OPEN

Waterfowl occurrence and residence time as indicators of H5 and H7 avian influenza in North American Poultry

John M. Humphreys [1,7*], Andrew M. Ramey[2], David C. Douglas[2], Jennifer M. Mullinax[3], Catherine Soos [4], Paul Link[5], Patrick Walther[6] & Diann J. Prosser[7]

Avian influenza (AI) affects wild aquatic birds and poses hazards to human health, food security, and wildlife conservation globally. Accordingly, there is a recognized need for new methods and tools to help quantify the dynamic interaction between wild bird hosts and commercial poultry. Using satellite-marked waterfowl, we applied Bayesian joint hierarchical modeling to concurrently model species distributions, residency times, migration timing, and disease occurrence probability under an integrated animal movement and disease distribution modeling framework. Our results indicate that migratory waterfowl are positively related to AI occurrence over North America such that as waterfowl occurrence probability or residence time increase at a given location, so too does the chance of a commercial poultry AI outbreak. Analyses also suggest that AI occurrence probability is greatest during our observed waterfowl northward migration, and less during the southward migration. Methodologically, we found that when modeling disparate facets of disease systems at the wildlife-agriculture interface, it is essential that multiscale spatial patterns be addressed to avoid mistakenly inferring a disease process or disease-environment relationship from a pattern evaluated at the improper spatial scale. The study offers important insights into migratory waterfowl ecology and AI disease dynamics that aid in better preparing for future outbreaks.

Avian influenza is a global concern and poses hazards to human health, food security, and wildlife conservation worldwide[1,2]. Domestic poultry operations are particularly vulnerable to avian influenza viruses maintained in wild bird hosts[3–6] as the viruses may be spread to poultry[7–10] via direct contact or by way of environmental contamination. Once introduced into a poultry operation, avian influenza viruses of the H5 and H7 hemagglutinin subtype can rapidly propagate through commercial flocks and mutate to the Highly Pathogenic Avian Influenza (HPAI) pathotype with increased virulence[1,11,12]. HPAI outbreaks can inflict direct stock mortality or necessitate that culling protocols be implemented to minimize the risk of disease spread. In rare instances, control efforts may also reduce the spread of viruses potentially lethal to humans, as was shown for Goose Guangdong (GsGD) lineage HPAI H5N1 and H7N9 in China[13,14].

Avian influenza viruses circulate among wild aquatic birds globally and taxa such as migratory waterfowl are considered to be natural biologic reservoirs[15]. Waterfowl are infected with varying virus subtypes, including those with the H5 and H7 hemagglutinin protein that have the potential to mutate to HPAI in poultry throughout the Neotropics and Nearctic[16–19].

During 2014 and 2015, it appears migratory waterfowl contributed to the introduction of GsGD lineage HPAI H5 viruses into North America[20–22] and subsequent spread in Canada and the U.S.[23–26]. This outbreak was the largest in U.S. history, affected wild and domestic birds in 21 U.S. States, resulted in the loss of approximately 50 million poultry, and caused estimated losses of more than 3 billion US dollars[27,28].

[1]Michigan State University, East Lansing, Michigan, USA. [2]U.S. Geological Survey, Alaska Science Center, Anchorage, Alaska, USA. [3]University of Maryland, College Park, Maryland, USA. [4]Environment and Climate Change Canada, Ecotoxicology and Wildlife Health Division, Saskatchewan, Canada. [5]Louisiana Department of Wildlife and Fisheries, Baton Rouge, Louisiana, USA. [6]U.S. Fish and Wildlife Service, Texas Chenier Plain Refuge Complex, Anahuac, Texas, USA. [7]U.S. Geological Survey, Patuxent Wildlife Research Center, Laurel, Maryland, USA. *email: jmh09r@my.fsu.edu

To assess the potential role of migratory waterfowl in transporting HPAI viruses throughout central and eastern Asia, several studies have examined the spatial and temporal overlap of waterfowl movement trajectories or utilization distributions with disease occurrence locations, poultry facilities, and key waterfowl habitats[3,29–33], however, there are a lack of comparable waterfowl movement studies investigating associations with avian influenza outbreaks in the Western Hemisphere. Indeed, there is a recognized need for new methods and tools to help quantify the dynamic interaction between wild bird hosts and commercial poultry in North America[34–36]. To help address this knowledge gap, we developed a combined animal movement and disease distribution model to evaluate the spatiotemporal relationship of migratory waterfowl to H5 and H7 subtype avian influenza detections in commercial poultry (AIPs) within North America.

For this study, we define AIPs as being any documented outbreak or infection caused by H5 or H7 subtype avian influenza in commercial poultry. Our central objective was to leverage multi-year animal movement telemetry and poultry disease data to characterize intra-annual patterns within a complex disease system. To accomplish this task, we used satellite-marked blue-winged teal (*Anas discors*) as a representative sample of the larger North American dabbling duck population and applied Bayesian joint hierarchical modeling to concurrently model species distributions, residency times, migration timing, and AIP occurrence probability. Blue-winged teal were chosen as a representative waterfowl species due to their widespread distribution and suspected role in redistributing avian influenza viruses throughout breeding grounds in the U.S. Northern Great Plains and Canada and overwintering habitats in Mexico[17], the Caribbean[37], and Central and northern South America[16,18,38]. Because migratory behavior and habitat preferences can differ between duck species[39], the selection of blue-winged teal as an archetypal dabbling duck species necessitates that several important caveats be considered. As examples, blue-winged teal have a tendency to start the spring migration later, undertake the fall migration earlier, and fly greater distances than other dabbling ducks[40,41]. These caveats are examined in the context of results' interpretation in the closing discussion.

