, coinciding with periods of increased waterfowl abundance, migration, and avian influenza circulation throughout North America (Olsen et al., 2006; Verhagen et al., 2015). Consequently, NDVI may act as an indicator of broader seasonal ecological conditions that influence both host movement patterns and opportunities for transmission. In contrast, the positive relationships with soil moisture and vapor pressure and the negative association with temperature range suggest that moist and thermally stable environments may facilitate environmental maintenance of the virus between shedding and exposure events (Lowen et al., 2007; Zhang et al., 2014; Ferenczi et al., 2016).

Hydrological processes appear to further influence transmission by shaping how hosts and environmental reservoirs are connected across the landscape. The negative association with precipitation weeks suggests that excessive rainfall may reduce transmission opportunities by increasing water turnover, diluting viral particles, and dispersing hosts across a broader area, whereas moderate and stable hydrological conditions may promote repeated use of shared aquatic habitats (Si et al., 2010; García-Sastre, 2010; Ferenczi et al., 2021). This interpretation is supported by the positive relationship between flow centrality and occupancy probability,

indicating that connected wetland, riparian, and floodplain systems may facilitate redistribution of environmentally shed viral particles among watersheds (Gorsich et al., 2021; Pepin et al., 2023). Among host groups, dabbling ducks exhibited the strongest positive association with transmission probability, consistent with their recognized role as major avian influenza reservoirs and their tendency to forage in shallow aquatic habitats where fecal–oral transmission is highly efficient (Xu et al., 2016; Fourment et al., 2017). Although diving ducks and gulls showed weaker relationships, their extensive movements among wetlands, coastal habitats, and human-modified landscapes may contribute to pathogen redistribution otherwise disconnected ecosystems. Collectively, these findings suggest that cross-species transmission emerges from the combined effects of seasonal host aggregation, environmental suitability, and watershed-scale connectivity rather than any single ecological driver alone.

3.4.3 Spatial Distribution and Surveillance Implications

The implications for surveillance are significant as these findings indicate that existing bird surveillance programs may not be sufficient to detect environmentally mediated cross-species transmission in connected watershed systems. The broadest extent of elevated transmission probability occurred during winter and spring, extending across Texas, the Gulf Coast, southern California, and much of the Mississippi River Basin, whereas summer and fall predictions were more spatially restricted to portions of these same regions. These seasonal patterns are ecologically consistent with periods of increased waterfowl abundance and movement, coupled with environmental conditions that favor viral persistence in aquatic habitats (Olsen et al., 2006; Verhagen et al., 2015; Bevins et al., 2014). The concentration of elevated probabilities along major hydrological systems further supports the importance of environmentally mediated transmission through connected wetlands, riparian corridors, and

floodplain habitats (Gorsich et al., 2021; Ferenczi et al., 2016). High transmission probabilities predicted in portions of New Mexico and western Texas were particularly noteworthy given the comparatively lower wetland availability and feral pig densities in these regions. One possible explanation is that limited aquatic habitats within semi-arid landscapes may concentrate wildlife activity around relatively few water sources, increasing opportunities for environmental contamination and indirect transmission despite the overall scarcity of suitable habitat (Breban et al., 2009; Hénaux & Samuel, 2011). However, these regions also exhibited elevated prediction uncertainty, suggesting that additional surveillance is needed to determine whether these patterns reflect localized transmission hotspots or limited data availability. While regions with high transmission probabilities and low uncertainty can be prioritized as high-confidence surveillance targets, areas with elevated uncertainty may represent important knowledge gaps where additional sampling could improve understanding of potential transmission pathways and refine future surveillance efforts.

From a surveillance perspective, these findings indicate that monitoring efforts focused primarily on migratory bird sampling or poultry outbreak response may overlook environmentally mediated transmission occurring within connected watershed systems. The predicted hotspots suggest that surveillance should prioritize watersheds where suitable environmental conditions, hydrological connectivity, and overlap between migratory waterfowl and resident feral pigs create opportunities for repeated cross-species exposure. In particular, high-priority watersheds across Texas, the Gulf Coast, southern California, and the Mississippi River Basin may benefit from integrated watershed-based surveillance in which waterfowl, feral pigs, and environmental samples such as surface water and sediments are collected concurrently. Such an approach could help distinguish isolated spillover events from sustained environmental

circulation and improve detection of transmission occurring at the wildlife–livestock interface (Stallknecht & Brown, 2007; Roche et al., 2014; Pepin et al., 2023).

3.4.4 Limitations and Future Directions

However, there are several caveats that need to be acknowledged for our study. The use of surveillance data was from several programs with differing sampling intensity and detection effort, which may lead to some heterogeneity that could affect occupancy estimates even after accounting for imperfect detection explicitly. Furthermore, fine-scale variation within localized wetlands, where transmission dynamics are likely to be very heterogeneous, may be masked by watershed-scale analyses. The occupancy framework also does not directly measure transmission intensity, and so cannot be used to distinguish between persistent local maintenance and repeated seasonal introduction. Lastly, the ecological associations found here are very supportive of environmentally mediated transmission between birds and feral pigs, but not necessarily direct mechanistic evidence of viral exchange. Future research combining viral genomics, environmental persistence experiments and fine scale movement data of hosts will be critical to understanding the mechanistic contribution of feral pigs in the transmission system of avian influenza and whether the identified high-risk watersheds are environmental maintenance systems, repeated introduction hotspots, or a combination of both.

3.5 Conclusions

This study shows that environmentally mediated processes within hydrologically connected wetland systems are a significant driver of the transmission of avian influenza across migratory birds and feral pigs. Wetland extent, hydrological connectivity, environmental moisture conditions, and water-associated avian host assemblages were identified as drivers of

occupancy for both host groups, indicating that shared aquatic habitats serve as ecological transmission interfaces where viral persistence and repeated indirect exposure promote cross-species transmission. The results from the occupancy model suggests feral pigs may be a significant but under-recognized part of the avian influenza transmission system because of their year-round presence in wetlands, floodplains and riparian habitats that are also used seasonally by migratory waterfowl. The results also indicate the importance of adopting a strategy to monitor feral pigs in ecologically connected wetland systems that go beyond bird-based surveillance to include watershed connectivity and environmental transmission pathways. Overall, this study offers a basis for building multi-host surveillance frameworks that are risk-based and can enhance early detection and preparedness for avian influenza emergence at the wildlife-livestock-environment interface.

References

Capua, I., and D. J. Alexander. 2004. Avian influenza: recent developments. *Avian Pathology* **33**: 393–404.

Datta, A., Banerjee, S., Finley, A. O., & Gelfand, A. E. (2016). Hierarchical nearest-neighbor Gaussian process models for large geostatistical datasets. *Journal of the American Statistical Association*, *111*(514), 800-812.

Davis, A. J., Kirby, J. D., Chipman, R. B., Nelson, K. M., Xifara, T., Webb, C. T., Wallace, R., Gilbert, A. T., & Pepin, K. M. (2019). Not all surveillance data are created equal—a multi-method dynamic occupancy approach to determine rabies elimination from wildlife. *Journal of Applied Ecology*, *56*(11), 2551-2561.

Devarajan, K., Morelli, T. L., & Tenan, S. (2020). Multi-species occupancy models: Review, roadmap, and recommendations. *Ecography*, *43*(11), 1612-1624.

Doser, J. W., Finley, A. O., & Banerjee, S. (2023). Joint species distribution models with imperfect detection for high-dimensional spatial data. *Ecology*, *104*(9), e4137.

Doser, J. W., Finley, A. O., Kéry, M., & Zipkin, E. F. (2022). spOccupancy: An R package for single-species, multi-species, and integrated spatial occupancy models. *Methods in Ecology and Evolution*, *13*(8), 1670-1678.

Finley, A. O., Datta, A., Cook, B. D., Morton, D. C., Andersen, H. E., & Banerjee, S. (2019). Efficient algorithms for Bayesian nearest neighbor Gaussian processes. *Journal of Computational and Graphical Statistics*, *28*(2), 401-414.

Fuglstad, G.-A., Simpson, D., Lindgren, F., & Rue, H. (2019). Constructing priors that penalize the complexity of Gaussian random fields. *Journal of the American Statistical Association*, *114*(525), 445-452.

Hao, Z., Cai, X., Ge, Y., Foody, G., Li, X., Yin, Z., Du, Y., & Ling, F. (2024). Resolving data gaps in global surface water monthly records through a self-supervised deep learning strategy. *Journal of Hydrology*, *640*, 131673.

Lewis, J. S., Corn, J. L., Mayer, J. J., Jordan, T. R., Farnsworth, M. L., Burdett, C. L., VerCauteren, K. C., Sweeney, S. J., & Miller, R. S. (2019). Historical, current, and potential population size estimates of invasive wild pigs (Sus scrofa) in the United States. *Biological Invasions*, *21*(7), 2373-2384.

Lindgren, F., Rue, H., & Lindström, J. (2011). An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, *73*(4), 423-498.

Mosher, B. A., Brand, A. B., Wiewel, A. N., Miller, D. A., Gray, M. J., Miller, D. L., & Grant, E. H. C. (2019). Estimating occurrence, prevalence, and detection of amphibian pathogens: insights from occupancy models. *The Journal of Wildlife Diseases*, *55*(3), 563-575.

Miller, R. S., Bevins, S. N., Cook, G., Free, R., Pepin, K. M., Gidlewski, T., & Brown, V. R. (2022). Adaptive risk-based targeted surveillance for foreign animal diseases at the wildlife-livestock interface. Transboundary and Emerging Diseases, 69, e2329–e2340. https://doi.org/10.1111/tbed.14576

Poggio, L., De Sousa, L. M., Batjes, N. H., Heuvelink, G., Kempen, B., Ribeiro, E., & Rossiter, D. (2021). SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. *Soil*, *7*(1), 217-240.

Rota, C. T., Ferreira, M. A., Kays, R. W., Forrester, T. D., Kalies, E. L., McShea, W. J., Parsons, A. W., & Millspaugh, J. J. (2016). A multispecies occupancy model for two or more interacting species. *Methods in Ecology and Evolution*, *7*(10), 1164-1173.

Smith, G. J. (2013). *The US Geological Survey Bird Banding Laboratory: An integrated scientific program supporting research and conservation of North American birds* (2331-1258).

Sullivan, B. L., Aycrigg, J. L., Barry, J. H., Bonney, R. E., Bruns, N., Cooper, C. B., Damoulas, T., Dhondt, A. A., Dietterich, T., & Farnsworth, A. (2014). The eBird enterprise: An integrated approach to development and application of citizen science. *Biological conservation*, *169*, 31-40.

Thornton, M., Shrestha, R., Wei, Y., Thornton, P., Kao, S., & Wilson, B. (2022). Daymet: monthly climate summaries on a 1-km grid for North America, version 4 R1. ORNL DAAC, Oak Ridge, Tennessee, USA. In.

Tingley, M. W., Nadeau, C. P., & Sandor, M. E. (2020). Multi-species occupancy models as robust estimators of community richness. *Methods in Ecology and Evolution*, *11*(5), 633-642.

Domanska-Blicharz, Katarzyna, Zenon Minta, Krzysztof Smietanka, Sylvie Marché, and Thierry Van Den Berg. "H5N1 high pathogenicity avian influenza virus survival in different types of water." *Avian Diseases* 54, no. s1 (2010): 734-737.

U.S. Department of Agriculture. 2013. *National H5/H7 Avian Influenza Surveillance Plan*. Animal and Plant Health Inspection Service, Veterinary Services. 3–25. https://www.aphis.usda.gov/media/document/1295/file

U.S. Department of Agriculture. 2016. Final Report for the 2014–2015 Outbreak of Highly Pathogenic Avian Influenza (HPAI) in the United States. 10–11. https://www.aphis.usda.gov/animal_health/emergency_management/downloads/hpai/2015-hpai-final-report.pdf

U.S. Department of Agriculture. 2022. *Implementation Plan for the National Strategy for Highly Pathogenic Avian Influenza Surveillance in Wild Migratory Birds*. Animal and Plant Health Inspection Service.chrome- https://www.aphis.usda.gov/sites/default/files/2025-26-wild-bird-ai-surveillance-implementation-plan.pdf

Bevins SN, Pedersen K, Lutman MW, Baroch JA, Schmit BS, et al. (2014) Large-scale avian influenza surveillance in wild birds throughout the United States. PLoS ONE: e104360.

Jareb C, Pepin KM, Miller RS, Sykora S, Shwiff SA, McKee SC. Agricultural and Ecological Resources Safeguarded by the Prevention of Wild Pig Population Expansion. *Biology*. 2024; 13(9):670. https://doi.org/10.3390/biology13090670

Blagodatski, A., K. Trutneva, O. Glazova, O. Mityaeva, L. Shevkova, E. Kegeles, N. Onyanov, K. Fede, A. Maznina and E. Khavina (2021). "Avian influenza in wild birds and poultry: dissemination pathways, monitoring methods, and virus ecology." Pathogens **10**(5): 630.

Bourret, V. (2018). "Avian influenza viruses in pigs: An overview." The Veterinary Journal **239**: 7-14.

Breban, R., J. M. Drake, D. E. Stallknecht and P. Rohani (2009). "The role of environmental transmission in recurrent avian influenza epidemics." PLoS computational biology **5**(4): e1000346.

Brown, V. R., M. C. Marlow, R. M. Maison, T. Gidlewski, R. Bowen and A. Bosco-Lauth (2019). "Current status and future recommendations for feral swine disease surveillance in the United States." Journal of animal science **97**(6): 2279-2282.

Brüssow, H. (2024). "The arrival of highly pathogenic avian influenza viruses in North America, ensuing epizootics in poultry and dairy farms and difficulties in scientific naming." Microbial Biotechnology **17**(12): e70062.

Caliendo, V., N. S. Lewis, A. Pohlmann, S. R. Baillie, A. C. Banyard, M. Beer, I. H. Brown, R. Fouchier, R. D. Hansen and T. K. Lameris (2022). "Transatlantic spread of highly pathogenic avian influenza H5N1 by wild birds from Europe to North America in 2021." Scientific reports **12**(1): 11729.

Clontz, L. M., K. M. Pepin, K. C. VerCauteren and J. C. Beasley (2021). "Behavioral state resource selection in invasive wild pigs in the Southeastern United States." Scientific reports **11**(1): 6924.

Datta, A., S. Banerjee, A. O. Finley and A. E. Gelfand (2016). "Hierarchical nearest-neighbor Gaussian process models for large geostatistical datasets." Journal of the American Statistical Association **111**(514): 800-812.

Devarajan, K., T. L. Morelli and S. Tenan (2020). "Multi-species occupancy models: Review, roadmap, and recommendations." Ecography **43**(11): 1612-1624.

Domanska-Blicharz, K., Z. Minta, K. Smietanka, S. Marché and T. Van Den Berg (2010). "H5N1 high pathogenicity avian influenza virus survival in different types of water." Avian Diseases **54**(s1): 734-737.

Doser, J. W., A. O. Finley and S. Banerjee (2023). "Joint species distribution models with imperfect detection for high-dimensional spatial data." Ecology **104**(9): e4137.

Doser, J. W., A. O. Finley, M. Kéry and E. F. Zipkin (2022). "spOccupancy: An R package for single-species, multi-species, and integrated spatial occupancy models." Methods in Ecology and Evolution **13**(8): 1670-1678.

Elledge, A. E., C. A. McAlpine, P. J. Murray and I. J. Gordon (2013). "Modelling habitat preferences of feral pigs for rooting in lowland rainforest." Biological Invasions **15**(7): 1523-1535.

Ferenczi, M., C. Beckmann and M. Klaassen (2021). "Rainfall driven and wild-bird mediated avian influenza virus outbreaks in Australian poultry." BMC Veterinary Research **17**(1): 306.

Ferenczi, M., C. Beckmann, S. Warner, R. Loyn, K. O'riley, X. Wang and M. Klaassen (2016). "Avian influenza infection dynamics under variable climatic conditions, viral prevalence is rainfall driven in waterfowl from temperate, south-east Australia." Veterinary research **47**(1): 23.

Finley, A. O., A. Datta, B. D. Cook, D. C. Morton, H. E. Andersen and S. Banerjee (2019). "Efficient algorithms for Bayesian nearest neighbor Gaussian processes." Journal of Computational and Graphical Statistics **28**(2): 401-414.

Fourment, M., A. E. Darling and E. C. Holmes (2017). "The impact of migratory flyways on the spread of avian influenza virus in North America." BMC evolutionary biology **17**(1): 118.

Fuglstad, G.-A., D. Simpson, F. Lindgren and H. Rue (2019). "Constructing priors that penalize the complexity of Gaussian random fields." Journal of the American Statistical Association **114**(525): 445-452.

García-Sastre, A. (2010). "Influenza virus receptor specificity: disease and transmission." The American journal of pathology **176**(4): 1584-1585.

Gorsich, E. E., C. T. Webb, A. A. Merton, J. A. Hoeting, R. S. Miller, M. L. Farnsworth, S. R. Swafford, T. J. DeLiberto, K. Pedersen and A. B. Franklin (2021). "Continental-scale

dynamics of avian influenza in US waterfowl are driven by demography, migration, and temperature." Ecological Applications **31**(2): e2245.
Gutierrez, R. A. and P. Buchy (2012). "Contaminated soil and transmission of influenza virus (H5N1)." Emerging infectious diseases **18**(9): 1530.
Hao, Z., X. Cai, Y. Ge, G. Foody, X. Li, Z. Yin, Y. Du and F. Ling (2024). "Resolving data gaps in global surface water monthly records through a self-supervised deep learning strategy." Journal of Hydrology **640**: 131673.
Harvey, J. A., J. M. Mullinax, M. C. Runge and D. J. Prosser (2023). "The changing dynamics of highly pathogenic avian influenza H5N1: Next steps for management & science in North America." Biological Conservation **282**: 110041.
Hénaux, V. and M. D. Samuel (2011). "Avian influenza shedding patterns in waterfowl: implications for surveillance, environmental transmission, and disease spread." The Journal of Wildlife Diseases **47**(3): 566-578.
Hubbard, L. E., C. E. Givens, E. A. Stelzer, M. L. Killian, D. W. Kolpin, C. M. Szablewski and R. L. Poulson (2023). "Environmental surveillance and detection of infectious highly pathogenic avian influenza virus in Iowa wetlands." Environmental Science & Technology Letters **10**(12): 1181-1187.
Huchhegowda, K., H. Murugkar, K. Nagaraja, N. S. Manoj Kumar and C. Tosh (2014). "Persistence of highly pathogenic avian influenza virus (H5N1) in soil samples at different temperatures." Adv Anim Vet Sci **2**: 248-254.
Lewis, J. S., J. L. Corn, J. J. Mayer, T. R. Jordan, M. L. Farnsworth, C. L. Burdett, K. C. VerCauteren, S. J. Sweeney and R. S. Miller (2019). "Historical, current, and potential population size estimates of invasive wild pigs (Sus scrofa) in the United States." Biological Invasions **21**(7): 2373-2384.
Lindgren, F., H. Rue and J. Lindström (2011). "An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach." Journal of the Royal Statistical Society Series B: Statistical Methodology **73**(4): 423-498.
Lowen, A. C., S. Mubareka, J. Steel and P. Palese (2007). "Influenza virus transmission is dependent on relative humidity and temperature." PLoS pathogens **3**(10): e151.
Martin, B. E., H. Sun, M. Carrel, F. L. Cunningham, J. A. Baroch, K. C. Hanson-Dorr, S. G. Young, B. Schmit, J. M. Nolting and K.-J. Yoon (2017). "Feral swine in the United States have been exposed to both avian and swine influenza A viruses." Applied and environmental microbiology **83**(19): e01346-01317.
Mena, A., M. E. von Fricken and B. D. Anderson (2025). "The impact of Highly Pathogenic Avian Influenza H5N1 in the United States: A scoping review of past detections and present outbreaks." Viruses **17**(3): 307.
Nelli, R. K., S. V. Kuchipudi, G. A. White, B. B. Perez, S. P. Dunham and K.-C. Chang (2010). "Comparative distribution of human and avian type sialic acid influenza receptors in the pig." BMC veterinary research **6**(1): 4.
Pepin, K. M., C. B. Leach, N. L. Barrett, J. W. Ellis, K. K. VanDalen, C. T. Webb and S. A. Shriner (2023). "Environmental transmission of influenza A virus in mallards." MBio **14**(5): e00862-00823.
Poggio, L., L. M. De Sousa, N. H. Batjes, G. Heuvelink, B. Kempen, E. Ribeiro and D. Rossiter (2021). "SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty." Soil **7**(1): 217-240.

Ramey, A. M., N. J. Hill, T. J. DeLiberto, S. E. Gibbs, M. Camille Hopkins, A. S. Lang, R. L. Poulson, D. J. Prosser, J. M. Sleeman and D. E. Stallknecht (2022). "Highly pathogenic avian influenza is an emerging disease threat to wild birds in North America." The Journal of Wildlife Management **86**(2): e22171.

Renaud, C., A. Osborn, E. J. Parmley, T. F Hatchette, J. LeBlanc, J. S. Weese, V. Misra, D. Yamamura, S. Forgie and S. Renwick (2023). "Highly pathogenic avian influenza: Unprecedented outbreaks in Canadian wildlife and domestic poultry." Journal of the Association of Medical Microbiology and Infectious Disease Canada **8**(3): 187-191.

Runstadler, J. A. and W. B. Puryear (2024). "The virus is out of the barn: the emergence of HPAI as a pathogen of avian and mammalian wildlife around the globe." American Journal of Veterinary Research **85**(5).

Shahid, M. A., M. Abubakar, S. Hameed and S. Hassan (2009). "Avian influenza virus (H5N1); effects of physico-chemical factors on its survival." Virology Journal **6**(1): 38.

Si, Y., T. Wang, A. K. Skidmore, W. F. de Boer, L. Li and H. H. Prins (2010). "Environmental factors influencing the spread of the highly pathogenic avian influenza H5N1 virus in wild birds in Europe." Ecology and Society **15**(3).

Smith, G. J. (2013). The US Geological Survey Bird Banding Laboratory: An integrated scientific program supporting research and conservation of North American birds, US Geological Surey.

Sullivan, B. L., J. L. Aycrigg, J. H. Barry, R. E. Bonney, N. Bruns, C. B. Cooper, T. Damoulas, A. A. Dhondt, T. Dietterich and A. Farnsworth (2014). "The eBird enterprise: An integrated approach to development and application of citizen science." Biological conservation **169**: 31-40.

Thornton, M., R. Shrestha, Y. Wei, P. Thornton, S. Kao and B. Wilson (2022). Daymet: monthly climate summaries on a 1-km grid for North America, version 4 R1. ORNL DAAC, Oak Ridge, Tennessee, USA.

Ud-Dean, S. M. (2010). "Structural explanation for the effect of humidity on persistence of airborne virus: seasonality of influenza." Journal of Theoretical Biology **264**(3): 822-829.

Verhagen, J. H., R. A. Fouchier and N. Lewis (2021). "Highly pathogenic avian influenza viruses at the wild–domestic bird interface in Europe: future directions for research and surveillance." Viruses **13**(2): 212.

Wu, T., C. Perrings, C. Shang, J. P. Collins, P. Daszak, A. Kinzig and B. A. Minteer (2020). "Protection of wetlands as a strategy for reducing the spread of avian influenza from migratory waterfowl." Ambio **49**(4): 939-949.

Xu, Y., P. Gong, B. Wielstra and Y. Si (2016). "Southward autumn migration of waterfowl facilitates cross-continental transmission of the highly pathogenic avian influenza H5N1 virus." Scientific reports **6**(1): 30262.

Yang, A., R. K. Boughton, R. S. Miller, B. Wight, W. M. Anderson, J. C. Beasley, K. C. VerCauteren, K. M. Pepin and G. Wittemyer (2021). "Spatial variation in direct and indirect contact rates at the wildlife-livestock interface for informing disease management." Preventive Veterinary Medicine **194**: 105423.

Zhang, Z., D. Chen, Y. Chen, B. Wang, Y. Hu, J. Gao, L. Sun, R. Li and C. Xiong (2014). "Evaluating the impact of environmental temperature on global highly pathogenic avian influenza (HPAI) H5N1 outbreaks in domestic poultry." International journal of environmental research and public health **11**(6): 6388-6399.

Chapter 4 : Drivers of Human Leptospirosis Exposure in Different Community Types of Southern Chile: A One Health Perspective

Keywords: leptospirosis; communities; environmental drivers; animal reservoirs; One Health; extreme gradient boosting; eco-epidemiology

Published in MDPI *Pathogens* 2024, 13, 687. https://doi.org/10.3390/pathogens13080687

Abstract

Leptospirosis is a zoonosis with global public health impact, particularly in poor socioeconomic settings in tropical regions. Transmitted through urine-contaminated water or soil from rodents, dogs, and livestock, leptospirosis causes over a million clinical cases annually. Risk factors include outdoor activities, livestock production, and substandard housing that foster high densities of animal reservoirs. This One Health study in southern Chile examined Leptospira serological evidence of exposure in people from urban slums, semi-rural settings, and farm settings, using the Extreme Gradient Boosting algorithm to identify key influencing factors. In urban slums, age, shrub terrain, distance to Leptospira-positive households, and neighborhood housing density were contributing factors. Human exposure in semi-rural communities was linked to environmental factors (trees, shrubs, and lower vegetation terrain) and animal variables (Leptospira-positive dogs and rodents and proximity to Leptospira-positive households). On farms, dog counts, animal Leptospira prevalence, and proximity to Leptospira-contaminated water samples were significant drivers. The study underscores that disease dynamics vary across landscapes, with distinct drivers in each community setting. This case study demonstrates how the integration of machine learning with comprehensive cross-sectional epidemiological and geospatial data provides valuable insights into leptospirosis eco-epidemiology. These insights are crucial for informing targeted public health strategies and generating hypotheses for future research.

4.1 Introduction

Leptospirosis, a zoonotic disease of global distribution caused by the pathogenic bacterial species Leptospira, poses a significant health risk to both humans and animals (Crecelius & Burnett, 2020). In humans, leptospirosis can cause asymptomatic infection, flu-like illness or sometimes jaundice, kidney failure, meningitis, or even death (Haake & Levett, 2015). Globally, it has been estimated that leptospirosis causes around a million clinical cases and 58000 deaths each year but, due to poor diagnosis and reporting, the actual burden is unknown (Costa et al., 2015). The transmission of Leptospira is a multifaceted process that includes a diverse range of hosts and reservoirs and operates through both direct and indirect pathways. Humans may acquire infections either through direct contact with infected animals, both wildlife and domestic animals, or indirectly through exposure to contaminated environments (soil, water) (Luna et al., 2020). It is considered an occupational hazard related to activities such as agriculture, sewage management, and animal husbandry (Guerra, 2013), as well as an infection associated with routine domestic or recreational activities that put individuals in contact with a Leptospira-contaminated environment. Rodents (rats, mice) serve as primary synanthropic reservoirs, although other wildlife species such as raccoons, skunks, opossums, foxes, and deer can also become infected (Ellis, 2015). Dogs transmit the pathogen to humans through either direct contact or urine-infected materials or environment (Bradley & Lockaby, 2023). Domesticated farm animals like cattle and sheep can also act as a carrier of Leptospira and people can become infected by direct contact with blood, aborted fetuses, vaginal discharge, or calving products from infected animals as well as indirectly through the urine-contaminated farm environment (Montes & Monti, 2021). Human-to-human infection is also possible via breast milk or sexual contact but is very rare (Bolin & Koellner, 1988; Harrison & Fitzgerald, 1988). More

importantly, people can become infected when exposed to water and soil that is contaminated with Leptospira (Claudia Muñoz-Zanzi et al., 2014). Communities with low socio-economic conditions are vulnerable to exposure given the inadequate sanitation and lack of safe drinking water (Pappas et al., 2008). Tropical climate areas where there is frequent flooding caused by high rainfall and natural disasters often result in high risk of Leptospira infection (Miller et al., 1991; Romero et al., 2003). Moreover, humidity and warm temperatures in tropical areas cause the pathogen to persist longer in the environment, specifically in soil, which in turn, increases the likelihood of exposure (Mwachui et al., 2015). The multifaceted nature of leptospirosis is driven by a variety of factors, encompassing ecological, animal, and anthropogenic elements. Geographical differences further contribute to the dynamic nature of the transmission, indicating that the intricate web of factors influencing the disease may differ significantly from one community to another. Understanding the interactions among these elements is crucial for developing targeted and effective strategies to mitigate the impact of leptospirosis in different regions. However, the existing literature is scarce about how leptospirosis can be present in different community settings and their unique relationships between the various transmission drivers.

The growing body of research on leptospirosis prevalence highlights the lack of a comprehensive investigation framework. While numerous studies have explored the prevalence of leptospirosis within different populations, a noticeable limitation arises from the absence of simultaneous consideration of human, animal, and environmental factors. The One Health approach advocates for an integrated and interdisciplinary perspective that acknowledges the interconnectedness of human, animal, and environmental health. Utilizing the principles of One Health in leptospirosis research facilitates a more holistic comprehension of disease dynamics

and transmission pathways. By scrutinizing the complex interactions among humans, animals, and their shared environment through adopting a One Health framework, researchers can gain insight into the intricate epidemiology of leptospirosis and devise more efficacious prevention and control measures.

Thus, the data presented herein constitute a case study integrating leptospirosis data from rodents, dogs, and the environment (all previously published), along with new livestock and human seroprevalence data from an eco-epidemiological study conducted in Chile (C. Muñoz-Zanzi et al., 2014). The Los Rios region in south-central Chile, a predominantly agricultural and farming area juxtaposed with scattered urban settlements, has been a location of several leptospirosis studies (Alexander, 1960). A previous serosurvey targeting individuals with occupational hazards revealed a seroprevalence of around 22%, underscoring the need to elucidate disease dynamics and effective control measures (Zamora et al., 1990). Additionally, a survey conducted in dogs in the area reported a leptospirosis seroprevalence of 25%, with variations across different community types, highlighting the interconnectedness of human-influenced environments and animal health (Lelu et al., 2015). Rodents in the region were shown to have kidney carriage of around 20%, further emphasizing the multifaceted nature of leptospirosis transmission pathways in the Chilean site (C. Muñoz-Zanzi et al., 2014).

Using a One Health framework, the objective of this study was to identify the intricate relationships between human Leptospira exposure and different drivers, including household characteristics, animal reservoirs, environmental conditions, and water sources, in three distinct community settings in the Los Rios region of Chile: urban slums, semi-rural and farm communities. We hypothesized that the drivers of human Leptospira exposure would be different in those three communities given different environmental settings and compositions of animal

reservoirs. Specifically, we expected that rodent-related variables and household conditions would drive the human Leptospira exposure in urban slum communities, while the exposure in semi-rural areas will be impacted by environmental conditions and household variables. In farm communities, we expect that livestock and wildlife-related variables would play an important role in driving human Leptospira exposure.

4.2 Materials and Methods

4.2.1 Study area

The study design and data come from a study on the eco-epidemiology of leptospirosis conducted in the Los Rios region in south-central Chile (Latitude: 39°15'S - 40°33'S, Longitude: 73°43'W-71°35'W) (Claudia Muñoz-Zanzi et al., 2014; C. Muñoz-Zanzi et al., 2014). The climate is characterized as temperate rainforest, with annual cumulative rainfall if 2588 mm and ranges from 1200 mm in the central valley up to 5000 mm in the Andes Mountain. Average temperature has less variation throughout the year, with 17°C in summer and 8°C in winter (C. Muñoz-Zanzi et al., 2014). Communities were selected to represent three community types (i) urban slums: informal settlements in the outskirts of a major city characterized by substandard housing, (ii) semi-rural: rural community settlements away from major cities where households are clustered together and (iii) farm: dispersed households, typically small family farms, located in a specific rural locality. Communities were chosen from areas that have the highest number of settlements in the region, specifically the central valley and the vicinity of the region's capital. Most of the communities were located within the elevation of 0-100 m above sea level, except for two households in farm communities which were at 100-200 m elevation above sea level (Mason et al., 2016). Individual households within each community were selected randomly and

enrolled based on willingness to participate in all components of the study, which included rodent trapping, sampling of domestic animals, sampling of surface water sources in the peri-domestic environment, and sampling of household members.

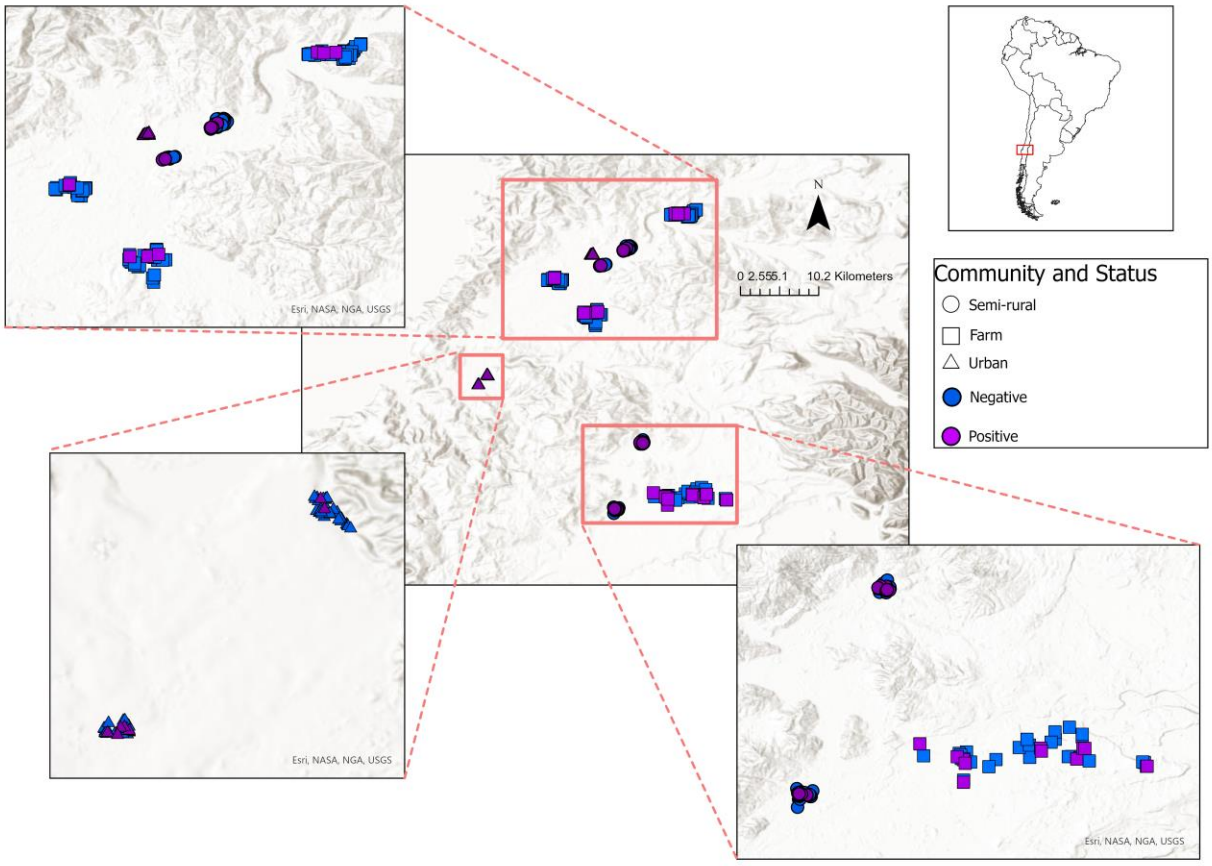


Figure 4-1: Study area and distribution of leptospirosis seropositive (purple) and seronegative (blue) participants across the various community types indicated by shape (semi-rural: circle, farm: square, urban slums: triangle).

4.2.2 Data collection

The data collection for the eco-epidemiology study was conducted between August 2010 and March 2012 and involved 422 households from 12 communities, 4 of each community type. A

questionnaire survey was carried out to obtain information on socio-demographic characteristics, housing conditions, presence of animals, and to characterize each sample.

Leptospirosis status in people and domestic animals- Data available included *Leptospira* exposure status of 907 people. Household members 13 years old and older who consented to participate provided a one-time serum sample to measure *Leptospira*-specific antibodies. Dogs and livestock present at the residence were also sampled to obtain serum for *Leptospira* antibody testing. The microscopic agglutination test (MAT) with a panel of 20 serovars, representing 17 serogroups was run on serum samples at Austral University, Chile *(L. interrogans* serovars Australis, Bratislava, Autumnalis, Bataviae, Canicola, Djasiman, Grippotyphosa, Icterohaemorrhagiae, Mankarso, Pomona, Pyrogenes, and Wolffi, *L. borgpetersenii* serovars Ballum, Javanica, and Tarassovi, *L. kirschneri* serovar Cynopteri, *L. santarosai* serovars Borincana, Alexi, and Georgia, and *L. weilii* serovar Celledoni) as described previously (Lelu et al., 2015). Titers of 1:100 or higher were considered positive to classify each individual (people and animals) as seropositive or seronegative for *Leptospira*. We calculated the cumulative number of seropositive animals for cattle (*bov_com_pos*), sheep (*ovi_com_pos*), dogs (*dog_com_pos*) and all animals (*animal_pos_com*) in each community.

Leptospirosis status in rodents- Trapped rodents were euthanized and kidneys were tested for presence of *Leptospira* DNA using Polymerase Chain Reaction (PCR) as previously described (C. Muñoz-Zanzi et al., 2014). Test results were used to create a variable representing the number of households within 100-meter with positive rodents weighted inversely by the distance from the house (*distance_pos_rod*).

***Leptospira* contamination in the peridomestic environment-** Surface water samples were collected from various sources and locations in the household environment and tested for the

presence of pathogenic *Leptospira* DNA using PCR as previously described (Mason et al., 2016; Claudia Muñoz-Zanzi et al., 2014). Results were then used to calculate variables that represent the proportion of positive puddle samples in the community (*puddle_pos*), the proportion of positive samples considering all water source types (puddles, rivers, wells, ponds, etc.) in the community (*water_prev_com*), and the number of households within 100-meter with positive water samples weighted inversely by distance from the house (*distance_pos_water*).

Participant characteristics- A questionnaire was used to collect individual socio-demographic characteristics (*age, sex*) and behaviors that could expose people to *Leptospira* such as gardening (*garden*), swimming (*swim*), cleaning animal barns (*clean_barn*), cleaning sewers (*clean_drain*), water drainage (*clean_water_drain*), slaughter animals (*slaughter*), milking (*milking*), and/or cleaning after animal birth (*clean_birth*).

Household animal characteristics- Rodent presence was obtained based on trapping efforts as the number of rodents trapped at the household (*rodent_count*) and the total for each community (*rodent_count_comm*), as previously described (C. Muñoz-Zanzi et al., 2014). A questionnaire was used to collect data on reported presence of rodents (e.g., observing rodents, seeing rodent droppings, seeing, or smelling rodent urine, whether gnawed boxes, food, or wood, holes in the walls were found, or hearing rodent noises) (C. Muñoz-Zanzi et al., 2014), as well as domestic animal-related variables such as the total number of cattle, sheep, and dogs in the household.

Spatial and environmental variables- Geospatial variables from multiple sources were used to describe the environmental settings of each community. Details on how variables were generated have been described previously (Munoz-Zanzi et al., 2016; C. Muñoz-Zanzi et al., 2014). Variables used for this analysis included the number of houses within a 100-meter radius (*house*) and the number of buildings within a 100-meter radius (*buildings*). Land cover was defined as

the dominated terrain within the surrounding radius and categorized as tree canopy (*tree*), lower vegetation (*lowveg*), shrub (*shrub*), or barren space (*field*). Bioclimatic variables used in this study to represent the climatic conditions of the study area (*bio1, bio2, bio12, bio15*) were collected from WorldClim (worldclim.org) (Fick & Hijmans, 2017).

Table 4-1: Description and source of the variables included in the study.

Type	Variable Name	Description	Source
Socio-demographic and household characteristics	*sex*	Sex of the person	Questionnaire
	age	Age of the person (in years)	
	clean_barn	Person cleans barns	
	clean_drain	Person cleans drains in the field	
	slaughter	Person butchers meat	
	milking	Person milks cows	
	clean_birth	Person cleans cow birth products	
	clean_water_drain	Person cleans water drains	
	clean_field	Person cleans fields	
	swim	Person swims	
	season	Sampling season	
	house	Number of houses within 100-meter radius	Derived from worldview-2 satellite imagery
	buildings	Number buildings within 100-meter radius	
Environmental	*elev*	Altitude of sampled household	Derived from worldview-2 satellite imagery
	FlowAcc	Difference in altitude compared to surroundings (higher numbers means greater slope downward)	
	tree	Square meters of tree-dominated terrain within 100-meter radius	
	lowveg	Square meters of lower-vegetation terrain within 100-meter radius (e.g. bushes and other short plants)	
	shrub	Square meters of shrub-dominated terrain within 100-meter radius	
	wetland	Square meters of wetland terrain within 100-meter radius	
	field	Square meters of field terrain within 250-meter radius	
	bio1	Annual mean temperature	worldclim.org
	bio2	Mean Diurnal Range (Mean of monthly (max temp – min temp))	
	bio12	Annual Precipitation	
	bio15	Precipitation Seasonality (Coefficient of Variation)	
	puddle_pos_com	Proportion of *Leptospira* positive puddles in the community	Laboratory testing
	water_prev_com	Proportion of *Leptospira* positive water samples in the community (all water source types)	

	distance_pos_water	Number of households within 100 meters with *Leptospira*-positive water samples weighted inversely by distance from house	Derived from worldview-2 satellite imagery
Animal	*rodent_count*	Number of rodents trapped in the household	Questionnaire
	rod_pos	Presence of positive rodents in the household	Derived
	rodent_count_com	Number of rodents trapped in the community	Questionnaire
	RodHHPrev	*Leptospira* prevalence in rodents at household level	Derived
	rodent_prev_com	*Leptospira* prevalence in rodents in the community	Derived
	distance_pos_rod	Number of households within 100-meter with *Leptospira*-positive rodents weighted inversely by distance from house	Derived from worldview-2 satellite imagery
	spdiv	Number of different domestic animal species in the household	Derived
	bov_count	Number of bovines in the household	Questionnaire
	bov_pos	Presence of seropositive bovines in the household	Derived
	BovHHPrev	*Leptospira* seroprevalence in bovines at household level	
	bov_com_pos	Number of seropositive bovines in the community	
	bov_com_prev	*Leptospira* seroprevalence in bovines at community level	
	ovi_count	Number of ovines in the household	Questionnaire
	ovi_pos	Presence of seropositive ovines in the household	Derived
	OviHHPrev	*Leptospira* seroprevalence in ovines at household level	
	ovi_pos_com	Number of seropositive ovines in the community	
	OviComPrev	*Leptospira* seroprevalence in ovines at community level	
	dog_count	Number of dogs in the household	Questionnaire
	dog_pos	Presence of seropositive dogs in the household	Derived
	DogHHPrev	*Leptospira* seroprevalence in dogs at household level	
	dog_com_pos	Number of seropositive dogs in the community	
	DogComPrev	*Leptospira* seroprevalence in dogs at community level	
	Anim_pos	Presence of seropositive animals in the household	
	AnimalHHPrev	*Leptospira* seroprevalence in farm animals at household level	
	animal_pos_com	Number of overall seropositive farm animals in the community	
	AnimCommPrev	*Leptospira* seroprevalence in farm animals at community level	

All human participants provided written consent to participate in this study, including the disclosure of their domestic and livestock information. The study protocol was approved by the University of Minnesota's Institutional Review Board (No. 0903 M62042) and Institutional

Animal Care and Use Committee (No. 0904A63201) and the Austral University's Human and Animal Ethics Committee (No. 01/09). More details about data collection can be found in (Munoz-Zanzi et al., 2016; C. Muñoz-Zanzi et al., 2014). See Table 1 for further information on data descriptions and sources.