Essential to our proposed method was taking steps to ensure that spatial correlation between satellite-marked birds and AIP locations did not bias model results and that the inclusion of locally-abundant (clustered) AIP records did not inflate model estimates or predictions. AIP records that are aggregated in time or space may signify non-typical increases in disease incidence that offer special insight into the system's ecology. Recognizing that disease clusters can be difficult to interpret and statistically problematic, we took steps to avoid mistakenly inferring a disease process or disease-environment relationship from a pattern evaluated at the improper spatial scale[42,43]. In the case of poultry AIPs, localized clusters may be borne of epidemiological mechanisms operating at a much finer spatial scale than those patterns arising from long-range viral dispersal by migratory birds[44]. For instance, failures in biosecurity may result in the unintentional transmission of viral pathogens among farms via contaminated farm equipment long after initial introduction by a wild bird[26,28,45]. To better differentiate between possible farm-to-farm transmission and novel wild bird introduction, our model was designed to accommodate multi-scaled spatial processes.

Materials and Methods

Animal protocols statement.
The authors confirm that all methods were carried out in accordance with relevant guidelines and regulations and that all experimental protocols were approved by the appropriate institutions and licensing committees. Protocols for capturing, handling, and instrumenting blue-winged teal included in this study were reviewed by institutional animal care and use committees (USGS Alaska Science Center Animal Care and Use Committee Animal Use permit no. 2014–5 and the University Committee on Animal Care and Supply University of Saskatchewan Animal Use protocol no. 20070039) and carried out under federal authority (US Department of the Interior Federal Bird Banding permit nos. 09072 and 23792, Federal Fish and Wildlife permit no. MB779238-2, Environment Canada Migratory Bird Banding permit no. 10458R).

Study domain and bird telemetry.
Our study area encompassed North America between 14.5° and 60.0° North Latitude and −137.5° and −66.2° West Longitude. This area includes the conterminous United States, Mexico, Cuba, the Bahamas, and major portions of Canada. Forty-two adult, male blue-winged teal were captured and marked. Twelve birds were captured on their summer range in the Saskatchewan and Alberta Canadian Provinces during August 11–15, 2013, with the remaining thirty birds captured at wintering areas along the Texas and Louisiana Gulf Coast in the United States during March 17–22, 2015. All birds were marked with 9.5 gram, solar-powered Platform Transmitter Terminals (PTTs) manufactured by Microwave Telemetry, Inc., Columbia, Maryland (USA) and then released in close proximity to their respective capture locations. The PTTs were attached using a Teflon harness secured over the sternum to have a dorsally extended antenna as outlined by Takekawa *et al.*[46]. The PTTs were programmed to have a regular duty cycle with a 10 hour on period every 48 hours. We used the Argos Data Collection and Location System (CLS America, Incorporated, Largo, Maryland, USA) to receive transmissions, which provided Doppler-based location estimates, a movement activity indicator, and a class index reporting the quality of each location estimate. To prepare telemetry for statistical modeling, we first applied a filtering algorithm[47] and then manually inspected movements and activity sensor records to exclude dates following likely bird mortality or PTT detachment. Locations exhibiting an unchanged activity indicator over two or more duty cycles were removed from the telemetry dataset. Table 1 details the bird-specific number of locations retained for analysis following data cleaning. Animations depicting the wild bird's used in this study and downloadable telemetry data are available from the U.S. Geological Survey, Alaska Science Center, Wildlife Tracking Data Collection (https://doi.org/10.5066/P9Z9BA9F)[48].

Environmental data.
To characterize habitat and environmental conditions over the study area, we utilized 250 meter resolution gridded soils data[49], the Global Topographic 30 Arc-Second Digital Elevation Model available from the U.S. Geological Survey (USGS) Earth Resources Observation and Science Center (https://www.usgs.