4.2.3 Extreme Gradient Boosting Model

XGBoost is a high-performing gradient classification and regression boosting machine learning algorithm that is widely used in disease prediction, identifying medical conditions, image classification and object detection (Chen & Guestrin, 2016; Ogunleye & Wang, 2020; Shaheed et al., 2023). This algorithm uses split techniques, which select the features that best separate the data into two groups. This method is well suited for our analysis because it identifies complex and non-linear relationships among the drivers and *Leptospira* exposure and adjusts the collinearity between the drivers (Ergul Aydin & Kamisli Ozturk, 2021). It also minimizes overfitting and achieves high predictive accuracy through regularization techniques such as pruning and shrinkage, and by ensemble learning with decision trees (Chen & Guestrin, 2016).

Table 4-2: Parameters that were used to tune the models.

Parameter	Description	Range	Interval
scale_pos_weight	Weight of positive class to address class imbalance	Neg/pos	Fixed
nrounds	Number of boosting rounds or iterations during the training process.	100-600	50
learning_rate	Learning rate for gradient boosting	0-1	0.01
max_depth	Maximum depth of the decision tree	0-10	1
min_child_weight	Minimum sum of instance weight (Hessian) needed in a child	0-10	1
gamma	Minimum loss reduction required to make a further partition on a leaf node	0-5	0.1
subsample	Fraction of training data to randomly sample during training	0-1	0.1

colsample_bytree	Fraction of features to be randomly sampled for each tree	0-1	0.1
objective	Learning task and objective function (binary classification in this case)	Binary:logistic	
Max_delta_step	Introduce an 'absolute' regularization capping the weight before apply eta correction.	1-10	0.1

We run separate XGBoost models for the three different community types. For each model fitting, we randomly split the observations into training and testing sets under the ratio of 80% and 20%. The training dataset was used to run the models and identify the drivers, while the testing dataset was used to estimate the accuracy of the model.

The performance of the XGBoost can be sensitive to their hyperparameters (Putatunda & Rama, 2018). After splitting, we tuned a variety of hyperparameters to optimize the performance of the model. A matrix of hyperparameters was provided by using the *expand.grid* function to find the best combination of the hyperparameters (Davagdorj et al., 2020). Specifically, the maximum tree depth, learning rate, and gamma were adjusted here, given that these variables generally exhibit the most significant impact on model performance (Srinivas & Katarya, 2022). Furthermore, we accounted for unbalanced classes, which refers to an imbalance between the number of positive and negative cases, by adjusting the model parameter *'scale_pos_weight'* (Farooq et al., 2022). This is estimated as the total number of negative cases divided by the total number of positive cases (See Table 2 for detailed information about the parameterization). The final XGBoost classification model was fitted using those tuned hyperparameters. To account for stochasticity in random split of training and testing set in model development, we performed 2000 iterations of our data splits and perform the final splitting processes for them. This allowed us to use an ensemble modeling approach that incorporated information and uncertainties from

multiple random split scenarios and created confidence intervals for evaluating the model and determining the importance of the drivers.

We also generated the variable importance (gain score'- represents how much a variable contributes to enhancing the model's predictions) on for every model iteration and took the average to show the most important drivers of leptospirosis in three community types and the partial dependency plots to show the relationship between potential drivers and *Leptospira* exposure. To evaluate the performance of the final model, we used the area under the receiver operating characteristics curve (AUC) value, which assesses how well the model can distinguish positive and negative cases (Jin & Ling, 2005). The XGBoost model development and the associated analyses were performed using the "*XGBoost*" package in R version 4.3.1 (Chen T, 2023; Team, 2020).

4.3 Results

4.3.1 Seroprevalence of Leptospirosis

Evidence of Leptospira exposure was evident in all community types with an overall seroprevalence in people of 6.0%. Seroprevalence was also 6% for each community type, while ranging by community from 3.7% to 10.3%. Leptospira occurrence is high with 20.4% of rodent kidney carriers and 13.5% of water samples classified as PCR positive. Among dogs owned by

household members, 26.8% were seropositive with a marked increase in urban dogs (45.1%)

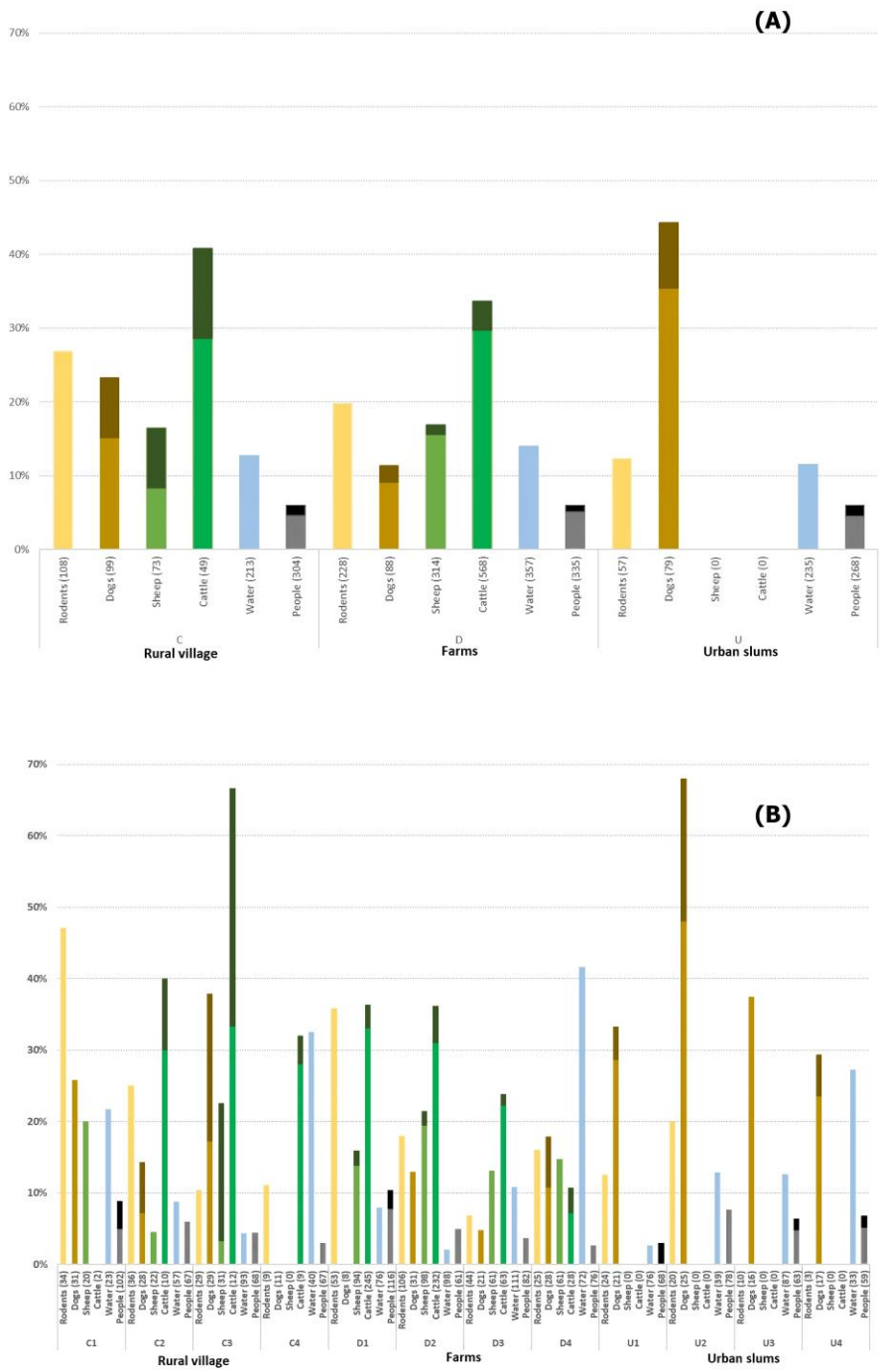


Figure 4-2: Leptospira positivity distribution in rodents (PCR kidney carriage positivity), most commonly sampled domestic animals (MAT seropositivity), water samples from the peri-domestic environment (PCR positivity for pathogenic species), and people (MAT seropositivity) (A) by the three community types and and (B) by the 12 communities, within each community type, Los Rios region, Chile Darker color in a bar represents the proportion of individuals with a MAT titer ≥1:400.

compared with rural dogs (22.3%). Sheep and cattle were the main livestock in the area with overall seroprevalences of 16.5% and 31.2%, respectively (Figure 4-2).

Overall, our study revealed varying seroprevalence in people across different demographic and behavioral categories. Males exhibited a higher seroprevalence (6.5%) than females (5.6%), while individuals engaging in swimming activities showed a higher seroprevalence (6.8%) compared to non-swimmers (4.0%). People living in households harboring positive rodents and dogs had greater seroprevalences (7.1% and 7.8%, respectively) than those without such animals (5.0% and 5.6%, respectively). Additionally, gardening and barn cleaning activities were associated with increased seroprevalence (6.7% and 6.8%, respectively) compared to individuals who did not perform these tasks (4.1% and 5.4%, respectively missing % for those). People who reported activities involving slaughtering and milking animals displayed higher seroprevalence (7.3% and 8.3%, respectively) than those who did not (5.7% and 5.8%, respectively % for those) (Table 4-3). Notably, the lowest seroprevalence was observed among participants involved in sewage drain cleaning activities (2.3%).

Table 4-3: Seroprevalence of Leptospirosis in different demographic and behavioral factors.

Variables	No of Participants	MAT positive	seroprevalence (%)	95% CI
Sex of the person				
Male	387	25	6.46	4.22-9.39
Female	520	29	5.57	3.77-7.91
Person swim's				
Yes	629	43	6.84	4.99-9.09
No	278	11	3.96	1.99-6.97
Positive rodents in the household				
Yes	407	29	7.13	4.82-10.07
No	500	25	5.00	3.26-7.29
Positive dog in the household				
Yes	128	10	7.81	4.13-7.51

No	779	44	5.64	3.81-13.90
Positive Cattle in the household				
Yes	299	16	5.35	3.09-8.54
No	608	38	6.25	4.46-8.48
Positive sheep in the household				
Yes	282	13	4.61	2.48-7.75
No	625	41	6.56	4.75-8.79
work in garden				
Yes	269	11	4.09	2.06-7.20
No	638	43	6.73	4.92-8.97
Clean barn				
Yes	337	23	6.82	4.37-10.06
No	570	31	5.44	3.73-7.63
Clean sewage drains				
Yes	46	1	2.27	0.05-11.53
No	861	53	6.15	4.64-7.97
Person slaughter				
Yes	123	9	7.31	3.40-13.43
No	784	45	5.74	4.22-7.60
Person milk animals				
Yes	48	4	8.33	2.32-19.98
No	859	50	5.82	4.35-7.60
Clean animal at birth				
Yes	96	5	5.21	1.71-11.73
No	811	49	6.04	4.50-7.91

4.3.2 Urban Slum Community

The final XGBoost model for urban slum communities that was used to predict the seropositive and seronegative participants for the testing data performs well with an average AUC value of 95.09% (range: 87.08%-98.36%). The most important drivers in urban slum communities were the square meters of shrub terrain in 100-meter radius (*shrub*) followed by age, and the number of houses in a 100-meter radius (*houses*) (Figure 4-3A). The probability of *Leptospira* exposure, in general, increases as the areas

of shrub terrain within 100-meter radius increase (Figure 4). Males have higher probability of getting exposed to *Leptospira* than females. The probability of exposure is positively correlated with the number of houses in a 100-meter radius. Additionally, model results revealed that environmental drivers such as areas of wetlands (*wetland*), tree terrain (*tree*), and lower vegetation (*lowveg*) in a 100-meter radius can

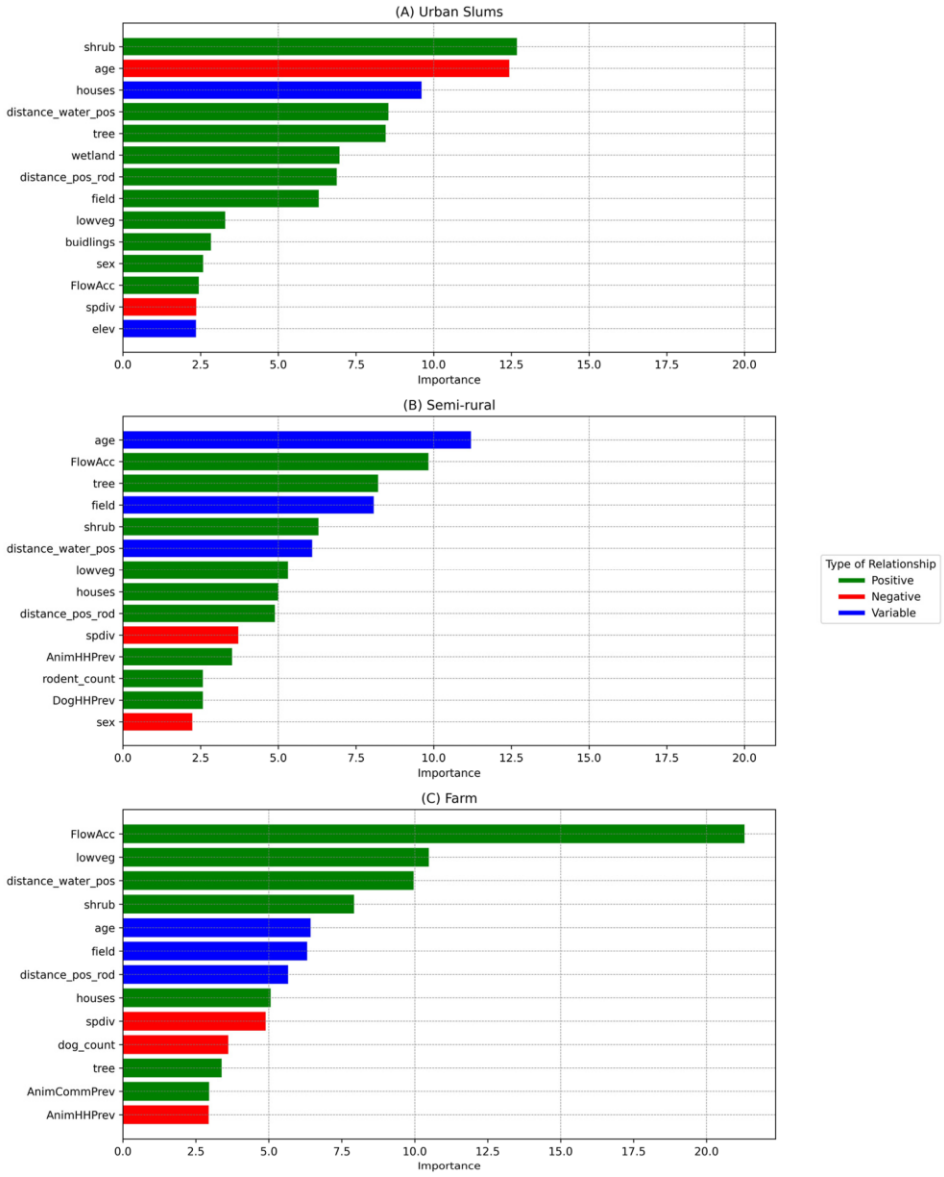


Figure 4-3: Variable Importance plot for the model results of the three community types, A) Urban slums, B) Semi-Rural, and C) Farm. Variables with more than a 2% importance frequency score are shown. The bars were color-coded for ease of interpretation of the overall trend in the predicted relationship between each variable and human exposure probability across the three different community types. Note that the relationship types depicted here were derived visually from Figure 4-4, Figure 4-5 and Figure 4-6 and were simplified to overall positive or negative, but actual relationships are non-linear.

positively impact *Leptospira* exposure. The local ecology of *Leptospira* was reflected in the relationship between probability of human exposure and the number of households within 100-meter with positive water samples (*distance_water_pos*) and the number of households within 100-meter with positive rodents (*distance_pos_rod*) (Figure 4-4).

4.3.3 Semi-rural Community

The AUC value from model performance for the model of semi-rural communities was 88.27% (range: 82.63–90.04%). The most important drivers of Leptospira exposure in semi-rural communities were age, slope of the terrain (FlowAcc), and tree and field areas (Figure 4-3B). The probability of Leptospira exposure was higher among the youngest and the oldest people, in particular those older than 60 years old (Figure 4-5). The probability of exposure increased when the slope of the terrain increased. There was a higher exposure probability when there was high tree coverage, while this effect was opposite for field area coverage.

With a relatively low importance score, several animal-related drivers emerged in the model (Figure 3B). Household seroprevalence of Leptospira in animals (AnimHHPrev), trapped rodent count (rodent_count), and household seroprevalence of Leptospira in dogs (DogHHPrev) were positively associated with exposure probability, while species diversity (spdiv) was negatively associated. Exposure probability generally increased as the number of households within 100 m with positive rodents (distance_pos_rod) increased (Figure 4-5).

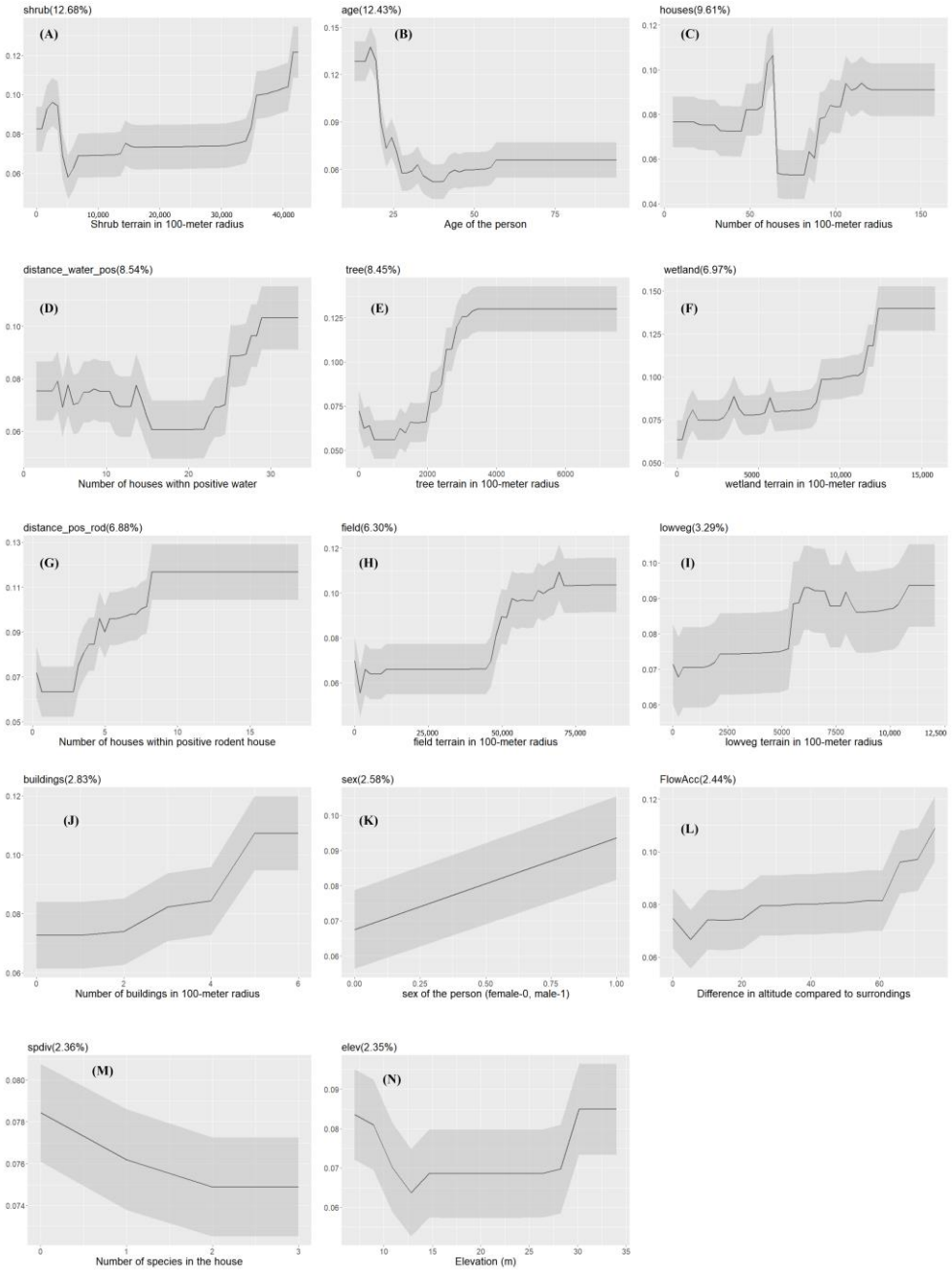


Figure 4-4: The partial dependency plots of covariates used in the model for urban slum communities. The gray shading area indicates the confidence intervals derived from the model iterations. The variable importance score is reported in the parentheses ((**A**–**N**): partial dependency of each variable in the model).

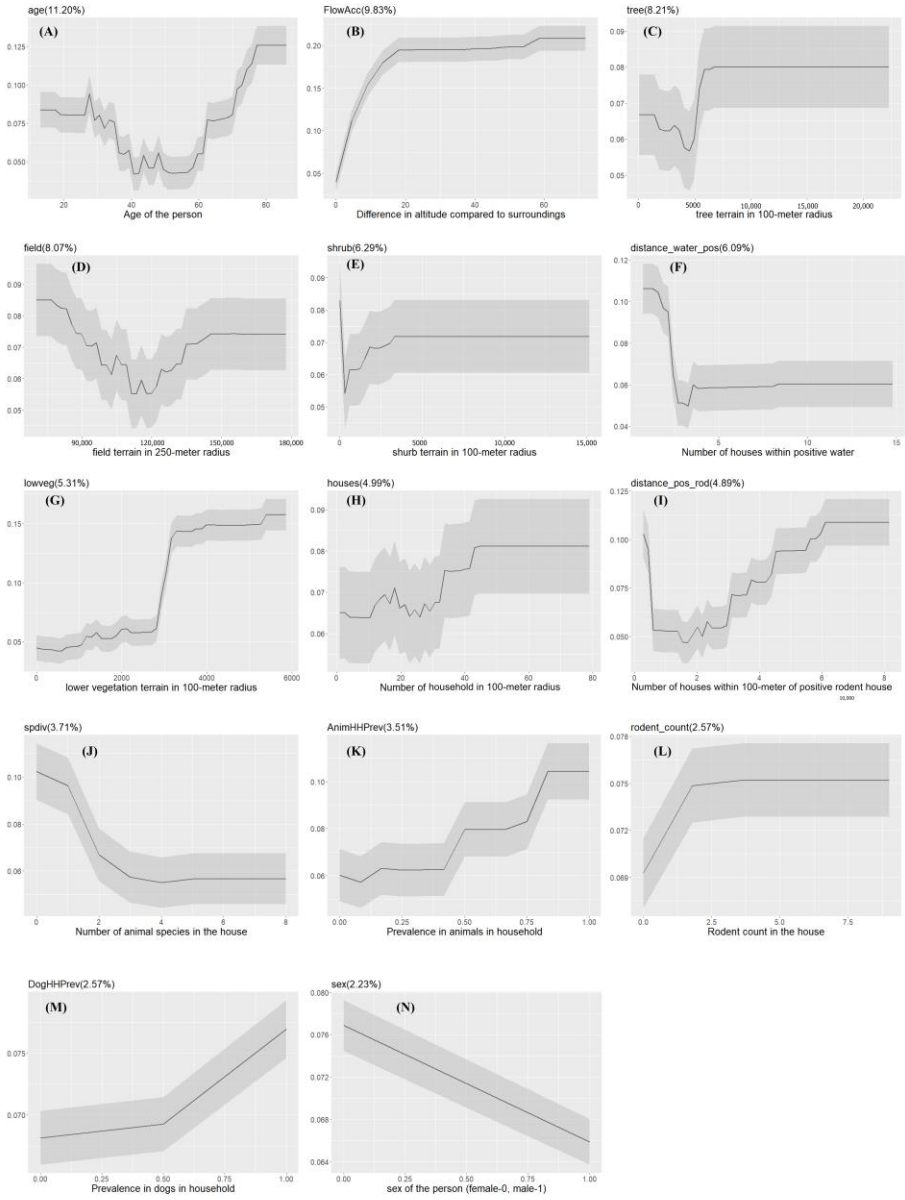


Figure 4-5: The partial dependency plots of covariates used in the model for semi-rural communities. The gray shading area indicates the confidence intervals derived from the model iterations. The variable importance score is reported in the parentheses ((**A**–**N**): partial dependency of each variable in the model).

4.3.4 Farm Community

The XGBoost model showed high prediction of seropositive and seronegative individuals with an AUC value of 91.74% (range: 84.56–95.13%). Slope of the terrain (FlowAcc), low vegetation area (lowveg), and the number of households within 100 m with positive water

samples (distance_water_pos) were significantly affecting human Leptospira exposure (Figure 4-3C). The probability of Leptospira exposure was positively correlated with terrain slope and low vegetation and negatively correlated with the number of households within 100 m with positive water samples (Figure 4-6).

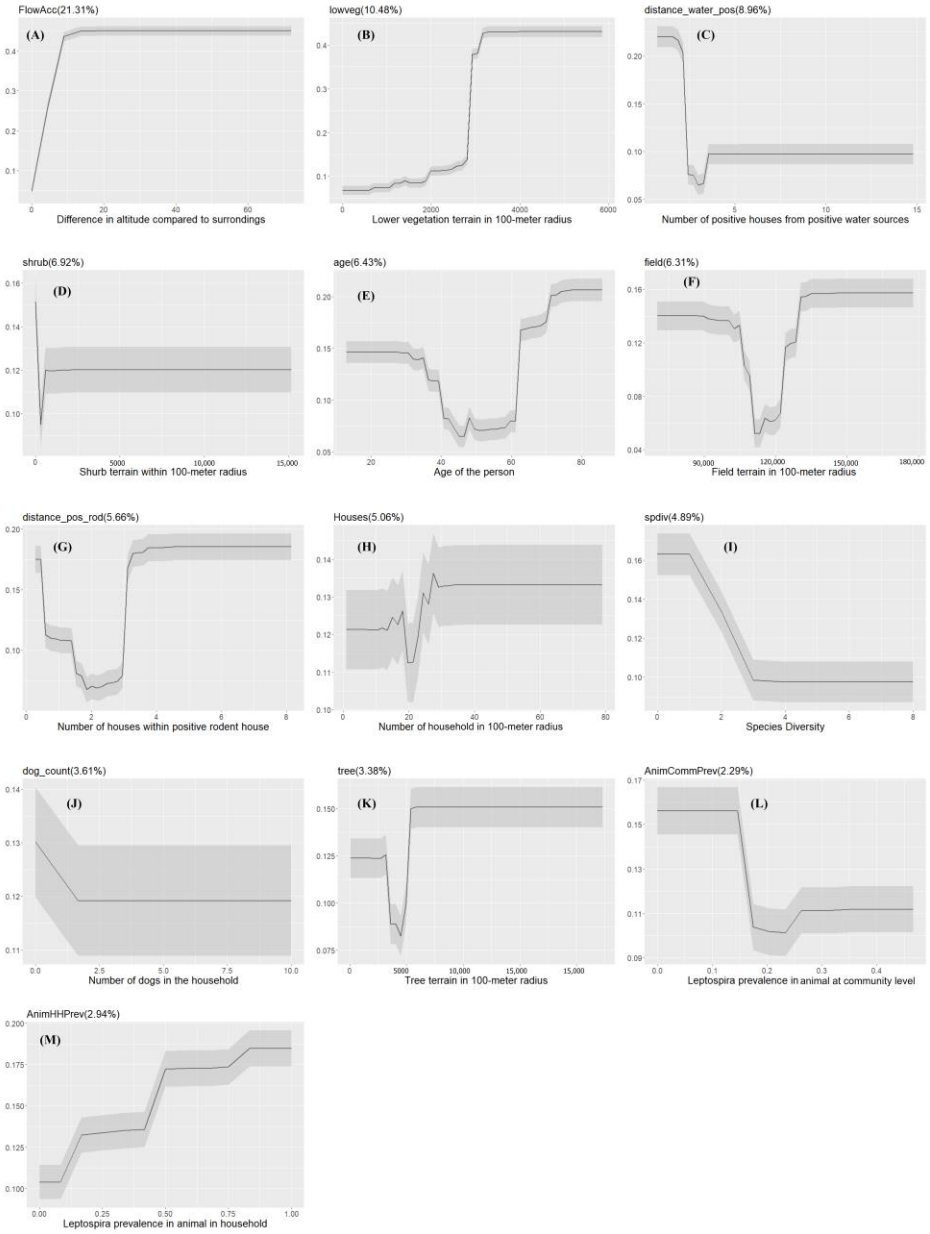


Figure 4-6: The partial dependency plots of covariates used in the model for farm communities. The gray shading area indicates the confidence intervals derived from the model iterations. The variable importance score is reported in the parentheses ((**A**–**M**): partial dependency of each variable in the model).

Other results showed that the exposure probability increased with the increasing seroprevalence of Leptospira among animals in a household (AnimHHPrev), while species diversity (spdiv), the number of dogs in the household (dog_count), and community-level seroprevalence in animals (AnimCommPrev) were negatively associated with the probability of Leptospira exposure. Age had a U-shape type of relationship, similar to the semi-rural community type, indicating a higher exposure probability for the younger and older ages. For the number of households within 100 m with positive rodents (distance_pos_rod), the relationship was variable, but it increased sharply when there were three or more households within 100 m with positive rodents (Figure 4-6).

4.4 Discussions

Leptospirosis, a zoonotic disease caused by the pathogenic spirochetes of the genus Leptospira, poses significant public health challenges worldwide, especially in tropical regions (Karpagam & Ganesh, 2020). It is also present in temperate regions and, using One Health principles, our case study unveiled the intricate landscape of leptospirosis transmission across varied community settings in Chile. We found distinct seroprevalences and risk factors associated with demographic, behavioral, environmental, and animal-related variables in those communities. Urban slum areas showed higher exposure probabilities linked to environmental factors like shrub terrain and positive water samples, while semi-rural and farm communities exhibited different patterns influenced by age, household characteristics, and animal prevalence. These findings shed light on the multifaceted nature of leptospirosis transmission, informing targeted interventions for reducing human exposure and enhancing public health efforts.

Our research combined data from systematic efforts to detect leptospirosis in humans, animals, and the peri-domestic environment, unveiling the presence of Leptospira across the

distinct community types in the study area of Chile and with varying seroprevalences among humans and domestic animals, rodent kidney carriages, and environmental contamination. The overall 6% seroprevalence in humans is similar to several studies in Colombia and Mexico in similar socio-economic settings (Romero et al., 2010; Alvarado-Esquivel et la., 2016; Benschop et al., 2009). Previous studies found that the inadequate living conditions encompassed dirt floors, proximity to sewage, and absence of proper sanitation as well as behavioral factors such as walking barefoot, having uncovered wounds, and gathering firewood are contributors to exposure (Mwachui et al., 2015); however, our modeling did not reveal any individual behaviors as significant risk factors. This may be because disease transmission is primarily determined by the living environment, where external factors can have a more significant impact on disease spread, overshadowing the influence of individual behaviors. It could also be because MAT antibodies are evidence of an exposure in the past that may not be reflected in the behaviors reported in the survey. Based on our overall descriptive analysis, males exhibited a higher seroprevalence (6.5%) compared to females (5.6%), suggesting potential gender-specific differences in exposure or susceptibility to infection which has often been reported in the existing literature (Carrero et al., 2017; Dias et al., 2007; Wynwood et al., 2014). However, when analyzing by community type, our XGBoost models revealed contrasting effects in which men had higher exposure probability than women in urban slum communities, but women had higher exposure probability than men in semi-rural communities. This finding may reflect differences in behavioral patterns, occupational activities, or exposure to contaminated environments between genders (Notobroto et al., 2021). Models consistently revealed age as an important factor but with different relationships. There was a negative association with age in urban slum communities, with a higher probability at younger ages, but it had was U-shape in semi-rural and

rural communities. The different trends could reflect different demographics and associated activities in the communities. For example, people from urban slum communities tended to be younger (Awoniyi et al., 2022) and the higher probability among this group could reflect their greater representation. In semi-rural and rural, higher exposure probability among the youngest (<40 years) and the oldest (>60 years) ages could reflect common recreational, occupational, and domestic activities with exposure to a contaminated environment or infected farm animals such as swimming, gardening, or livestock management. The high exposure probability among older adults may also be a reflection of a longer period at risk to become exposed.

Several environmental variables such as shrub terrain, wetlands, tree terrain, lower vegetation, and field terrain showed to impact the likelihood of Leptospira exposure, suggesting the importance of considering landscape heterogeneity in assessing exposure risk across different types of communities. The identified positive correlation with vegetation covers surrounding the households such as the presence of trees and lower vegetation (i.e., bushes) in all community types and shrubs in urban communities highlights the interconnectedness of wildlife habitats and leptospirosis risk (Benschop et al., 2009). Those land covers can act as critical habitats for wildlife reservoirs, notably rodents, increasing the likelihood of human-animal contact and subsequent disease transmission (Farooq et al., 2022; Jin & Ling, 2005; Carrero et al., 2017). Additionally, the shaded, humid conditions under vegetation cover can modify microclimatic conditions by trapping and retaining moisture in the soil, creating damp conditions that are ideal for the survival of Leptospira outside of its hosts. This extended survival in moist soil increases the duration during which humans can encounter Leptospira, further facilitating transmission (Moseley et al., 2018; Chiani et al., 2023; Caley & Ramsey, 2001). Furthermore, the presence of wetlands was positively associated with the likelihood of exposure in semi-rural and farm

communities which is consistent with the notion that they can act as reservoirs for Leptospira (Cucchi et al., 2019; Cunha et al., 2022) facilitating its survival and dissemination through water-borne transmission routes (Goarant, 2016; Luna et al., 2020). The presence of open fields can also be a contributing factor because it can lead to trash accumulation, attracting dogs and rodents, as well as standing water. This factor manifested differently as it was positively correlated with the probability of exposure in urban slums; however, the pattern was different for semi-rural and farm communities, likely due to interactions with other vegetation features and the built-in environment. It is widely recognized that downward slopes enhance surface water flow, leading to the accumulation of stagnant water bodies and moist soil ideal for Leptospira survival (Chiani et al., 2023). The flow accumulation variable used in this analysis was found to be important in the three community types. Derived from slope data, this variable aimed to capture land features that facilitate water movement downhill during rainfall or flooding which results in resuspension of Leptospira present in the soil and sediment, leading to areas with high environmental contamination and high exposure risk (Cunha et al., 2022).

Leptospirosis has been associated with urbanization in which the conditions of high-density and low-income housing are highly suitable for Leptospira contamination and transmission (Belsare et al., 2022; Caley & Ramsey, 2001). This effect of high-density housing was evident across all three community types when measured by the number of houses in a 100-meter radius. Additionally, the number of households within a 100-meter radius with water samples positive for pathogenic Leptospira emerged as a factor positively associated with exposure risk for the urban slum community type, further supporting the idea of exposure risk from of *Leptospira* contaminated peri-domestic environment in these vulnerable communities (Kembhavi et al., 2021; Zakharova et al., 2021). However, this relationship differed in semi-

rural and farm community types when exposure risk was the highest when there were few houses with positive water samples in the immediate vicinity. This could be an effect of water sources, which are often shared or communal, making contamination more widespread, or that rural households have greater exposure to environmental factors such as open fields or livestock, making localized household contamination less relevant (Daniels et al., 2023). Additionally, mobility patterns and water usage behaviors in rural communities might lead to a mismatch between household proximity to contaminated water and actual exposure risk (Galan et al., 2021).

Although animal-related variables were present across all communities, it was more evident in semi-rural and farm communities. The variable measuring the number of houses with Leptospira-positive rodents in the surrounding area was found to impact exposure risk in all three types of communities. There was a clear positive trend in urban slum communities, which is consistent with the findings regarding water sample contamination. This is also consistent with the general knowledge that synanthropic rodents are ubiquitous reservoirs of Leptospira (Sluydts et al., 2022). Although with a lower importance level in the model results compared with other variables, dog household seroprevalence in semi-rural communities was associated with increased exposure risk. Dogs are recognized reservoir hosts for leptospirosis, capable of shedding the bacterium in their urine and contaminating the environment (Céspedes Z et al., 2003). The often-limited access to veterinary care and vaccination in low-resource settings leads to underdiagnosis and untreated infections (Brockmann et al., 2016; Maze et al., 2018). A high seroprevalence of Leptospira in dogs within households could be correlated to increased likelihood of human exposure because of transmission from the dogs or a dog-contaminated environment. Alternatively, both dogs and people could be subject to the similar sources of

Leptospira. Although also of low importance level in the modeling results, it is worth noting that the seroprevalence of Leptospira in all farm animals within the household was positively associated with exposure risk in semi-rural and farm communities. Households in these communities had a variety of livestock such as cattle, sheep, and pigs which can contribute to environmental contamination and pose a risk to household members through occupational and/or domestic activities (Anderson et al., 2023; Daud et al., 2018; Harran et al., 2023). The variable could represent a measurement of the underlying Leptospira exposure risk at the household level. The inverse relationship observed between increasing species diversity and leptospirosis exposure more importantly in rural communities, may reflect several ecological and epidemiological factors. A low diversity of domestic animal species in the farm system investigated here could indicate higher abundance of more competent hosts that contribute to increase transmission and higher likelihood of human exposure (Moseley et al., 2018). This association could also be the result of farm management practices and other activities correlated with both likelihood of exposure in people and species diversity.

Our study has several limitations. Firstly, the cross-sectional nature of the study design limits our ability to capture the temporal change in the relationship between identified risk factors and Leptospira exposure in community participants. Leptospirosis status in the community participants was measured using presence of antibodies which indicate an exposure sometime in the past, and although the population in the study communities were stable regarding time of residence in the sampled locations, some of the factors may not reflect the same conditions as the time of the exposure. Additionally, the use of self-reported data and reliance on recall of the participants may have introduced information bias to reported behaviors. Furthermore, the multi-strain and multi-host transmission dynamics of Leptospira are complex,

but our study focused solely on the broad relationships between various eco-epidemiological drivers and human exposure, but not on the underlying mechanisms or pathways of transmission or how drivers may interact with one another to influence exposure risk. The benefit of using cross-sectional sampling is that it allows for a cost-efficient comprehensive investigation and is particularly useful as an initial approach in areas with limited knowledge of the leptospirosis situation. Findings can generate hypothesis for future research incorporating longitudinal study designs with incidence cases and genomic approaches for strain identification in humans, animals, and the environment and provide further insights into the transmission dynamics of leptospirosis.

4.5 Conclusions

Across urban slums, semi-rural, and farm communities, the overall human seroprevalence by community type was similar; however, our study results show the intricate interplay between environmental, socio-demographic, and animal-related factors in shaping leptospirosis transmission dynamics across these different community types. In urban slum areas, densely populated environments and altered landscapes contribute to disease risk, with environmental and rodent-related factors driving the transmission. Similarly, semi-rural communities exhibit complex interactions between environmental features and animal-related variables influencing exposure probability. In contrast, farm communities present unique challenges characterized by the coexistence of agricultural practices and human habitation. Here, environmental factors interact with the presence of livestock to shape disease dynamics. Emphasizing a One Health approach that recognizes the interconnectedness of human, animal, and environmental health is paramount to addressing the complex nature of leptospirosis transmission. The case study presented here is an example of the integration of efforts across disciplines for data collection at

the human–environment–animal interface and novel methods such as machine learning for the analysis of complex data. Adopting a holistic approach that considers the health of humans, animals, and ecosystems is essential for achieving sustainable disease control and ensuring the health and prosperity of communities worldwide.

References

Alexander, A. D. (1960). The distribution of leptospirosis in Latin America. *Bull World Health Organ*, *23*(1), 113-125.