Tag	Deployment	Last	Mortality	Duration	Median	Sample	Latitude	Longitude	Region
131009	8/13/2013	10/26/2013	Hunting	74	29	558	52.04	−107.10	Saskatchewan
131010	8/13/2013	3/26/2014	—	225	40	1142	52.04	−107.10	Saskatchewan
131011	8/14/2013	4/27/2015	—	620	48	2730	50.48	−112.12	Alberta
131012	8/14/2013	11/5/2013	—	83	41	464	50.48	−112.12	Alberta
131013	8/14/2013	10/24/2013	Hunting	71	37	453	52.04	−107.10	Saskatchewan
131014	8/11/2013	10/24/2013	—	73	31	523	52.04	−107.10	Saskatchewan
131015	8/12/2013	4/8/2014	—	239	43	1149	52.04	−107.10	Saskatchewan
131016	8/12/2013	1/25/2014	—	165	42	803	52.04	−107.10	Saskatchewan
131017	8/12/2013	10/21/2013	—	70	32	516	52.04	−107.10	Saskatchewan
131018	8/12/2013	10/17/2013	—	66	32	468	52.04	−107.10	Saskatchewan
131019	8/15/2013	11/29/2013	—	106	42	593	50.48	−112.12	Alberta
131020	8/15/2013	10/17/2013	—	63	39	408	50.48	−112.12	Alberta
131009	3/17/2015	10/25/2015	—	222	46	1087	30.01	−94.14	Texas
131013	3/19/2015	9/26/2015	—	191	47	898	29.61	−94.53	Texas
135841	3/17/2015	12/9/2015	Hunting	267	49	1165	30.01	−94.14	Texas
135842	3/17/2015	11/12/2015	—	240	57	816	30.01	−94.14	Texas
135843	3/17/2015	10/31/2015	—	228	47	1055	30.01	−94.14	Texas
135844	3/17/2015	5/28/2015	—	72	49	335	30.01	−94.14	Texas
135845	3/18/2015	4/28/2015	—	41	52	181	29.67	−94.44	Texas
135846	3/18/2015	10/2/2015	Hunting	198	43	1061	29.67	−94.44	Texas
135847	3/18/2015	8/17/2015	—	152	56	565	29.67	−94.44	Texas
135848	3/18/2015	10/27/2015	—	223	46	1051	29.67	−94.44	Texas
135849	3/18/2015	9/27/2016	—	559	53	2129	29.67	−94.44	Texas
135850	3/19/2015	3/31/2015	—	12	53	57	29.61	−94.53	Texas
135851	3/19/2015	5/11/2015	—	53	54	205	29.61	−94.53	Texas
135852	3/19/2015	9/30/2016	—	561	56	1918	29.61	−94.53	Texas
135853	3/19/2015	5/5/2015	—	47	53	211	29.61	−94.53	Texas
135854	3/19/2015	10/18/2015	—	213	48	938	29.61	−94.53	Texas
135855	3/22/2015	10/21/2015	—	213	48	955	30.55	−91.85	Louisiana
135856	3/22/2015	7/12/2015	—	112	48	532	30.55	−91.85	Louisiana
135857	3/22/2015	7/22/2015	—	122	51	519	30.55	−91.85	Louisiana
135858	3/22/2015	5/8/2015	—	47	47	226	30.55	−91.85	Louisiana
135859	3/22/2015	7/17/2015	—	117	46	569	30.55	−91.85	Louisiana
135860	3/22/2015	12/23/2016	—	642	50	2647	30.55	−91.85	Louisiana
135861	3/22/2015	4/13/2016	—	388	50	1338	30.55	−91.85	Louisiana
135862	3/22/2015	8/14/2015	—	145	55	528	30.55	−91.85	Louisiana
135863	3/22/2015	4/17/2015	—	26	50	125	30.55	−91.85	Louisiana
135864	3/21/2015	10/28/2015	Hunting	221	48	1019	30.46	−91.57	Louisiana
135865	3/21/2015	4/7/2015	—	17	51	89	30.46	−91.57	Louisiana
135866	3/21/2015	5/31/2015	—	71	53	290	30.46	−91.57	Louisiana
135867	3/21/2015	11/23/2015	—	247	47	1141	30.46	−91.57	Louisiana
135868	3/21/2015	4/7/2015	—	17	49	84	30.46	−91.57	Louisiana

Table 1. Telemetry summary. The PTT transmitter identifier (Tag), date of deployment (Deployment), last transmission date (Last), bird mortality type if known (Mortality), duration of record (Duration, days), median sampling interval between locations (Median, minutes), total locations (Sample), and the coordinates (Latitude, Longitude) and geographic region (Region) of deployment for records retained for analysis.

gov/centers/eros), the 300 meter 2009 GlobCover Land Cover Maps provided by the European Space Agency (http://due.esrin.esa.int/page_globcover.php), and 30 meter resolution topographic roughness and topographic wetness indices provided by the ENVIREM dataset[50]. Because our goal with incorporating environmental variables was to characterize waterfowl habitat broadly and not to infer species-specific preferences or tolerances, we decomposed gridded soils data (17 different soil attributes) and elevation derived variables (topographical roughness, topographical wetness, elevation, and slope) using separate Principal Components Analyses and selected the resulting first component from each to represent soil and topographical conditions respectively. The correlation of the retained first component to other attributes included during decomposition are provided for soils in Table 2 and topography in Table 3. In addition to simplifying our analysis, the decomposition of variables also aided in avoiding multicollinearity issues during the model fitting process[51]. Recognizing the critical importance of wetlands and surface waters to waterfowl, we also calculated the Euclidean distance (km) of each telemetry location

3

Attribute	r (%)	Description
AWCh1	−0.21 (0.01)	Available water capacity (Horizon 1)
AWCh2	−0.20 (0.01)	Available water capacity (Horizon 2)
AWCh3	−0.20 (0.01)	Available water capacity (Horizon 3)
AWCtS	−0.46 (0.03)	Saturated water content
BLDFIE	0.48 (0.03)	Bulk density (kg/m^3)
CECSOL	−0.23 (0.01)	Cation exchange capacity (cmolc/kg)
CLYPPT	0.18 (0.00)	Clay content mass fraction
CRFVOL	−0.68 (0.15)	Coarse fragments volumetric
OCDENS	0.20 (0.59)	Organic carbon density (kg/m^3)
OCSTHA	−0.21 (0.02)	Organic carbon stock (tons/ha)
ORCDRC	−0.32 (0.06)	Organic carbon content (g/kg)
PHIHOX	0.28 (0.03)	Soil pH $\times$ 10 (H_2O)
PHIKCL	0.23 (0.01)	Soil pH $\times$ 10
SLTPPT	0.38 (0.02)	Silt content mass fraction
SNDPPT	−0.44 (0.01)	Sand content mass fraction
TEXMHT	−0.08 (0.01)	Texture class (USDA system)
WWP1	−0.15 (0.00)	Available water capacity (wilting point)

Table 2. Soils analysis. Correlation of 17 soil attributes to decomposed component retained as soil covariate. Soil attribute name (Attribute), Pearson correlation (r) with parenthetical proportion of variance explained (%), and attribute description. All correlations were significant at the $\alpha = 0.001$ threshold.

| Attribute	r (%)
Slope	−0.51 (0.00)
Elevation	−0.59 (0.01)
Compound Topographic Index	0.93 (0.81)
Terrain Roughness Index	−0.52 (0.18)

Table 3. Topographical analysis. Correlation of 4 topographical attributes to decomposed component retained as a topographical covariate. Attribute name (Attribute), Pearson correlation (r) with parenthetical proportion of variance explained (%), and attribute description. All correlations were significant at the $\alpha = 0.001$ threshold.

to the nearest surface water class in the GlobCover land cover dataset using the Program R[52] and the **spatstat** package[53]. As detailed in subsection 2.6, we modeled distance to water as a non-linear covariate to avoid treating water presence as a bird occurrence prerequisite.

Poultry and human density data. For estimates of poultry abundance, we used the chicken and duck raster layers from the Food and Agriculture Organization of the United Nations (FAO) Gridded Livestock data-base[54]. The rasters were provided at a 0.083333 decimal degree resolution and report the number of poultry head per square kilometer ($head/km^2$). Human population density estimates were obtained as a continuous raster (30 meter) from NASA's Socioeconomic Data Center that gave the number of persons per square kilometer (persons/km^2)[55]. The distribution of poultry abundance relative to our study domain is shown in Fig. 1 with telemetry tracks overlaid for marked birds.

Avian influenza event in poultry data. AIP data were obtained from the web-based Global Animal Disease Information System (EMPRES-i) maintained by the FAO (http://empres-i.fao.org/eipws3g/, accessed 12/01/19). EMPRES-i is intended to support veterinary services and related organizations by providing access and analysis of global animal disease information. We initially downloaded all AIP data for the period between January 01, 2004 and December 01, 2019 and then filtered records to include only those with an FAO "Confirmed" status, geographically located within our study domain, affecting commercial poultry, and associated with virus subtypes having a H5 or H7 hemagglutinin surface protein. The H5 and H7 subtypes were chosen due to our interest in the HPAI pathotype, however, both HPAI (n = 440) and LPAI (n = 49) events were included in the resulting 489 record dataset. A secondary consideration in limiting the study to the H5 and H7 subtypes was that these subtypes are required to be reported to the World Organization for Animal Health (OIE) giving us some confidence in our sample. This is not the case with other virus subtypes, which are reported voluntarily and may therefore introduce sample bias. As a final step in preparing AIP data, we randomly selected 97 AIPs (approximately 20%) as a hold-out dataset to aid in later model validation. Locations of all events are shown by virus subtype in Fig. 2.
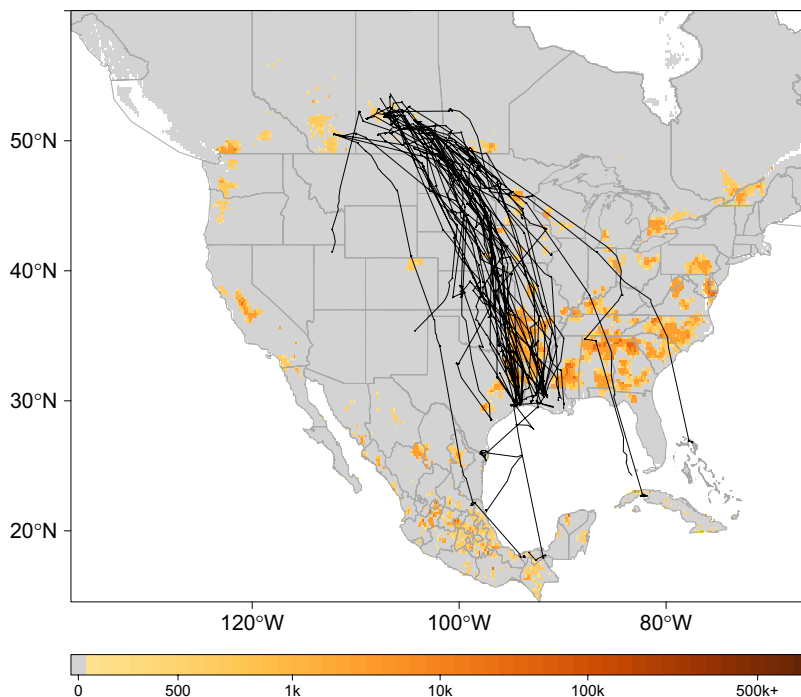
Figure 1. Track Map. Telemetry tracks for 42 blue-winged teal over grid depicting poultry abundance[54]. Black lines show individual bird tracks, legend describes estimated number of poultry (chickens and ducks) per km². To better display poultry abundance in this figure, zero was defined as less than 100 poultry/km².
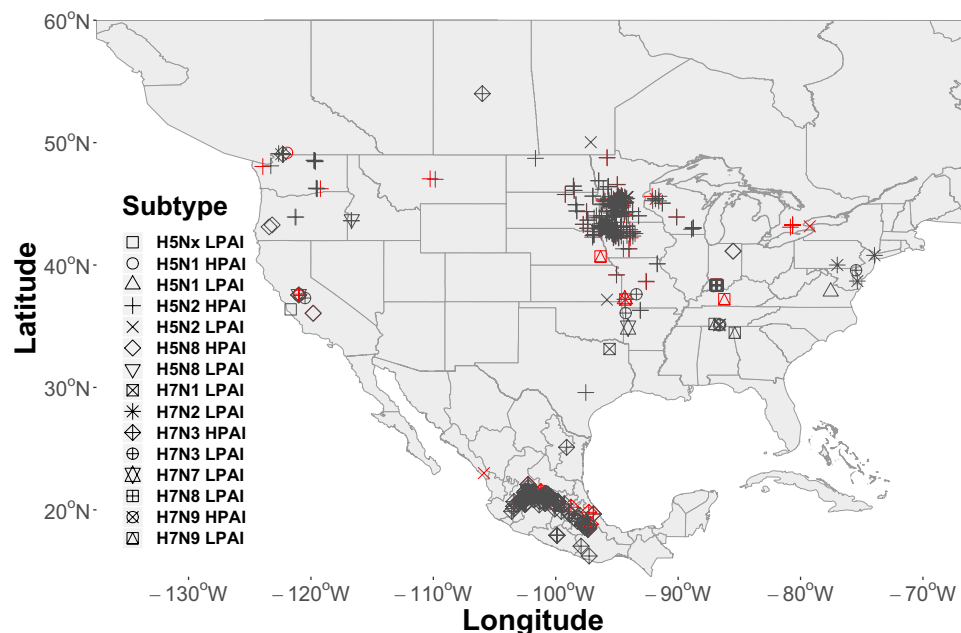
Figure 2. Locations of Avian Influenza Events (AIP) by virus subtype. AIP cover the period 2004–2019 and included 440 events used for model training and 49 (20%) randomly selected for model validation. Locations shown in the color red signify those retained for validation.