Anderson, T., Hamond, C., Haluch, A., Toot, K., Nally, J. E., LeCount, K., & Schlater, L. K. (2023). Animals Exposed to Leptospira Serogroups Not Included in Bacterins in the United States and Puerto Rico. *Trop Med Infect Dis*, *8*(3). https://doi.org/10.3390/tropicalmed8030183

Belsare, A. V., Gompper, M. E., Mason, M., & Munoz-Zanzi, C. (2022). Investigating Leptospira dynamics in a multi-host community using an agent-based modelling approach. *Transboundary and Emerging Diseases*, *69*(6), 3780-3789. https://doi.org/https://doi.org/10.1111/tbed.14750

Bolin, C. A., & Koellner, P. (1988). Human-to-Human Transmission of Leptospira interrogans by Milk. *Journal of Infectious Diseases*, *158*(1), 246-247. https://doi.org/10.1093/infdis/158.1.246

Bradley, E. A., & Lockaby, G. (2023). Leptospirosis and the Environment: A Review and Future Directions. *Pathogens*, *12*(9), 1167. https://www.mdpi.com/2076-0817/12/9/1167 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10538202/pdf/pathogens-12-01167.pdf

Brockmann, S. O., Ulrich, L., Piechotowski, I., Wagner-Wiening, C., Nockler, K., Mayer-Scholl, A., & Eichner, M. (2016). Risk factors for human Leptospira seropositivity in South Germany. *Springerplus*, *5*(1), 1796. https://doi.org/10.1186/s40064-016-3483-8

Caley, P., & Ramsey, D. (2001). Estimating disease transmission in wildlife, with emphasis on leptospirosis and bovine tuberculosis in possums, and effects of fertility control. *Journal of Applied Ecology*, *38*(6), 1362-1370. https://doi.org/https://doi.org/10.1046/j.0021-8901.2001.00676.x

Céspedes Z, M., Ormaeche M, M., Condori, P., Balda J, L., & Glenny A, M. (2003). Prevalencia de Leptospirosis y factores de riesgo en personas con antecedentes de fiebre en la Provincia de Manu, Madre de Dios, Perú. *Revista Peruana de Medicina Experimental y Salud Publica*, *20*, 80-185. http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S1726-46342003000400002&nrm=iso

Chen, T., & Guestrin, C. (2016). *XGBoost: A Scalable Tree Boosting System* Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, California, USA. https://doi.org/10.1145/2939672.2939785

Chen T, H. T., Benesty M, Khotilovich V, Tang Y, Cho H, Chen K, Mitchell R, Cano I, Zhou T, Li M, Xie J, Lin M, Geng Y, Li Y, Yuan J. (2023). *xgboost: Extreme Gradient Boosting*. . In (Version R package version 1.7.5.1)

Chiani, Y. T., Jacob, P., Mayora, G., Aquino, D. S., Quintana, R. D., & Mesa, L. (2023). Presence of Leptospira spp. in a Mosaic of Wetlands Used for Livestock Raising under Differing Hydroclimatic Conditions. *Appl Environ Microbiol*, *89*(6), e0197122. https://doi.org/10.1128/aem.01971-22

Costa, F., Hagan, J. E., Calcagno, J., Kane, M., Torgerson, P., Martinez-Silveira, M. S., Stein, C., Abela-Ridder, B., & Ko, A. I. (2015). Global Morbidity and Mortality of Leptospirosis:

A Systematic Review. *PLoS Negl Trop Dis*, *9*(9), e0003898. https://doi.org/10.1371/journal.pntd.0003898

Crecelius, E. M., & Burnett, M. W. (2020). Leptospirosis. *Journal of Special Operations Medicine*, *20*(4), 121. https://doi.org/10.55460/8ybj-0dlp

Cucchi, K., Liu, R., Collender, P. A., Cheng, Q., Li, C., Hoover, C. M., Chang, H. H., Liang, S., Yang, C., & Remais, J. V. (2019). Hydroclimatic drivers of highly seasonal leptospirosis incidence suggest prominent soil reservoir of pathogenic Leptospira spp. in rural western China. *PLoS Negl Trop Dis*, *13*(12), e0007968. https://doi.org/10.1371/journal.pntd.0007968

Cunha, M., Costa, F., Ribeiro, G. S., Carvalho, M. S., Reis, R. B., Nery, N., Jr., Pischel, L., Gouveia, E. L., Santos, A. C., Queiroz, A., Wunder, E. A., Jr., Reis, M. G., Diggle, P. J., & Ko, A. I. (2022). Rainfall and other meteorological factors as drivers of urban transmission of leptospirosis. *PLoS Negl Trop Dis*, *16*(4), e0007507. https://doi.org/10.1371/journal.pntd.0007507

Daud, A. B., Mohd Fuzi, N. M. H., Wan Mohammad, W. M. Z., Amran, F., Ismail, N., Arshad, M. M., & Kamarudin, S. (2018). Leptospirosis and Workplace Environmental Risk Factors among Cattle Farmers in Northeastern Malaysia. *Int J Occup Environ Med*, *9*(2), 88-96. https://doi.org/10.15171/ijoem.2018.1164

Davagdorj, K., Pham, V. H., Theera-Umpon, N., & Ryu, K. H. (2020). XGBoost-Based Framework for Smoking-Induced Noncommunicable Disease Prediction. *Int J Environ Res Public Health*, *17*(18). https://doi.org/10.3390/ijerph17186513

Ellis, W. A. (2015). Animal leptospirosis. *Curr Top Microbiol Immunol*, *387*, 99-137. https://doi.org/10.1007/978-3-662-45059-8_6

Ergul Aydin, Z., & Kamisli Ozturk, Z. (2021). *Performance Analysis of XGBoost Classifier with Missing Data*.

Farooq, Z., Rocklöv, J., Wallin, J., Abiri, N., Sewe, M. O., Sjödin, H., & Semenza, J. C. (2022). Artificial intelligence to predict West Nile virus outbreaks with eco-climatic drivers. *Lancet Reg Health Eur*, *17*, 100370. https://doi.org/10.1016/j.lanepe.2022.100370

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. *International Journal of Climatology*, *37*(12), 4302-4315. https://doi.org/https://doi.org/10.1002/joc.5086

Goarant, C. (2016). Leptospirosis: risk factors and management challenges in developing countries. *Res Rep Trop Med*, *7*, 49-62. https://doi.org/10.2147/RRTM.S102543

Guerra, M. A. (2013). Leptospirosis: Public health perspectives. *Biologicals*, *41*(5), 295-297. https://doi.org/10.1016/j.biologicals.2013.06.010

Haake, D. A., & Levett, P. N. (2015). Leptospirosis in Humans. In (pp. 65-97). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-45059-8_5

Harran, E., Pinot, A., Kodjo, A., Djelouadji, Z., Le Gudayer, M., Sionfoungo Daouda, S., Groud, K., Lattard, V., & Ayral, F. (2023). Identification of Pathogenic Leptospira kirschneri Serogroup Grippotyphosa in Water Voles (Arvicola terrestris) from Ruminant Pastures in Puy-de-Dome, Central France. *Pathogens*, *12*(2). https://doi.org/10.3390/pathogens12020260

Harrison, N. A., & Fitzgerald, W. R. (1988). Leptospirosis--can it be a sexually transmitted disease? *Postgrad Med J*, *64*(748), 163-164. https://doi.org/10.1136/pgmj.64.748.163

Jin, H., & Ling, C. X. (2005). Using AUC and accuracy in evaluating learning algorithms. *IEEE Transactions on Knowledge and Data Engineering*, *17*(3), 299-310. https://doi.org/10.1109/TKDE.2005.50

Karpagam, K. B., & Ganesh, B. (2020). Leptospirosis: a neglected tropical zoonotic infection of public health importance-an updated review. *Eur J Clin Microbiol Infect Dis*, *39*(5), 835-846. https://doi.org/10.1007/s10096-019-03797-4

Kembhavi, R. S., Velhal, G. D., & Shah, A. K. (2021). Epidemiological determinants of leptospirosis in rural and urban districts of Maharashtra, India. *J Family Med Prim Care*, *10*(9), 3361-3367. https://doi.org/10.4103/jfmpc.jfmpc_674_21

Lelu, M., Muñoz-Zanzi, C., Higgins, B., & Galloway, R. (2015). Seroepidemiology of leptospirosis in dogs from rural and slum communities of Los Rios Region, Chile. *BMC Veterinary Research*, *11*(1), 31. https://doi.org/10.1186/s12917-015-0341-9

Luna, J., Salgado, M., Tejeda, C., Moroni, M., & Monti, G. (2020). Assessment of Risk Factors in Synanthropic and Wild Rodents Infected by Pathogenic Leptospira spp. Captured in Southern Chile. *Animals*, *10*(11), 2133. https://www.mdpi.com/2076-2615/10/11/2133
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7697743/pdf/animals-10-02133.pdf

Mason, M. R., Encina, C., Sreevatsan, S., & Muñoz-Zanzi, C. (2016). Distribution and Diversity of Pathogenic Leptospira Species in Peri-domestic Surface Waters from South Central Chile. *PLOS Neglected Tropical Diseases*, *10*(8), e0004895. https://doi.org/10.1371/journal.pntd.0004895

Maze, M. J., Cash-Goldwasser, S., Rubach, M. P., Biggs, H. M., Galloway, R. L., Sharples, K. J., Allan, K. J., Halliday, J. E. B., Cleaveland, S., Shand, M. C., Muiruri, C., Kazwala, R. R., Saganda, W., Lwezaula, B. F., Mmbaga, B. T., Maro, V. P., & Crump, J. A. (2018). Risk factors for human acute leptospirosis in northern Tanzania. *PLoS Negl Trop Dis*, *12*(6), e0006372. https://doi.org/10.1371/journal.pntd.0006372

Miller, D. A., Wilson, M. A., & Beran, G. W. (1991). Relationships between prevalence of Leptospira interrogans in cattle, and regional, climatic, and seasonal factors. *Am J Vet Res*, *52*(11), 1766-1768.

Montes, V., & Monti, G. (2021). Pathogenic Leptospira spp. Seroprevalence and Herd-Level Risk Factors Associated with Chilean Dairy Cattle. *Animals*, *11*(11), 3148. https://www.mdpi.com/2076-2615/11/11/3148
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8614305/pdf/animals-11-03148.pdf

Moseley, M., Rahelinirina, S., Rajerison, M., Garin, B., Piertney, S., & Telfer, S. (2018). Mixed Leptospira Infections in a Diverse Reservoir Host Community, Madagascar, 2013-2015. *Emerg Infect Dis*, *24*(6), 1138-1140. https://doi.org/10.3201/eid2406.180035

Munoz-Zanzi, C., Campbell, C., & Berg, S. (2016). Seroepidemiology of toxoplasmosis in rural and urban communities from Los Rios Region, Chile. *Infect Ecol Epidemiol*, *6*, 30597. https://doi.org/10.3402/iee.v6.30597

Muñoz-Zanzi, C., Mason, M., Encina, C., Astroza, A., & Romero, A. (2014). Leptospira Contamination in Household and Environmental Water in Rural Communities in Southern Chile. *International Journal of Environmental Research and Public Health*, *11*(7), 6666-6680. https://doi.org/10.3390/ijerph110706666

Muñoz-Zanzi, C., Mason, M., Encina, C., Gonzalez, M., & Berg, S. (2014). Household characteristics associated with rodent presence and Leptospira infection in rural and urban communities from Southern Chile. *Am J Trop Med Hyg*, *90*(3), 497-506. https://doi.org/10.4269/ajtmh.13-0334

Mwachui, M. A., Crump, L., Hartskeerl, R., Zinsstag, J., & Hattendorf, J. (2015). Environmental and Behavioural Determinants of Leptospirosis Transmission: A Systematic Review. *PLOS Neglected Tropical Diseases*, *9*(9), e0003843. https://doi.org/10.1371/journal.pntd.0003843

Notobroto, H. B., Mirasa, Y. A., & Rahman, F. S. (2021). Sociodemographic, behavioral, and environmental factors associated with the incidence of leptospirosis in highlands of Ponorogo Regency, Province of East Java, Indonesia. *Clinical Epidemiology and Global Health*, *12*. https://doi.org/10.1016/j.cegh.2021.100911

Ogunleye, A., & Wang, Q. G. (2020). XGBoost Model for Chronic Kidney Disease Diagnosis. *IEEE/ACM Trans Comput Biol Bioinform*, *17*(6), 2131-2140. https://doi.org/10.1109/tcbb.2019.2911071

Pappas, G., Papadimitriou, P., Siozopoulou, V., Christou, L., & Akritidis, N. (2008). The globalization of leptospirosis: worldwide incidence trends. *Int J Infect Dis*, *12*(4), 351-357. https://doi.org/10.1016/j.ijid.2007.09.011

Putatunda, S., & Rama, K. (2018). *A Comparative Analysis of Hyperopt as Against Other Approaches for Hyper-Parameter Optimization of XGBoost* Proceedings of the 2018 International Conference on Signal Processing and Machine Learning, Shanghai, China. https://doi.org/10.1145/3297067.3297080

Romero, E. C., Bernardo, C. C., & Yasuda, P. H. (2003). Human leptospirosis: a twenty-nine-year serological study in São Paulo, Brazil. *Rev Inst Med Trop Sao Paulo*, *45*(5), 245-248. https://doi.org/10.1590/s0036-46652003000500002

Shaheed, K., Abbas, Q., Hussain, A., & Qureshi, I. (2023). Optimized Xception Learning Model and XgBoost Classifier for Detection of Multiclass Chest Disease from X-ray Images. *Diagnostics (Basel)*, *13*(15). https://doi.org/10.3390/diagnostics13152583

Sluydts, V., Sarathchandra, S. R., Piscitelli, A. P., Van Houtte, N., Gryseels, S., Mayer-Scholl, A., Bier, N. S., Htwe, N. M., & Jacob, J. (2022). Ecology and distribution of Leptospira spp., reservoir hosts and environmental interaction in Sri Lanka, with identification of a new strain. *PLoS Negl Trop Dis*, *16*(9), e0010757. https://doi.org/10.1371/journal.pntd.0010757

Srinivas, P., & Katarya, R. (2022). hyOPTXg: OPTUNA hyper-parameter optimization framework for predicting cardiovascular disease using XGBoost. *Biomedical Signal Processing and Control*, *73*, 103456. https://doi.org/https://doi.org/10.1016/j.bspc.2021.103456

Team, R. C. (2020). *R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.* In (Version 4.3.1)

Zakharova, O. I., Korennoy, F. I., Iashin, I. V., Toropova, N. N., Gogin, A. E., Kolbasov, D. V., Surkova, G. V., Malkhazova, S. M., & Blokhin, A. A. (2021). Ecological and Socio-Economic Determinants of Livestock Animal Leptospirosis in the Russian Arctic. *Front Vet Sci*, *8*, 658675. https://doi.org/10.3389/fvets.2021.658675

Zamora, J., Riedemann, S., Montecinos, M. I., & Cabezas, X. (1990). [Serological survey of human leptospirosis in a high risk population in Chile]. *Rev Med Chil*, *118*(3), 247-252. (Encuesta serológica de leptospirosis humana en ocupaciones de alto riesgo en Chile.)

Chapter 5 : Synthesis, Conclusions and Future Directions

5.1 Integrative Synthesis of Research Findings

A central challenge in disease ecology is understanding how pathogens move through increasingly complex ecological systems. While many studies focus on individual hosts, single species, or specific transmission pathways, infectious diseases often emerge from interactions occurring across multiple ecological levels (Engering et al., 2013; Parrish et al., 2008; Coker et al., 2011; Adisasmito et al., 2022). The overarching objective of this dissertation was to investigate transmission dynamics across a continuum of ecological complexity, progressing from transmission within wildlife populations, to transmission between host species, and ultimately to transmission occurring across wildlife–livestock–human systems within a One Health framework. By examining these different transmission scenarios, this dissertation sought to provide a broader understanding of the ecological processes that facilitate pathogen persistence and spread.

The progression of chapters reflects an expansion in the scale and complexity of transmission processes. Chapter 2 focused on transmission within a single wildlife host population and demonstrated the importance of movement behavior and contact patterns in generating opportunities for pathogen spread. These findings are consistent with the growing recognition that contact structure and host movement are fundamental determinants of infectious disease transmission in wildlife populations (Craft, 2015; White et al., 2017). Building upon this framework, Chapter 3 expanded the investigation to cross-species transmission, emphasizing the role of ecological communities, shared environments, and landscape characteristics in facilitating

pathogen exchange among host species. This shift from individual-level interactions to community-level processes illustrates how transmission dynamics become increasingly influenced by broader ecological contexts as additional host species are incorporated into the system (Choudhury et al., 2022; Becker et al., 2020).

Chapter 4 extended this progression by examining disease transmission within a One Health framework that explicitly incorporated wildlife, domestic animals, humans, and environmental reservoirs. Unlike the previous chapters, which focused on transmission within and between wildlife hosts, this chapter considered how environmental, biological, and socio-demographic factors interact to influence human exposure. The analyses demonstrated that the relative importance of these factors varied among urban slum, semi-rural, and agricultural communities, emphasizing that zoonotic disease transmission is inherently context-dependent and cannot be explained by a single ecological or epidemiological driver (Costa et al., 2015; Mwachui et al., 2015).

Taken together, the three chapters demonstrate that the mechanisms governing transmission evolve as disease systems become increasingly complex. Within a single host population, transmission opportunities were primarily determined by host movement and behavior. Across species, transmission depended on ecological overlap and shared environmental conditions. At the human–animal interface, transmission emerged from interactions among environmental reservoirs, animal hosts, and human activities. This progression illustrates that understanding environmentally mediated infectious diseases requires moving beyond individual host populations and considering the ecological processes that connect wildlife, livestock, humans, and their shared environments. Such an integrated perspective forms the foundation of

contemporary One Health approaches to infectious disease research and management (Engering et al., 2013; Adisasmito et al., 2022).

5.2 Implications for Disease Ecology, Surveillance, and One Health

A major contribution of this dissertation is the demonstration that transmission can be investigated through ecological processes even when transmission events themselves are difficult or impossible to observe directly. Direct observation of pathogen movement among hosts is rarely feasible in wildlife and environmental systems. Consequently, disease ecology often relies on indirect indicators of transmission potential and pathogen movement (Craft, 2015; White et al., 2017; Manlove et al., 2022). Across the three chapters, transmission was examined through different ecological manifestations, including contact opportunities among individuals, shared occupancy among host species, and exposure pathways linking environmental reservoirs, animals, and humans. Collectively, these approaches illustrate how transmission can be conceptualized as a series of ecological processes that create opportunities for pathogen movement rather than solely as documented infection events.

The dissertation also highlights the value of distinguishing between transmission potential, transmission risk, and realized disease outcomes. The feral pig contact networks quantified opportunities for transmission, the avian influenza analyses identified ecological conditions associated with pathogen exchange among species, and the leptospirosis analyses examined factors associated with human exposure within complex socio-ecological systems. Although these represent different stages of the disease process, each contributes important information about how pathogens move through ecological systems. This perspective aligns with contemporary disease ecology frameworks that recognize transmission as a multistage process

linking contact, exposure, infection, and disease outcomes (Engering et al., 2013; Manlove et al., 2022; Becker et al., 2020).

Finally, this dissertation emphasizes that many infectious diseases are governed by processes that occur before infection is ever detected. Changes in host behavior, species interactions, environmental conditions, and exposure pathways may alter transmission dynamics long before changes become apparent in surveillance data. By focusing on these upstream ecological processes, this dissertation supports a preventive rather than reactive view of disease ecology, where understanding the conditions that facilitate transmission becomes as important as documenting infection itself (Allen et al., 2017; Hayman et al., 2023). Such an approach is increasingly important for anticipating zoonotic disease emergence and developing proactive surveillance strategies in rapidly changing ecological systems.

The findings of this dissertation suggest that infectious disease surveillance can be strengthened by shifting from host-centered approaches toward transmission-centered approaches. Chapter 2 demonstrated that transmission opportunities among feral pigs are concentrated within specific contact networks and locations rather than being uniformly distributed across populations. Similarly, Chapter 3 showed that cross-species avian influenza occupancy is associated with identifiable ecological conditions, particularly wetlands, hydrologically connected landscapes, and areas supporting large populations of migratory birds. Together, these findings indicate that surveillance efforts may be improved by prioritizing ecological settings where transmission is most likely to occur rather than focusing solely on areas where hosts are present. Such approaches may increase surveillance efficiency while providing earlier detection of emerging disease risks (Craft, 2015; White et al., 2017).

The results also highlight the importance of integrating environmental information into disease risk assessment frameworks. Across both avian influenza and leptospirosis systems, environmental conditions consistently influenced transmission opportunities, pathogen persistence, or exposure risk. These findings suggest that environmental surveillance, including monitoring of wetlands, water resources, landscape connectivity, and other ecological indicators, should be considered a complementary component of disease surveillance programs. Incorporating environmental data alongside traditional host-based surveillance may improve the ability to identify transmission hotspots and anticipate changes in disease risk before widespread outbreaks occur (Engering et al., 2013; Becker et al., 2020; Bradley & Lockaby, 2023).

Finally, the dissertation reinforces the practical value of One Health approaches for addressing environmentally mediated and multi-host diseases. Chapter 4 demonstrated that human exposure emerges from interactions among wildlife reservoirs, domestic animals, environmental contamination, and socio-ecological conditions, while the avian influenza analyses highlighted ecological linkages among wildlife species operating across shared landscapes. These findings suggest that disease prevention strategies will be most effective when public health agencies, wildlife managers, agricultural stakeholders, and environmental organizations coordinate surveillance and intervention efforts. As environmental change and increasing human–animal interactions continue to reshape disease systems globally, integrated One Health frameworks will become increasingly important for reducing zoonotic disease risks and improving preparedness for future disease emergence (Adisasmito et al., 2022; Hayman et al., 2023).

5.3 Limitations and Future Research Directions

Several limitations should be considered when interpreting the findings of this dissertation. First, each chapter relied on observational or surveillance-based datasets that may not fully capture all transmission events occurring within the study systems. The feral pig contact analyses quantified potential transmission opportunities rather than direct pathogen transmission, while the avian influenza analyses were based on surveillance detections that may be influenced by variation in sampling effort across space and time. Similarly, the leptospirosis analyses relied on available environmental, animal, and human exposure data, which may not fully represent the complexity of transmission pathways operating within individual communities. In Chapter 2, backward elimination was used for variable selection, which may have introduced confounding by removing variables that could still influence the relationships among predictors. In addition, vapor pressure was included as a climatic variable instead of vapor pressure deficit (VPD), which may better represent atmospheric conditions relevant to pathogen persistence and transmission.

Additional uncertainty arises from the inherent complexity of multi-host disease systems. Pathogen transmission is influenced by numerous ecological, behavioral, environmental, and socio-economic factors that are difficult to measure simultaneously. Although the analytical frameworks used in this dissertation incorporated a broad range of covariates and ecological processes, unmeasured factors may also contribute to observed transmission patterns. Furthermore, the findings represent specific pathogen systems and study regions, and caution should be exercised when extrapolating results to other diseases, geographic areas, or ecological contexts without appropriate validation.

The findings of this dissertation highlight several opportunities for future research. One important direction involves integrating movement data, pathogen surveillance, and environmental monitoring within unified analytical frameworks. Combining host movement information with pathogen occurrence data could improve understanding of how transmission pathways emerge across wildlife communities and how environmental conditions influence disease spread. Such approaches may be particularly valuable for identifying mechanisms linking within-species transmission to broader cross-species transmission processes.

Future research should also focus on strengthening predictive One Health frameworks that simultaneously incorporate wildlife, livestock, human, and environmental components of disease systems. Advances in remote sensing, environmental surveillance, genomic epidemiology, and computational modeling provide new opportunities to investigate pathogen transmission across ecological scales. Expanding these approaches to additional zoonotic diseases and geographic regions would improve understanding of how environmental change, habitat modification, and shifting species interactions influence future disease emergence and spillover risk.