Statistical analysis. We constructed a two-level, spatially explicit model with shared components to evaluate the spatial and temporal relationship of waterfowl to AIPs across North America. The model's first level estimated waterfowl residence time across the study area, whereas, the model's second level provided the probability of an AIP as conditioned on that waterfowl residence time. More generally, our model can be represented as:

$$y_1(i) = \alpha_1 + \beta \cdot X_1(i) + Z_1(i) + e_1(i) \tag{1}$$

$$y_2(i)|y_1(i) = \alpha_2 + \beta \cdot X_2(i) + Z_2(i) + Z_{shared}(i) + e_2(i) \tag{2}$$

where the first level [Eq. 1] estimates waterfowl residence time (y_1) at locations i ($i = 1, 2, 3, \ldots, n$) and the second level [Eq. 2] gives AIP probability (y_2) based on the past event (H5 or H7 AI detection) occurrence or absence (1, 0) at location i and conditioned on waterfowl residence time (y_1). This formulation offers independent intercepts for each model level (α_j), covariate matrices ($\beta \cdot X_j$) for level-specific linear and fixed effects, and uncorrelated error terms (e_j). Although independent spatial fields (Z_1 and Z_2) help account for spatial dependencies within each model level, we added a third spatial field (Z_{shared}) to explicitly serve as a shared scaling parameter between model levels[56,57]. The shared field is a re-scaled copy of Z_1 that controls for spatial autocorrelation and interaction occurring between waterfowl telemetry and AIP locations. As an initial step in specifying the spatial fields, we created a triangulated mesh over our study area using procedures described by Lindgren and Rue[58]. In addition to facilitating parameterization of the continuous spatial fields needed to account for spatial dependencies, the mesh provided a computationally efficient alternative to characterizing our continent-sized study domain using raster grids or areal polygons. The mesh included 9,555 nodes and had an outer extension large enough to avoid edge effects around study area boundaries. Due to our large study area, the mesh was projected using three-dimensional Cartesian coordinates scaled to one Earth radius (6,371 km). As a component of sensitivity analyses, meshes constructed with 4,897 nodes and 12,651 nodes were also evaluated and found to exhibit a mean spatial effect within 10% of that estimated using 9,555 nodes; therefore, we selected the middle-ground to balance precision and processing time.

To accommodate the joint modeling approach, our dependent variable was structured as a bivariate matrix with the first column corresponding to residence times calculated from telemetry data and the second column being a vector of 1's and 0's designating a past AIP occurrence (1) or absence (0) at each location. Because bird and AIP locations did not always geographically coincide, our model incorporated spatial misalignments; however, this was not problematic as the model is capable of providing estimates across continuous space and for all locations within the study domain[56]. To acquire initial residence time estimates for model fitting, we disaggregated individual bird telemetry tracks (Fig. 1) into 1 minute temporal intervals and then summed them based on mesh node proximity (i.e., "natural neighborhoods"). To achieve this, we performed a Voronoi tessellation around mesh nodes to identify each node's respective area of influence[56]. Area of influence refers to the bounded area surrounding each node in which all enclosed locations share that node as their closest point[59]. The process used to disaggregate tracks into equal time intervals was adapted from Sumner[60], but, modified to allow for summation of time intervals based on our tessellated neighborhoods as opposed to a regular grid. Each column of the bivariate matrix exhibited a different distribution, therefore, we specified model likelihoods as,

$$y_1(i) \sim \text{Gamma}(a_i, b_i) \text{ and } y_2(i) \sim \text{Binomial}(\pi_i)$$

where, residence time was always a positive value with shape and scale parameters (a_i, b_i) such that $a_i/b_i = \mu_i = E(v_i) \cdot Exposure$. Defining *Exposure* as the area of each node's tessellated neighborhood, the residence-time linear predictor v_i was of the form,

$$\begin{aligned} \log(v_i) &= \alpha_1 + \beta_1 \text{Soil}_i + \beta_2 \text{Topography}_i \\ &\quad + \beta_3 \text{WetDistance}_i + Z_1(i). \end{aligned} \tag{3}$$

here, a log link function was used with α_1 as the intercept, β_1 the soil variable coefficient, β_2 the topography coefficient, β_3 the coefficient for linear distance to nearest wetland, and Z_1 as the spatial field to control for spatial dependencies and other unobserved latencies (errors) that may have resulted from the aggregating process undertaken to estimate residence time.

The corresponding linear predictor for the binomially distributed AIP occurrence vector relied on the logit function and was specified as,

$$\begin{aligned} \text{logit}(\pi_i) &= \alpha_2 + \beta_1 \text{Poultry}_i + \beta_2 \text{Population}_i \\ &\quad + \beta_3 \cdot f_1 \text{Cluster}_i + \beta_4 \cdot f_2 \text{Displacement}_{i,t} \\ &\quad + Z_2(i) + Z_{shared}(i), \end{aligned} \tag{4}$$

where α_2 is an intercept, β_1 is the coefficient for poultry density and β_2 is the human population density coefficient. As described for Eq. 2, Z_2 provides a spatial field to account for domain-wide spatial dependencies and latencies among AIP locations, whereas, Z_{shared} controls for the spatial relationship between AIP and waterfowl locations. To supplement poultry and human population density variables, we developed two additional model covariates; one to account for fine-scale spatial structure among AIP locations (f_1 Cluster) and the second to examine the temporal relationship between seasonal waterfowl migration and the timing of AIPs (f_2 Displacement).

As shown by Fig. 2, the geographic distribution of AIPs across North America exhibits broad regional coverage (e.g., the Midwestern U.S.) as well as more localized or concentrated "clustering" at some locations (e.g., Jalisco, Mexico or Minnesota, USA). We speculate that the epidemiological processes that gave rise to the fine-scaled, cluster patterns were likely different than those that instigated the broad-scale, region-wide coverage. More specifically, it is possible that farm-to-farm transmission or "lateral spread" produced the observed clusters,
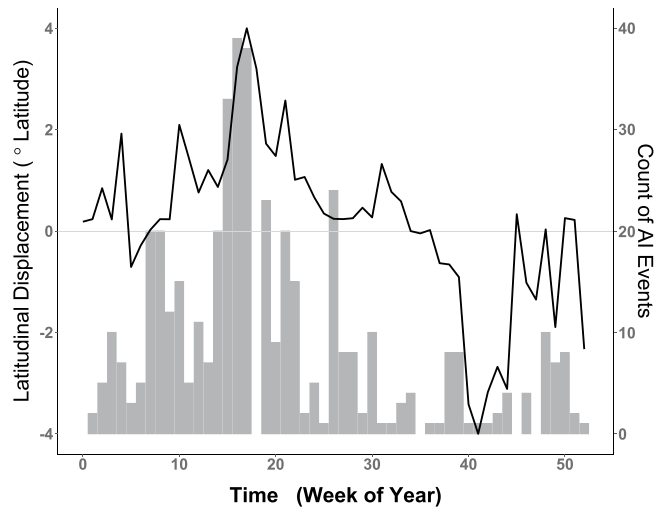
Figure 3. Weekly comparison of historic AI Events (AIPs) with waterfowl latitudinal displacement. Left vertical axis corresponds to smooth black line and represents weekly latitudinal displacement of waterfowl measured in degrees latitude. Horizontal axis provides the week of year. Dotted line intersecting zero signifies no net waterfowl displacement, with portions above zero indicating relative northward movement and portions below showing net southward change. Vertical axis at right corresponds to gray histogram for the weekly counts of all AIP summed over the period 2004–2019. Note apparent temporal correlation between waterfowl spring migration (Weeks 5–18) and increasing AIP as well as that between fall migration (weeks 35–42) and decreased AIP counts.

whereas, the broader coverage may be better explained by waterfowl introduction (see, Discussion for further explanation). We chose to address these different multi-scaled spatial processes using two separate spatial covariates. The spatial field (Z_2) was used as a domain-wide spatial covariate, while fine-scale spatial structure associated with AIP clusters was modeled using a random walk process across AIP nearest neighbor distances (km). As used here, a random walk (f_1 Cluster) is similar to a regression spline in that it permits distances between AIPs to follow a smooth, curvilinear distribution without assuming a mean response. The benefit of using the fine-scale spatial effect is that it quantifies the spatial variability due to proximity between AIP locations, thereby, reducing the model's capacity to attribute that variation to waterfowl or other covariates.