5.4 Concluding Remarks

The research presented in this dissertation demonstrates that infectious disease transmission is fundamentally an ecological process shaped by interactions among hosts, environments, and landscapes. Through the examination of within-species transmission among feral pigs, cross-species avian influenza transmission between wildlife hosts, and environmentally mediated leptospirosis transmission within a One Health framework, this work illustrates how transmission dynamics become increasingly complex as additional ecological components are incorporated into disease systems. Despite differences among pathogens and

host communities, transmission consistently emerged from interactions occurring across multiple ecological scales rather than from any single factor acting independently.

Collectively, these findings reinforce the importance of integrating disease ecology, spatial epidemiology, movement ecology, and One Health approaches to better understand pathogen transmission and emergence. As environmental change, land-use modification, and human–animal interactions continue to intensify globally, addressing future infectious disease challenges will require frameworks capable of capturing these interconnected processes. By examining transmission across diverse host scenarios and ecological contexts, this dissertation contributes to a more comprehensive understanding of how pathogens move through complex wildlife–livestock–human systems and provides a foundation for future research, surveillance, and disease prevention efforts.

References

Adisasmito, W. B., Almuhairi, S., Behravesh, C. B., Bilivogui, P., Bukachi, S. A., Casas, N., Becerra, N. C., Charron, D. F., Chaudhary, A., & Zanella, J. R. C. (2022). One Health: A new definition for a sustainable and healthy future. *PLoS Pathogens, 18*(6), e1010537. https://doi.org/10.1371/journal.ppat.1010537

Becker, D. J., Albery, G. F., Kessler, M. K., Lunn, T. J., Falvo, C. A., Czirják, G. Á., Martin, L. B., & Plowright, R. K. (2020). Macroimmunology: The drivers and consequences of spatial patterns in wildlife immune defence. *Journal of Animal Ecology, 89*(4), 972–995. https://doi.org/10.1111/1365-2656.13166

Bradley, E. A., & Lockaby, G. (2023). Leptospirosis and the environment: A review and future directions. *Pathogens, 12*(9), 1167. https://doi.org/10.3390/pathogens12091167

Choudhury, P. R., Saha, T., Goel, S., Shah, J. M., & Ganjewala, D. (2022). Cross-species virus transmission and its pandemic potential. *Bulletin of the National Research Centre, 46*(1), 18. https://doi.org/10.1186/s42269-022-00747-8

Coker, R., Rushton, J., Mounier-Jack, S., Karimuribo, E., Lutumba, P., Kambarage, D., Pfeiffer, D. U., Stärk, K., & Rweyemamu, M. (2011). Towards a conceptual framework to support one-health research for policy on emerging zoonoses. *The Lancet Infectious Diseases, 11*(4), 326–331. https://doi.org/10.1016/S1473-3099(10)70312-1

Costa, F., Hagan, J. E., Calcagno, J., Kane, M., Torgerson, P., Martinez-Silveira, M. S., Stein, C., Abela-Ridder, B., & Ko, A. I. (2015). Global morbidity and mortality of leptospirosis: A systematic review. *PLoS Neglected Tropical Diseases, 9*(9), e0003898. https://doi.org/10.1371/journal.pntd.0003898

Craft, M. E. (2015). Infectious disease transmission and contact networks in wildlife and livestock. *Philosophical Transactions of the Royal Society B: Biological Sciences, 370*(1669), 20140107. https://doi.org/10.1098/rstb.2014.0107

Engering, A., Hogerwerf, L., & Slingenbergh, J. (2013). Pathogen–host–environment interplay and disease emergence. *Emerging Microbes & Infections, 2*(1), 1–7. https://doi.org/10.1038/emi.2013.5

Hayman, D. T. S., Adisasmito, W. B., Almuhairi, S., Behravesh, C. B., Bilivogui, P., Bukachi, S. A., Casas, N., Becerra, N. C., Charron, D. F., Chaudhary, A., Ciacci Zanella, J. R., Cunningham, A. A., Dar, O., Debnath, N., Dungu, B., Farag, E., Gao, G. F., Khaitsa, M., Machalaba, C., ... Koopmans, M. (2023). Developing One Health surveillance systems. *One Health, 17*, 100617. https://doi.org/10.1016/j.onehlt.2023.100617

Manlove, K., Wilber, M., White, L., Bastille-Rousseau, G., Yang, A., Gilbertson, M. L. J., Craft, M. E., Cross, P. C., Wittemyer, G., & Pepin, K. M. (2022). Defining an epidemiological

landscape that connects movement ecology to pathogen transmission and pace-of-life. *Ecology Letters, 25*(8), 1760–1782. https://doi.org/10.1111/ele.14032

Mwachui, M. A., Crump, L., Hartskeerl, R., Zinsstag, J., & Hattendorf, J. (2015). Environmental and behavioural determinants of leptospirosis transmission: A systematic review. *PLoS Neglected Tropical Diseases, 9*(9), e0003843. https://doi.org/10.1371/journal.pntd.0003843

Parrish, C. R., Holmes, E. C., Morens, D. M., Park, E.-C., Burke, D. S., Calisher, C. H., Laughlin, C. A., Saif, L. J., & Daszak, P. (2008). Cross-species virus transmission and the emergence of new epidemic diseases. *Microbiology and Molecular Biology Reviews, 72*(3), 457–470. https://doi.org/10.1128/MMBR.00004-08

White, L. A., Forester, J. D., & Craft, M. E. (2017). Using contact networks to explore mechanisms of parasite transmission in wildlife. *Biological Reviews, 92*(1), 389–409. https://doi.org/10.1111/brv.12236

Supsplementary Files

Supplementary Files for Chapter 2

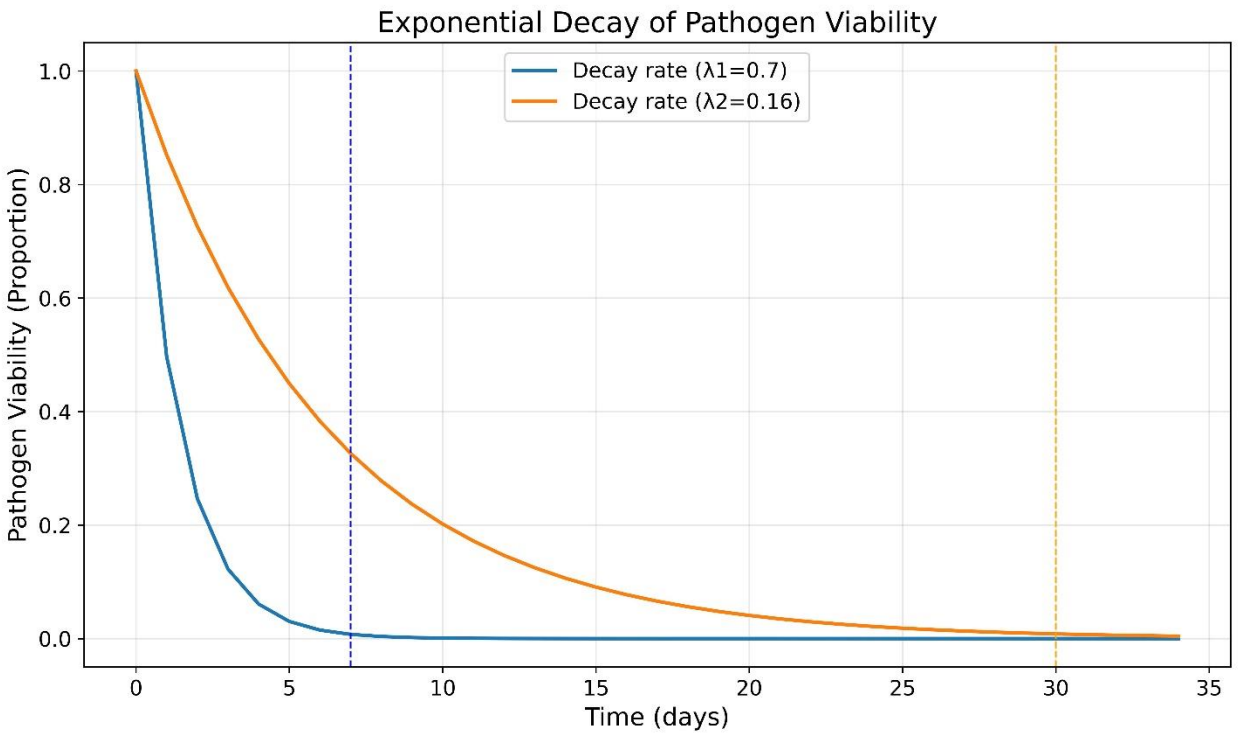


Figure S2-1. Pathogen viability with exponential decay for two survival windows (Dashed lines showing the days at which pathogen load in the environment becomes zero).

Table S2-1. Network matrices among 4 different models in 7-day and 30-day pathogen decay

	Pathogen decay model		**High-use area model**		**Behavioral model**		**Integrated model**	
Matrices	**Indirect (7 days)**	**Indirect (30 days)**	**Indirect (7 days)**	**Indirect (30 days)**	**Indirect (7 days)**	**Indirect (30 days)**	**Indirect (7 days)**	**Indirect (30 days)**
Modularity	0.50	0.44	0.52	0.46	0.42	0.37	0.42	0.37
Transitivity	0.81	0.84	0.85	0.87	0.77	0.79	0.77	0.79
Edge Density	0.60	0.65	0.58	0.63	0.50	0.55	0.50	0.54
Assortativity	0.29	0.17	0.43	0.30	0.27	0.31	0.26	0.30
Outdegree	norm (μ =9.53, σ = 3.63), x = [2,15]	norm (μ =10.35, σ = 3.58), x = [2,15]	norm (μ =9.35, σ = 3.44), x = [2, 13]	norm (μ =10.12, σ = 3.36), x = [2, 14]	norm (μ =8, σ = 3.24), x = [2, 13]	norm (μ =8.76, σ = 3.32), x = [2, 13]	norm (μ =7.94, σ = 3.28), x = [2, 13]	norm (μ =8.71, σ = 3.37), x = [2, 13]
Indegree	norm (μ =9.53, σ = 3.24), x = [2,13]	Gamma (α= 234.68, β= 0.23), x = [3,15]	norm (μ =9.35, σ = 3.53), x = [2, 13]	norm (μ =10.12, σ = 3.82), x = [3, 15]	norm (μ =8, σ = 3.33), x = [2, 13]	expon (λ = 2), x = [2,13]	norm (μ =7.94, σ = 3.40), x = [2, 13]	Gamma (α= 271.96, β= 0.21), x = [2,13]
Outstrength	Lognorm (α= 0.40, β = 1035.58) x = [49.64,1886]	Lognorm (α= 0.46, β = 2396.48) x = [148.19,5506.52]	Lognorm (α = 0.40, β = 1032.70), x = [64.92, 1921.07]	Lognorm (α =0.46, β = 2363.73), x = [212.36, 5542.75]	expon (λ = 13.85), x = [13.85, 1084.24]	Lognorm (α =0.80, β = 620.02), x = [22.09, 2722.64]	expon (λ = 13.56), x = [13.56, 1093.69]	Lognorm (α =0.79, β = 645.17), x = [21.58, 2752.60]
Instrength	norm (μ =689.81, σ = 429.105), x = [42.89,1527.19]	Lognorm (α= 0.35, β = 3256.64) x = [140.28,3946.87]	norm (μ =702.53, σ = 428.17), x = [90.78, 1541.65]	Lognorm (α = 0.43, β = 2566.81), x = [257.17, 3961.40]	expon (λ = 3.40), x = [3.40, 1031.91]	expon (λ = 15.64), x = [15.64, 3112.58]	expon (λ = 18.26), x = [18.26, 1039.28]	expon (λ = 78.52), x = [78.52, 3139.28]
Eigenvector (mean)	0.16	0.15	0.16	0.16	0.12	0.13	0.12	0.13
Closeness (mean)	0.67	0.72	0.64	0.71	0.58	0.62	0.58	0.62
Betweenness (mean)	0.17	0.16	0.14	0.12	0.16	0.17	0.15	0.15

SEIR Model Specifications

We parameterize the SEIR epidemic model to properly assess disease-specific transmission dynamics and provide realistic simulation of outbreak progression. Adequate parameterization enables the model to mimic biological processes including incubation periods,

infectious periods, and contact rates. The environmental compartment explicitly models pathogen accumulation through shedding from infected individuals, its subsequent decay in the external environment, and eventual acquisition by susceptible hosts (Brouwer et al., 2017). Parameterization reflects the distinct biological properties of the pathogens: Influenza A virus exhibits rapid decay in external environments while *Brucella suis* shows prolonged environmental stability in contaminated water and soil (Buch et al., 2023). Stochastic transmission (β), incubation (σ), recovery (γ), shedding (ξ), and environmental acquisition (α) rates capture natural infection heterogeneity at host and environmental interfaces (Table S2), according to recent empirical research (Antolin, 2008; Buch et al., 2023). Deterministic pathogen decay rates (λ) capture differential rapid versus long-term environmental survival, as a function of exposure duration, according to pathogen longevity traits. Initial infection proportion (I_0) was initialized to simulate initial outbreak seeding in controlled environments. In our simulations, we conducted 1,000 stochastic iterations over a 365-day period per run, thereby capturing natural variability and enabling a robust assessment of how environmental persistence interacts with contact network characteristics to drive transmission dynamics (Dureau et al., 2013).

Key epidemiological metrics, including R_0 and time to peak incidence, were estimated from the simulation outcomes. By integrating these metrics, along with extinction probabilities where applicable, our model provides a comprehensive framework to evaluate how variations in contact structures and environmental transmission pathways influence the spread and persistence of disease within a population.

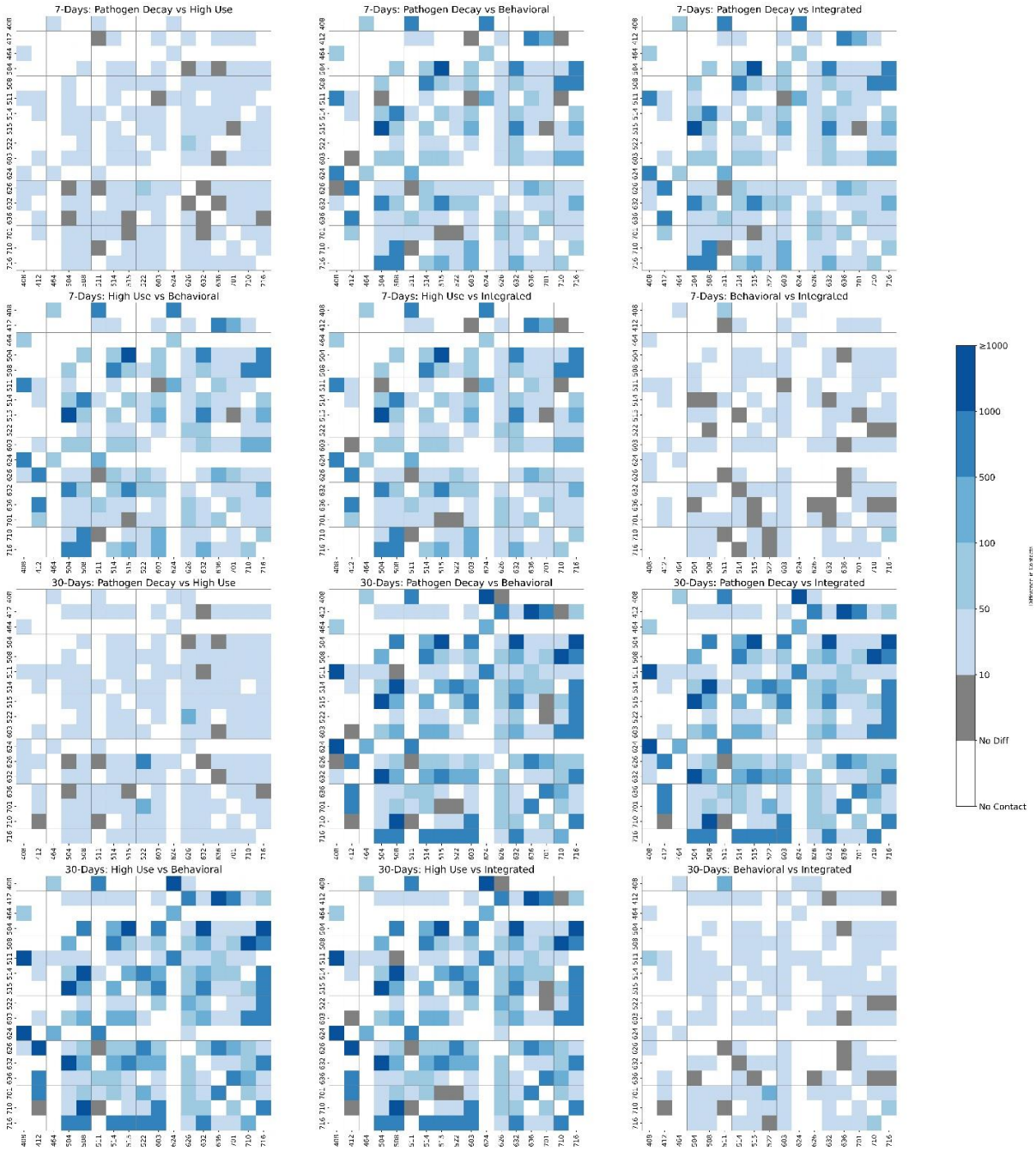


Figure S2-2: Pairwise comparison of the difference of contact between different models.

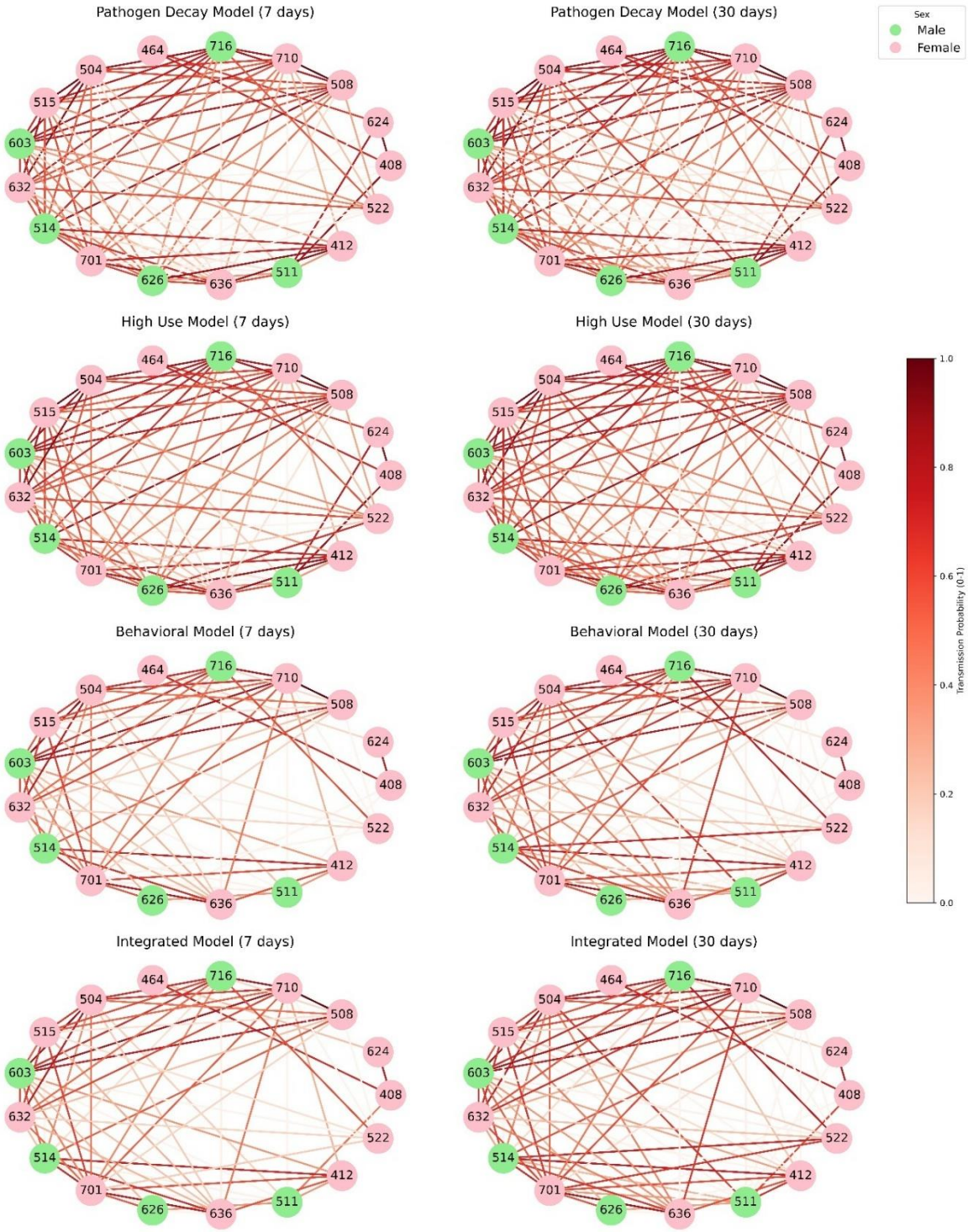


Figure S2-3: Probability of transmission network, based on the SEIR model.

Sensitivity Analysis for high-use area:

We have tested the 50 and 90 percentiles to define the high-use area and how these thresholds influence the identification of relevant environmentally-mediated contacts.

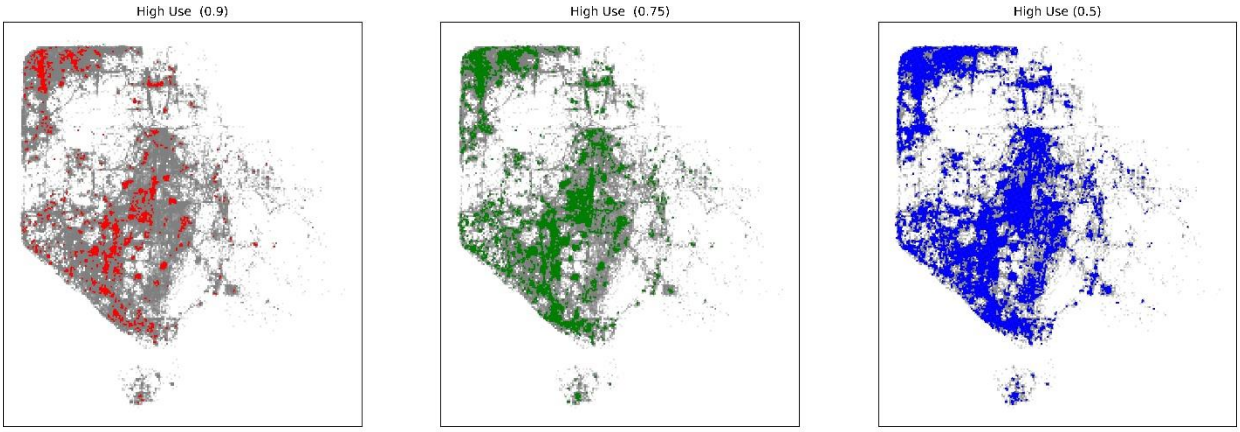


Figure S2-4: High-use areas under different definitions oh visitation rates with 50th , 75th and 90th percentile

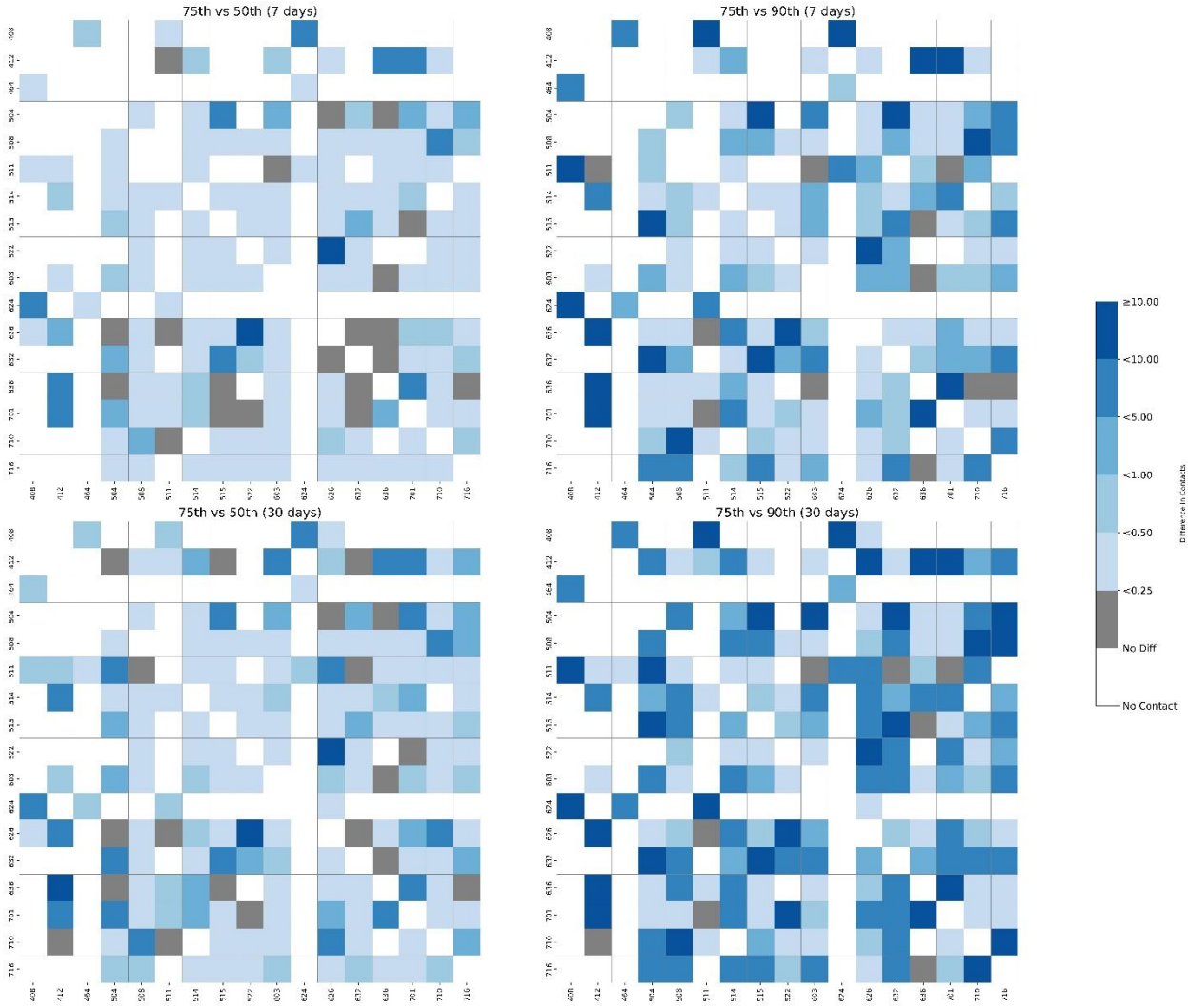


Figure S2-5: Contact kernel differences with sensitivity analysis of high-use area definition, with 50th ,75th and 90th percentile.

Hidden Markov Model for Behavioral State

In order to infer host behavioral states from GPS movement tracks, we fitted individual three-state Gaussian Hidden Markov Models (HMMs) to each of the collared pigs on the basis of standardized spatial and temporal features. Each model took as input x and y coordinates (UTM-projected) and a normalized timestamp variable, which equated to relative times scaled between 0 and 1. All of the features were z-transformed to zero mean and unit variance prior to model fitting. For each subject, we independently trained an HMM with three hidden states and a full

covariance structure using the *GaussianHMM* function in the Python package *hmmlearn*. Training was conducted using the Expectation-Maximization algorithm for a maximum of 200 iterations with a convergence tolerance. Where convergence failed due to too little data, that subject was excluded from behavioral modeling.

After convergence, the most probable series of behavioral states was decoded through the Viterbi algorithm. The emergent hidden states were a posteriori annotated as 'resting', 'foraging', and 'moving' based on their movement patterns. Specifically, we computed step length and turning angle distributions for each state and assigned behavioral labels by rank order: the state with the shortest step lengths and highest turning angle variance was labeled as 'resting'; the middle state as 'foraging'; and the state with the longest, most directional steps as 'moving'. This biologically informed partitioning adapts conventions from movement ecology literature (Pohle et al., 2017; Clontz et al., 2021) and aligns with patterns expected in feral pig behavior. Annotated states were then applied to environmentally-mediated contact events by acquisition-relevant or deposition-relevant behavior.

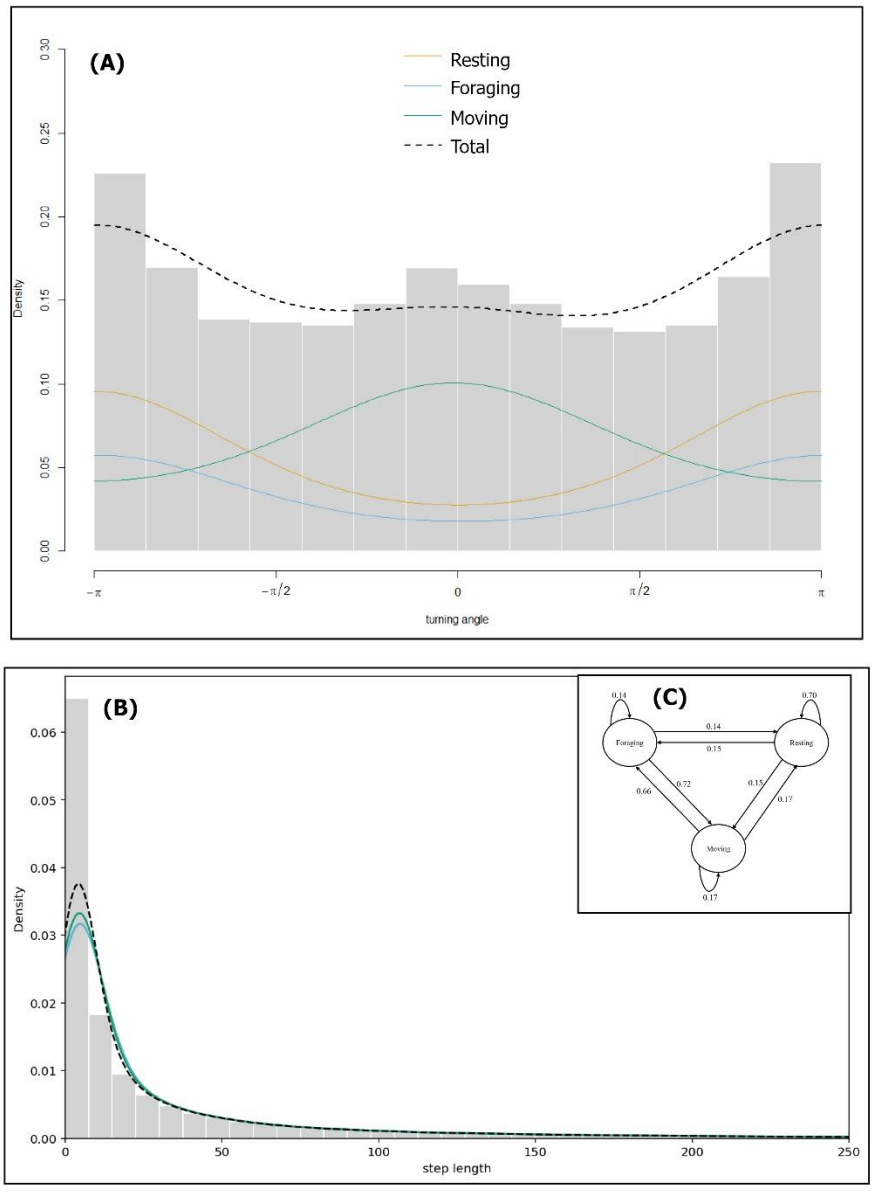


Figure S2-6: Distribution of the step length and turning angle of the HMM used to determine the behavioral states, (A) Based on turning angle, (B) Based on step length and (C) The transition probability among different states with self-transition.

Pairwise Spatial Distribution of contacts:

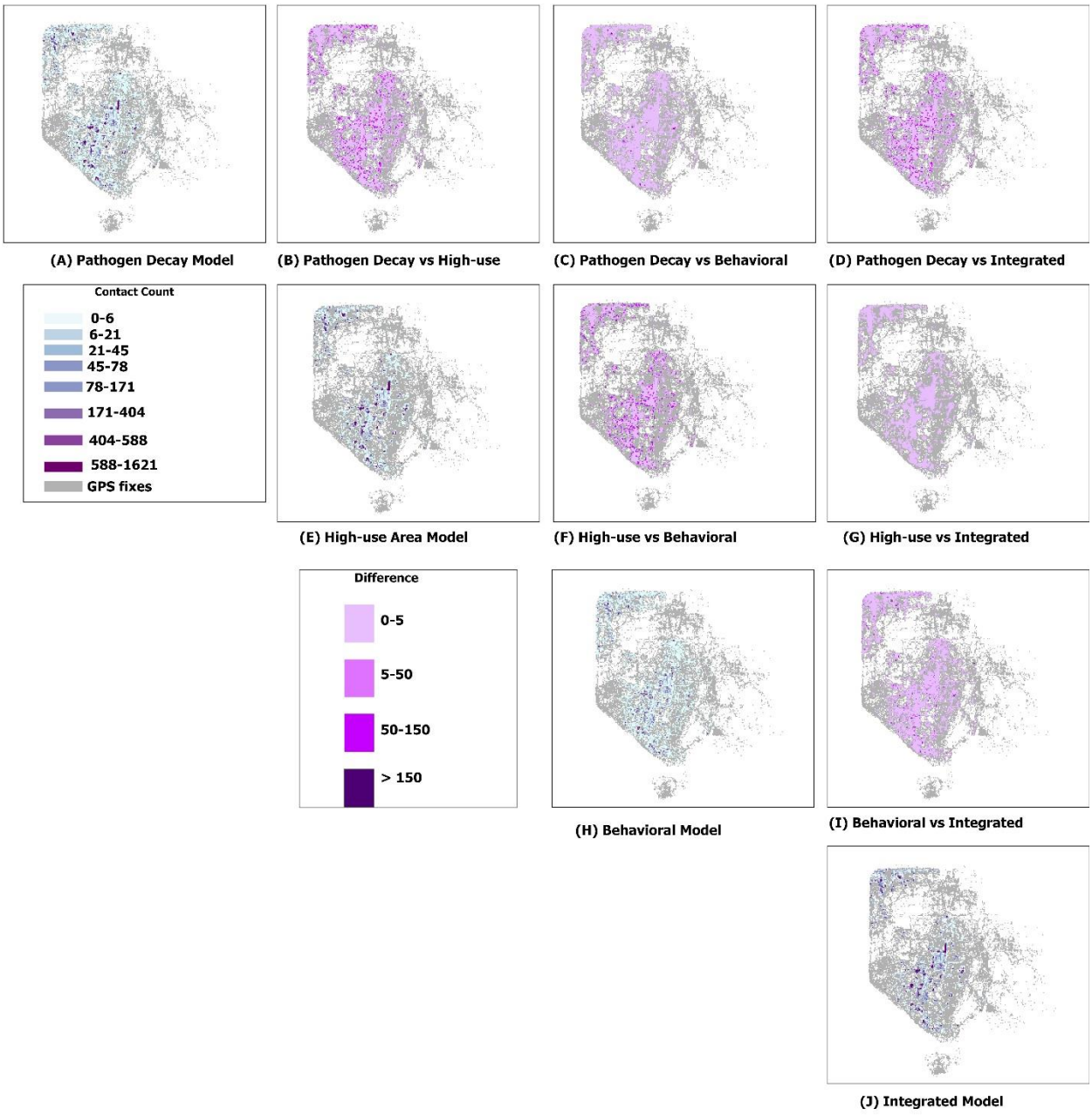


Figure S2-7: Distribution and difference of contacts among different model for 7-days pathogen decay scenario.

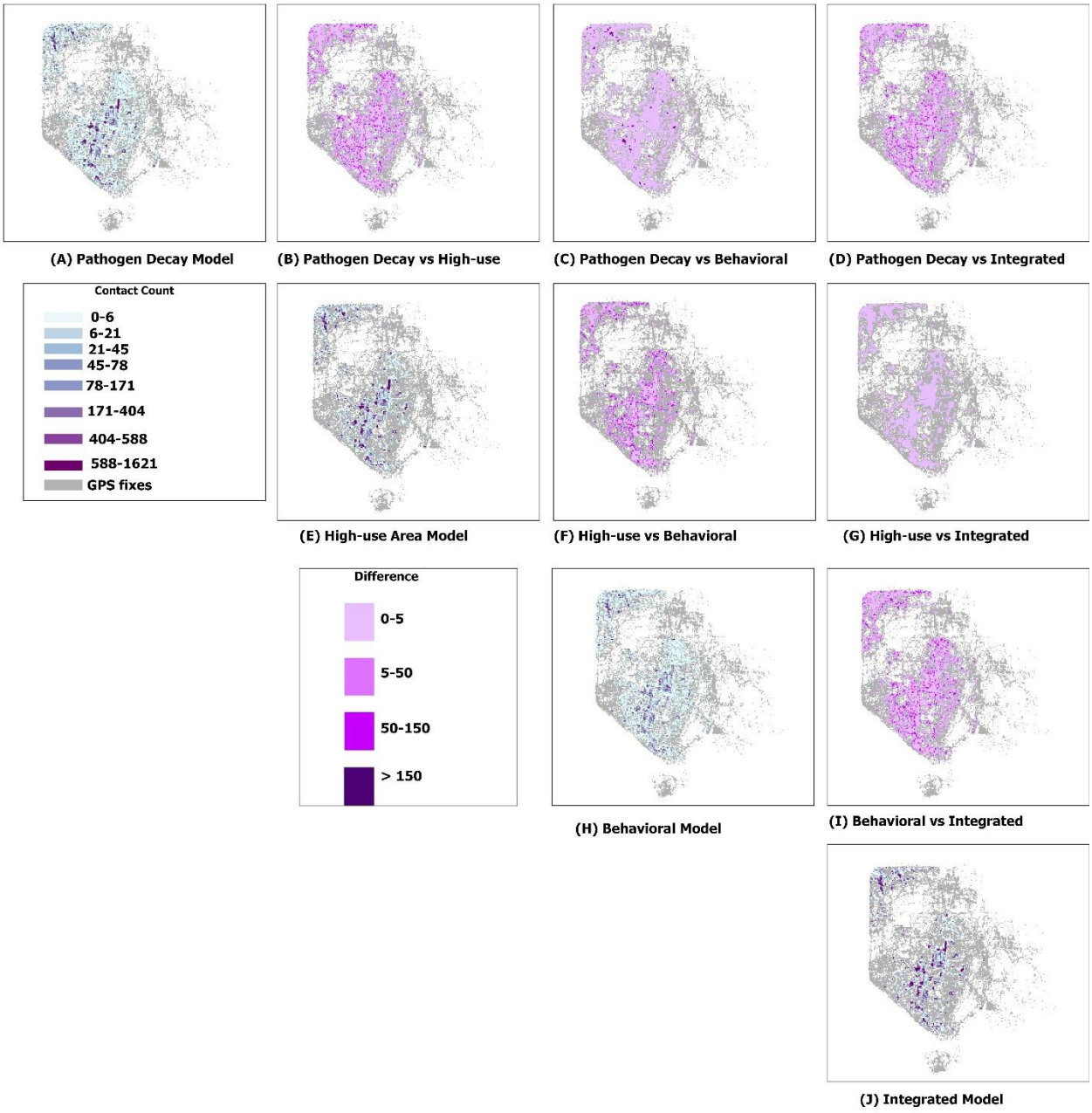


Figure S2-8: Distribution and difference of contacts among different model for 30-days pathogen decay scenario.

Supplementary files for Chapter 3

Table S3-1: Model initialization and prior specifications

Parameter	Description	Specification
z (latent state)	Initial occupancy state	Initialized as max detection across replicates
alpha.comm	Community-level detection intercept	Fixed at 0 (initial value)
beta.comm	Community-level occupancy coefficients	Fixed at 0 (initial value)
beta, alpha	Species-level coefficients	Initialized at 0
tau.sq.beta	Variance of occupancy coefficients	Inverse-Gamma (a = 2, b = 1)
tau.sq.alpha	Variance of detection coefficients	Inverse-Gamma (a = 2, b = 1)
tau.sq.theta	Variance of random effects	Inverse-Gamma (a = 2, b = 1)
sigma.sq	Spatial variance parameter	Inverse-Gamma (a = 2, b = 0.5)
phi	Spatial decay parameter	Uniform (1, 3)
beta.comm.normal	Prior for occupancy coefficients	Normal (mean = 0, var = 2.72)
alpha.comm.normal	Prior for detection coefficients	Normal (mean = −2, var = 2)

Table S3-2: Performance metrics and model selection results for all 123 candidate multi-species Bayesian hierarchical occupancy models evaluated across alternative ecological predictor sets and spatial covariance structures.