As with clustering, a random walk was also implemented to quantify the temporal relationship between seasonal waterfowl migration and the timing of AIPs. As shown in Fig. 3, summing raw AIP counts (all years, 2004–2019) by week reveals an apparent correlative relationship of AIP timing to the mean weekly latitudinal displacement exhibited by waterfowl. To capture the timing of waterfowl movement as a model covariate, we calculated latitudinal displacement by finding the weekly mean latitude occupied by marked birds (all years, 2013–2016) and then performing a one-week lag subtraction between weeks. The function f_2 Displacement introduces the resulting time-ordered vector as a non-parametric, dynamic effect that captures waterfowl latitudinal change at each weekly time step t ($t = 1, 2, 3, \ldots, 52$).

We used a Penalized Complexity (PC) framework[61,62] for the three spatial fields (Z_j) with priors scaled to one Earth radius to reflect the size and projection of our 3D triangulated mesh. Because all spatial fields occurred over a common mesh and spatial domain, PC priors were set with both the spatial range and standard deviation quantile and probability: $\rho_0 = 1$ and $p_\rho = 0.5$. The PC priors for the spatial effect were intended as non-informative and can be interpreted as specifying a 0.5 probability that the spatial range is within 6,371 km (i.e., one Earth radius). Sensitivity analyses for probability values between 0.2 and 0.8 in 0.1 increments assuming a $\rho_0 = 1$ revealed minimal change to model estimates. Flat PC priors were likewise applied for the random walk effects modeled as f_1 Cluster and f_2 Displacement, which were specified as $\mu = 1$ and $\alpha = 0.0001$. We then combined priors and likelihoods using Integrated Laplace Approximation[63].

In addition to performing joint estimation of waterfowl residence time and AIP occurrence probability as described above, we also explored an alternative model formulation to jointly estimate waterfowl occurrence probability with AIP occurrence probability. That is, rather than using residence time to inform AIP probability, the alternate formulation used mere waterfowl occurrence to estimate AIP probability as a way of comparing the relative importance of bird presence versus the amount of time spent at a location. Under the alternative model, we assumed that all locations with a residence time equal to, or exceeding 24 hours constituted an occurrence (1) with unobserved locations (0) used to characterize background environmental conditions. To correspond with the derived presence-background vector, we substituted a binomial likelihood for the gamma in Eq. 3, and exchanged a logit link function for the log. The alternative binomial-binomial configuration included all of the same spatial and environmental covariates as described for the Gamma-binomial.

Model	DIC	WAIC	lCPO	Description
Model1	2134.42	1963.92	0.178	Space Only
Model2	1247.12	1215.90	0.062	Space + Clustering
Model3	1114.29	1084.77	0.054	Space + Time
Model4	1039.38	1019.96	0.051	Space + Time + Covariates
Model5	842.08	824.30	0.043	Joint Model (Duration)
Model6	836.47	821.35	0.042	Joint Model (Occurrence)
Model7	3710.55	3715.98	0.184	Non-Spatial + Covariates

Table 4. AIP model comparison. Deviance information criterion (DIC), Watanabe-Akaike information criterion (WAIC), and log-conditional predictive ordinate (lCPO). Lower values indicate improved parsimony. Joint models displayed in this table, metrics reflect the AIP occurrence probability level of the model only (Eq. 2).

Model evaluation.   To evaluate model accuracy, we performed model-based estimation for each AIP location and event time in the independent testing dataset that was randomly selected and excluded from model development. We calculated the percent correctly classified (PCC), sensitivity (proportion of correctly predicted presences), specificity (proportion of correctly predicted absences), and the area under the receiver operating characteristic curve (AUC) using the Presence Absence R package[64], and then estimated the True Skill Statistic (TSS) by subtracting a value of one from the sum of the sensitivity and specificity estimates[65]. We also used tools available in the Presence Absence package to determine the probability thresholds that best maximize sensitivity, specificity, and the TSS. Probability thresholds were determined by comparing the input data to model estimates to ascertain the probability value (threshold) that best captured known disease occurrences.

In addition to validating with an independent out-of-sample testing data set, we applied a comparative approach to assess the overall performance of our joint waterfowl residence time and AIP occurrence probability model. In total, seven different models were constructed and then compared to the full joint model detailed in subsection 2.6, one of which (Model6) was the alternative formulation described in the closing paragraph of subsection 2.6. Recognizing that application of parsimony metrics to spatial models can be problematic[66,67], we chose three different parsimony measures; the deviance information criterion (DIC), the Watanabe-Akaike information criterion (WAIC), and the log-conditional predictive ordinate (lCPO). Generally, the DIC and WAIC are comparable, however, the DIC sometimes under-penalizes random effects[67]; thus, we opted to use both the DIC and WAIC as well as leave-one-out cross-validation (lCPO). Descriptions for all comparative models are provided with accompanying DIC, WAIC, and lCPO in the results section.