Model ID	Covariance Structure	Model Formula	R-hat	ESS	K-fold CV (deviance score)
M1	Spherical - Pig HR*	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks	1.0016	4368	3336.96
M2	Spherical - Pig HR	wetland_area_ha + lswi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_geese_swans	1.0014	5127	3342.15
M3	Spherical - Pig HR	pct_wetland + patch_density + area_mn + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + mean_diving_ducks + mean_gulls	1.0018	4875	3350.42
M4	Spherical - Pig HR	pct_wetland + edge_density + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_shorebirds	1.0015	5483	3361.87
M5	Spherical - Pig HR	ndvi + prcp_weeks + temp_sd + vp + thi + srad + sm + mean_diving_ducks + mean_geese_swans	1.0017	4639	3368.54
M6	Spherical - Pig HR	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + In_Degree + Betweenness + mean_dabbling_ducks + mean_diving_ducks	1.0013	5926	3374.19
M7	Spherical - Pig HR	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + Out_Degree + Eigenvector + Bridging_Index + mean_geese_swans + mean_shorebirds	1.0019	4412	3381.73
M8	Spherical - Pig HR	pct_wetland + lswi + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + rel_dabbling_ducks + mean_diving_ducks + mean_geese_swans	1.0016	5218	3393.48

M9	Spherical - Pig HR	wetland_area_ha + ndvi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks + mean_gulls	1.0015	4784	3402.91
M10	Spherical - Pig HR	wetland_area_ha + water_frequency + inundation_variability + prcp_spi + temp_range + rh_sd + sm + mean_dabbling_ducks + mean_pelagic_seabirds	1.0017	4547	3411.26
M11	Spherical - Bird Local HR*	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks	1.0018	6895	3344.72
M12	Spherical - Bird Local HR	wetland_area_ha + lswi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_geese_swans	1.002	7356	3357.11
M13	Spherical - Bird Local HR	pct_wetland + patch_density + area_mn + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + mean_diving_ducks + mean_gulls	1.0019	7028	3348.31
M14	Spherical - Bird Local HR	pct_wetland + edge_density + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_shorebirds	1.0021	7154	3356.48
M15	Spherical - Bird Local HR	ndvi + prcp_weeks + temp_sd + vp + thi + srad + sm + mean_diving_ducks + mean_geese_swans	1.0017	7482	3364.19
M16	Spherical - Bird Local HR	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + In_Degree + Betweenness + mean_dabbling_ducks + mean_diving_ducks	1.002	7346	3340.01
M17	Spherical - Bird Local HR	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + Out_Degree + Eigenvector + Bridging_Index + mean_geese_swans + mean_shorebirds	1.0019	7261	3381.27
M18	Spherical - Bird Local HR	pct_wetland + lswi + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + rel_dabbling_ducks + mean_diving_ducks + mean_geese_swans	1.002	6986	3392.15
M19	Spherical - Bird Local HR	wetland_area_ha + ndvi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks + mean_gulls	1.0018	7410	3401.73
M20	Spherical - Bird Local HR	wetland_area_ha + water_frequency + inundation_variability + prcp_spi + temp_range + rh_sd + sm + mean_dabbling_ducks + mean_pelagic_seabirds	1.0021	7074	3413.66
M21	Spherical – HUC*	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks	1.0007	7851	3410.84
M22	Spherical - HUC	wetland_area_ha + lswi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_geese_swans	1.0006	7924	3417.53
M23	Spherical - HUC	pct_wetland + patch_density + area_mn + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + mean_diving_ducks + mean_gulls	1.0008	7718	3423.96
M24	Spherical - HUC	pct_wetland + edge_density + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_shorebirds	1.0005	8168	3406.61
M25	Spherical - HUC	ndvi + prcp_weeks + temp_sd + vp + thi + srad + sm + mean_diving_ducks + mean_geese_swans	1.0005	8235	3430.22
M26	Spherical - HUC	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + In_Degree + Betweenness + mean_dabbling_ducks + mean_diving_ducks	1.0007	7684	3437.51
M27	Spherical - HUC	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + Out_Degree + Eigenvector + Bridging_Index + mean_geese_swans + mean_shorebirds	1.0006	8012	3445.87
M28	Spherical - HUC	pct_wetland + lswi + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + rel_dabbling_ducks + mean_diving_ducks + mean_geese_swans	1.0009	7543	3453.18
M29	Spherical - HUC	wetland_area_ha + ndvi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks + mean_gulls	1.0007	7795	3461.42
M30	Spherical - HUC	wetland_area_ha + water_frequency + inundation_variability + prcp_spi + temp_range + rh_sd + sm + mean_dabbling_ducks + mean_pelagic_seabirds	1.0008	7629	3470.35
M31	Exponential - 5 Neighbors	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks	1.0009	7325	3418.63
M32	Exponential - 5 Neighbors	wetland_area_ha + lswi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_geese_swans	1.0008	7461	3425.17
M33	Exponential - 5 Neighbors	pct_wetland + patch_density + area_mn + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + mean_diving_ducks + mean_gulls	1.001	7094	3433.82
M34	Exponential - 5 Neighbors	pct_wetland + edge_density + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_shorebirds	1.0009	7218	3440.95

M35	Exponential - 5 Neighbors	ndvi + prcp_weeks + temp_sd + vp + thi + srad + sm + mean_diving_ducks + mean_geese_swans	1.0008	7540	3448.26
M36	Exponential - 5 Neighbors	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + In_Degree + Betweenness + mean_dabbling_ducks + mean_diving_ducks	1.0008	7512	3412.84
M37	Exponential - 5 Neighbors	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + Out_Degree + Eigenvector + Bridging_Index + mean_geese_swans + mean_shorebirds	1.0007	7786	3409.98
M38	Exponential - 5 Neighbors	pct_wetland + lswi + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + rel_dabbling_ducks + mean_diving_ducks + mean_geese_swans	1.0011	6937	3457.83
M39	Exponential - 5 Neighbors	wetland_area_ha + ndvi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks + mean_gulls	1.0009	7176	3466.12
M40	Exponential - 5 Neighbors	wetland_area_ha + water_frequency + inundation_variability + prcp_spi + temp_range + rh_sd + sm + mean_dabbling_ducks + mean_pelagic_seabirds	1.001	7018	3475.48
M41	Exponential - 10 Neighbors	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks	1.0004	5562	3411.92
M42	Exponential - 10 Neighbors	wetland_area_ha + lswi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_geese_swans	1.0005	5384	3418.57
M43	Exponential - 10 Neighbors	pct_wetland + patch_density + area_mn + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + mean_diving_ducks + mean_gulls	1.0004	5627	3426.84
M44	Exponential - 10 Neighbors	pct_wetland + edge_density + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_shorebirds	1.0006	5218	3432.95
M45	Exponential - 10 Neighbors	ndvi + prcp_weeks + temp_sd + vp + thi + srad + sm + mean_diving_ducks + mean_geese_swans	1.0005	5481	3440.17
M46	Exponential - 10 Neighbors	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + In_Degree + Betweenness + mean_dabbling_ducks + mean_diving_ducks	1.0004	5664	3448.89
M47	Exponential - 10 Neighbors	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + Out_Degree + Eigenvector + Bridging_Index + mean_geese_swans + mean_shorebirds	1.0007	4976	3456.31
M48	Exponential - 10 Neighbors	pct_wetland + lswi + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + rel_dabbling_ducks + mean_diving_ducks + mean_geese_swans	1.0003	5745	3409.36
M49	Exponential - 10 Neighbors	wetland_area_ha + ndvi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks + mean_gulls	1.0005	5294	3465.42
M50	Exponential - 10 Neighbors	wetland_area_ha + water_frequency + inundation_variability + prcp_spi + temp_range + rh_sd + sm + mean_dabbling_ducks + mean_pelagic_seabirds	1.0006	5138	3474.76
M51	Exponential - 15 Neighbors	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks	1.001	7268	3322.11
M52	Exponential - 15 Neighbors	wetland_area_ha + lswi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_geese_swans	1.0012	6894	3328.54
M53	Exponential - 15 Neighbors	pct_wetland + patch_density + area_mn + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + mean_diving_ducks + mean_gulls	1.0013	6748	3335.42
M54	Exponential - 15 Neighbors	pct_wetland + edge_density + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_shorebirds	1.001	7316	3341.95
M55	Exponential - 15 Neighbors	ndvi + prcp_weeks + temp_sd + vp + thi + srad + sm + mean_diving_ducks + mean_geese_swans	1.0012	6985	3348.27
M56	Exponential - 15 Neighbors	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + In_Degree + Betweenness + mean_dabbling_ducks + mean_diving_ducks	1.0011	7152	3356.73
M57	Exponential - 15 Neighbors	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + Out_Degree + Eigenvector + Bridging_Index + mean_geese_swans + mean_shorebirds	1.0014	6621	3364.48
M58	Exponential - 15 Neighbors	pct_wetland + lswi + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + rel_dabbling_ducks + mean_diving_ducks + mean_geese_swans	1.0012	6847	3372.36
M59	Exponential - 15 Neighbors	wetland_area_ha + ndvi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks + mean_gulls	1.0011	7077	3318.69

M60	Exponential - 15 Neighbors	wetland_area_ha + water_frequency + inundation_variability + prcp_spi + temp_range + rh_sd + sm + mean_dabbling_ducks + mean_pelagic_seabirds	1.0013	6715	3380.91
M61	Matern	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks	1.0013	7412	3420.16
M62	Matern	wetland_area_ha + lswi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_geese_swans	1.0015	7084	3426.73
M63	Matern	pct_wetland + patch_density + area_mn + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + mean_diving_ducks + mean_gulls	1.0016	6917	3434.82
M64	Matern	pct_wetland + edge_density + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_shorebirds	1.0014	7231	3416.48
M65	Matern	ndvi + prcp_weeks + temp_sd + vp + thi + srad + sm + mean_diving_ducks + mean_geese_swans	1.0013	7568	3441.35
M66	Matern	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + In_Degree + Betweenness + mean_dabbling_ducks + mean_diving_ducks	1.0015	7146	3448.94
M67	Matern	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + Out_Degree + Eigenvector + Bridging_Index + mean_geese_swans + mean_shorebirds	1.0014	7298	3457.12
M68	Matern	pct_wetland + lswi + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + rel_dabbling_ducks + mean_diving_ducks + mean_geese_swans	1.0017	6724	3465.88
M69	Matern	wetland_area_ha + ndvi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks + mean_gulls	1.0015	7012	3474.61
M70	Matern	wetland_area_ha + water_frequency + inundation_variability + prcp_spi + temp_range + rh_sd + sm + mean_dabbling_ducks + mean_pelagic_seabirds	1.0016	6859	3483.27
M71	Gaussian - 5 Neighbors	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks	1.0011	7642	3428.16
M72	Gaussian - 5 Neighbors	wetland_area_ha + lswi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_geese_swans	1.0012	7418	3434.52
M73	Gaussian - 5 Neighbors	pct_wetland + patch_density + area_mn + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + mean_diving_ducks + mean_gulls	1.0011	7584	3442.37
M74	Gaussian - 5 Neighbors	pct_wetland + edge_density + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_shorebirds	1.0013	7261	3451.86
M75	Gaussian - 5 Neighbors	ndvi + prcp_weeks + temp_sd + vp + thi + srad + sm + mean_diving_ducks + mean_geese_swans	1.0012	7447	3460.18
M76	Gaussian - 5 Neighbors	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + In_Degree + Betweenness + mean_dabbling_ducks + mean_diving_ducks	1.0006	7865	3423.97
M77	Gaussian - 5 Neighbors	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + Out_Degree + Eigenvector + Bridging_Index + mean_geese_swans + mean_shorebirds	1.0011	7695	3468.43
M78	Gaussian - 5 Neighbors	pct_wetland + lswi + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + rel_dabbling_ducks + mean_diving_ducks + mean_geese_swans	1.0014	7016	3478.72
M79	Gaussian - 5 Neighbors	wetland_area_ha + ndvi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks + mean_gulls	1.0012	7314	3488.55
M80	Gaussian - 5 Neighbors	wetland_area_ha + water_frequency + inundation_variability + prcp_spi + temp_range + rh_sd + sm + mean_dabbling_ducks + mean_pelagic_seabirds	1.0013	7157	3497.21
M81	Gaussian - 10 Neighbors	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks	1.0011	7825	3425.41
M82	Gaussian - 10 Neighbors	wetland_area_ha + lswi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_geese_swans	1.0012	7568	3432.73
M83	Gaussian - 10 Neighbors	pct_wetland + patch_density + area_mn + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + mean_diving_ducks + mean_gulls	1.001	8144	3439.88
M84	Gaussian - 10 Neighbors	pct_wetland + edge_density + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_shorebirds	1.0013	7284	3447.16
M85	Gaussian - 10 Neighbors	ndvi + prcp_weeks + temp_sd + vp + thi + srad + sm + mean_diving_ducks + mean_geese_swans	1.0011	7916	3455.82

M86	Gaussian - 10 Neighbors	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + In_Degree + Betweenness + mean_dabbling_ducks + mean_diving_ducks	1.0008	8073	3419.65
M87	Gaussian - 10 Neighbors	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + Out_Degree + Eigenvector + Bridging_Index + mean_geese_swans + mean_shorebirds	1.0012	7631	3464.27
M88	Gaussian - 10 Neighbors	pct_wetland + lswi + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + rel_dabbling_ducks + mean_diving_ducks + mean_geese_swans	1.0014	7127	3473.91
M89	Gaussian - 10 Neighbors	wetland_area_ha + ndvi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks + mean_gulls	1.0012	7488	3482.36
M90	Gaussian - 10 Neighbors	wetland_area_ha + water_frequency + inundation_variability + prcp_spi + temp_range + rh_sd + sm + mean_dabbling_ducks + mean_pelagic_seabirds	1.0013	7269	3491.74
M91	Gaussian - 15 Neighbors	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks	1.0011	7742	3428.28
M92	Gaussian - 15 Neighbors	wetland_area_ha + lswi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_geese_swans	1.0012	7487	3433.94
M93	Gaussian - 15 Neighbors	pct_wetland + patch_density + area_mn + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + mean_diving_ducks + mean_gulls	1.0011	7684	3441.52
M94	Gaussian - 15 Neighbors	pct_wetland + edge_density + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_shorebirds	1.0013	7348	3449.76
M95	Gaussian - 15 Neighbors	ndvi + prcp_weeks + temp_sd + vp + thi + srad + sm + mean_diving_ducks + mean_geese_swans	1.0012	7563	3458.81
M96	Gaussian - 15 Neighbors	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + In_Degree + Betweenness + mean_dabbling_ducks + mean_diving_ducks	1.0005	7965	3421.73
M97	Gaussian - 15 Neighbors	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + Out_Degree + Eigenvector + Bridging_Index + mean_geese_swans + mean_shorebirds	1.0011	7816	3467.24
M98	Gaussian - 15 Neighbors	pct_wetland + lswi + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + rel_dabbling_ducks + mean_diving_ducks + mean_geese_swans	1.0014	7095	3476.88
M99	Gaussian - 15 Neighbors	wetland_area_ha + ndvi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks + mean_gulls	1.0012	7412	3485.63
M100	Gaussian - 15 Neighbors	wetland_area_ha + water_frequency + inundation_variability + prcp_spi + temp_range + rh_sd + sm + mean_dabbling_ducks + mean_pelagic_seabirds	1.0013	7241	3495.07
M101	Gaussian - Full GP	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks	1.0011	7624	3438.54
M102	Gaussian - Full GP	wetland_area_ha + lswi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_geese_swans	1.0012	7398	3445.12
M103	Gaussian - Full GP	pct_wetland + patch_density + area_mn + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + mean_diving_ducks + mean_gulls	1.0008	7868	3433.27
M104	Gaussian - Full GP	pct_wetland + edge_density + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_shorebirds	1.0013	7217	3453.69
M105	Gaussian - Full GP	ndvi + prcp_weeks + temp_sd + vp + thi + srad + sm + mean_diving_ducks + mean_geese_swans	1.0011	7544	3461.74
M106	Gaussian - Full GP	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + In_Degree + Betweenness + mean_dabbling_ducks + mean_diving_ducks	1.0012	7356	3470.58
M107	Gaussian - Full GP	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + Out_Degree + Eigenvector + Bridging_Index + mean_geese_swans + mean_shorebirds	1.0013	7183	3479.83
M108	Gaussian - Full GP	pct_wetland + lswi + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + rel_dabbling_ducks + mean_diving_ducks + mean_geese_swans	1.0014	6841	3489.16
M109	Gaussian - Full GP	wetland_area_ha + ndvi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks + mean_gulls	1.0012	7287	3498.42
M110	Gaussian - Full GP	wetland_area_ha + water_frequency + inundation_variability + prcp_spi + temp_range + rh_sd + sm + mean_dabbling_ducks + mean_pelagic_seabirds	1.0015	6698	3508.95

M111	Non-spatial	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks	1.0002	7967	3303.23
M112	Non-spatial	wetland_area_ha + lswi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_geese_swans	1.0002	8033	3304.04
M113	Non-spatial	pct_wetland + patch_density + area_mn + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + mean_diving_ducks + mean_gulls	1.0004	7825	3305.91
M114	Non-spatial	pct_wetland + edge_density + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_shorebirds	1.0003	8118	3301.32
M115	Non-spatial	ndvi + prcp_weeks + temp_sd + vp + thi + srad + sm + mean_diving_ducks + mean_geese_swans	1.0005	7684	3311.48
M116	Non-spatial	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + In_Degree + Betweenness + mean_dabbling_ducks + mean_diving_ducks	1.0007	7215	3318.72
M117	Non-spatial	wetland_area_ha + ndvi + prcp_spi + temp_range + rh_sd + Flow_Centrality + Out_Degree + Eigenvector + Bridging_Index + mean_geese_swans + mean_shorebirds	1.0006	7481	3324.36
M118	Non-spatial	pct_wetland + lswi + prcp_spi + temp_range + rh_sd + sm + Flow_Centrality + rel_dabbling_ducks + mean_diving_ducks + mean_geese_swans	1.0005	7864	3330.91
M119	Non-spatial	wetland_area_ha + ndvi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks + mean_gulls	1.0006	7346	3307.27
M120	Non-spatial	wetland_area_ha + water_frequency + inundation_variability + prcp_spi + temp_range + rh_sd + sm + mean_dabbling_ducks + mean_pelagic_seabirds	1.0004	7742	3338.27
M121	Non-spatial with log-pig positive count**	log_neighbour_pig_cases + wetland_area_ha + ndvi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks + mean_gulls	0.9999	8118	3301.32
M122	Non-spatial with log-bird positive count**	log_neighbour_bird_cases + wetland_area_ha + ndvi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks + mean_gulls	1.0002	8033	3304.04
M123	Non-spatial with log positive count (bird and pig)**	log_neighbour_positive_cases +wetland_area_ha + ndvi + prcp_weeks + temp_range + vp + sm + Flow_Centrality + mean_dabbling_ducks + mean_diving_ducks + mean_gulls	1.0002	7967	3303.23

**Pig HR* indicates that the cutoff distance for the spherical covariance function was defined using the average feral pig home-range radius. *Bird Local HR* indicates that the cutoff distance was defined using the average local movement radius of wild birds. *HUC* indicates that the cutoff distance was defined using the mean spatial extent (diameter) of the HUC-8 watershed. Under the spherical covariance model, spatial correlation declines with distance and becomes exactly zero beyond the specified cutoff distance.

**Non-spatial with log-pig positive count includes the natural log-transformed and standardized number of avian influenza-positive feral pig detections within first-degree neighboring HUC-8 watersheds as an additional predictor. Non-spatial with log-bird positive count includes the natural log-transformed and standardized number of avian influenza-positive wild bird detections within first-degree neighboring HUC-8 watersheds as an additional predictor. Non-spatial with log positive count (bird and pig) includes the natural log-transformed and standardized combined number of avian influenza-positive detections from both feral pigs and wild birds within first-degree neighboring HUC-8 watersheds as an additional predictor. First-degree neighbors were defined as HUC-8 watersheds sharing a common boundary with the focal watershed.

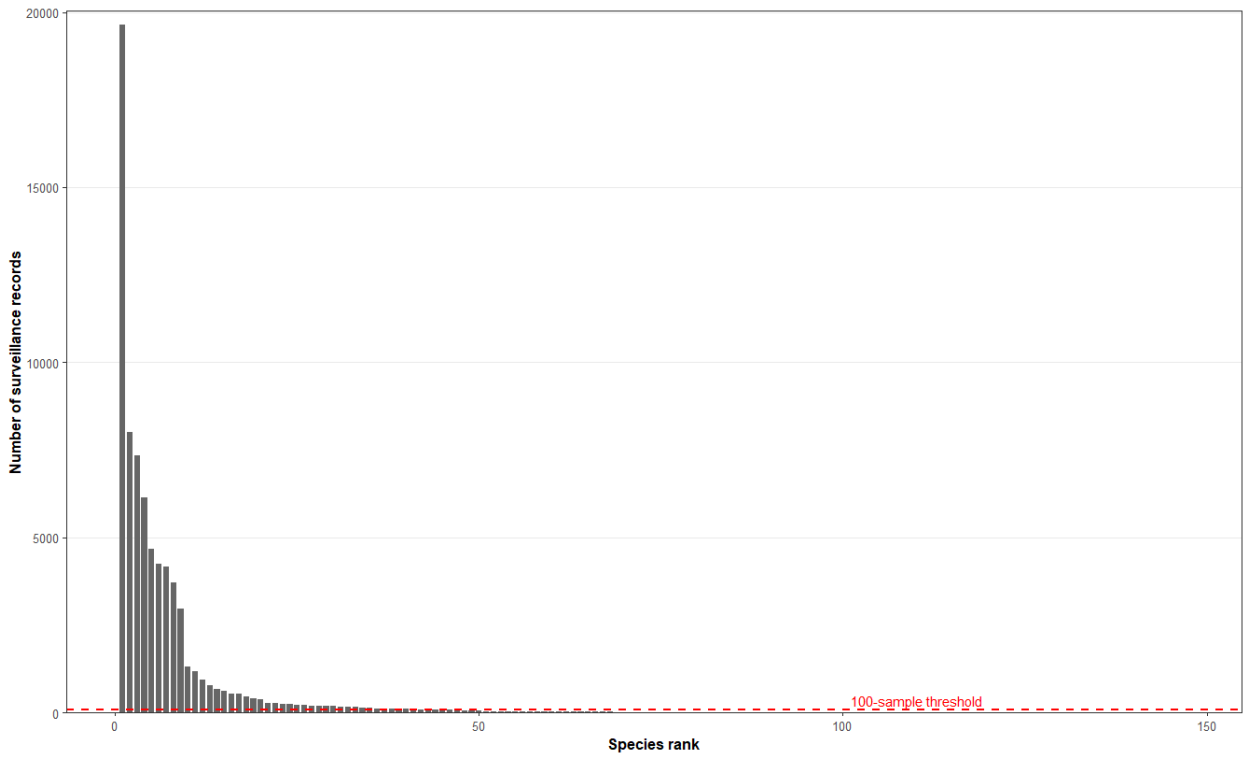


Figure S3-1: Distribution of surveillance sample sizes across wild bird species included in the avian influenza dataset. The dashed red line indicates the minimum sample size threshold (100 samples) used for species inclusion in occupancy modeling analyses.