Results

Table 4 lists all models used to assess performance relative to the full model described in subsection 2.6. Note that in Table 4, the DIC, WAIC, and lCPO indicate improved parsimony as additional covariates were added, and that joint models combining waterfowl telemetry (Model5 and Model6) with AIP occurrences exhibited the best overall performance. As a visual comparison, Fig. 4(A) maps the smoothed random field densities produced by the base spatial model with no covariates (Model1) in comparison to a spatial model with all covariates, but, excluding waterfowl telemetry (Model4, Fig. 4(B)), and the joint models incorporating waterfowl residence time (Model5, Fig. 4(C)) and occurrence probability (Model6, Fig. 4(D)). Each of the sub-figures in Fig. 4 display a zero mean domain density and can be interpreted in a manner similar to mapped spatial residuals. The map's warm colors (oranges and reds) indicate regions where actual values are greater than those estimated by the model and cool colors (blues) highlight regions where actual values fall below those estimated by model. Note that random field densities for both joint models are reduced relative to those resulting from Model1 and Model4, but, the joint models still underestimate AIPs over areas in the upper Midwestern and Northwestern U.S., as well as central Mexico. Residual error indicates that additional factors or variables are needed to fully explain the AIPs at these locations.

Tables 5 and 6 summarize model covariates for the waterfowl residence time and occurrence probability models jointly fit with AIP occurrence probability. Observe that the soils covariate is important as judged by its credible intervals for both Model5 and Model6, whereas, the topography covariate is not. The soils and topography covariates likely correlate to waterfowl habitat preferences, but, this variation is better captured by soils than by topography. Soils and topography have been successfully used to model wetland occurrence probability in other studies[68]. The distance to water covariate is negative for both models suggesting that as distance from water increases, both waterfowl occurrence probability and waterfowl residence time decrease. Coefficients for poultry abundance are similar across models signifying an approximate 52.9% $[(1/(1 + \exp(0.12))) \times 100\%]$ increase in AIP occurrence probability for every additional 100,000 poultry head/km^2 above the domain average while holding all other covariates constant. The f_1 Cluster effect is plotted in Fig. 5 and indicates that the probability of an AIP occurrence increases non-linearly as proximity to a known AIP decreases below approximately 300 km. More simply, the closer a location is to a detection of H5 or H7 subtype influenza virus, the more likely it is to be subject to an AIP itself. The lower credible interval in Fig. 5 is inclusive of zero at distances greater than about 170 km, meaning that there is a low posterior probability of equal occurrence odds at longer distances. Figure 6 displays the functional relationship of AIP occurrence to time-structured waterfowl latitudinal displacement. The effect is important over most of the year and suggests that the greatest probability of an AIP occurs at approximately the 18th week of the year (late April - early May), a period largely coinciding with the peak of northward mean waterfowl latitudinal displacement (spring migration, see Fig. 3).
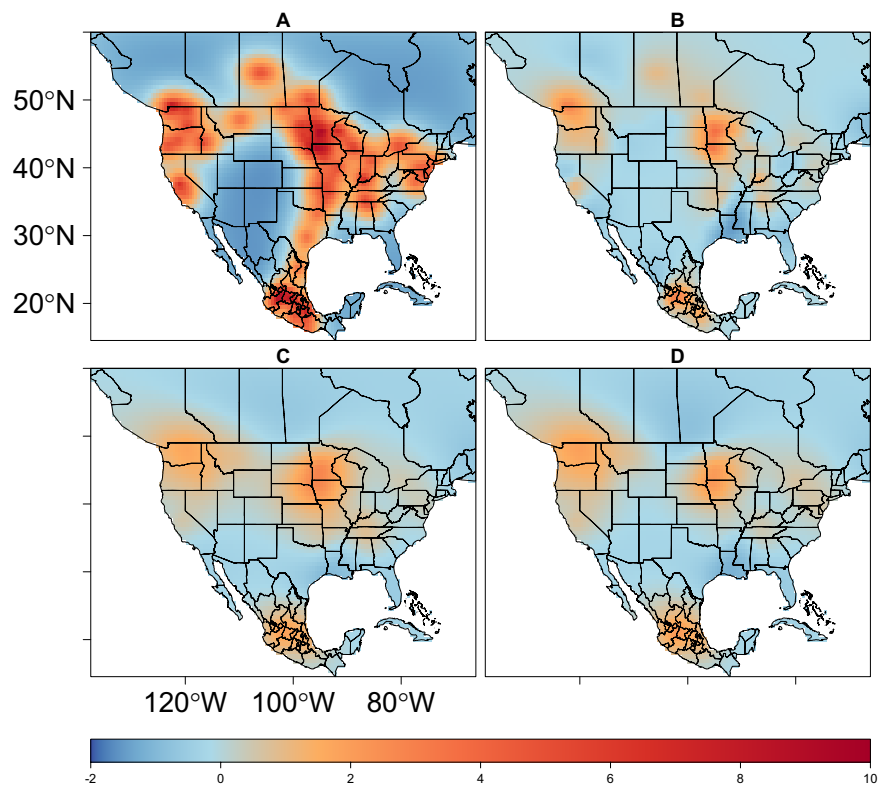
Figure 4. Random field density comparison. Spatial field resulting from the (**A**) base model (Model1), (**B**) base model with clustering effect and other covariates, (**C**) joint-model for residence time (Model5), and (**D**) joint-model for occurrence (Model6). Red areas highlight where models under predict AIP, cooler blue regions indicate where models over predict. Maps are plotted on the same scale using the same palette and can be interpreted as "mapped residuals". Note that random field densities for models incorporating telemetry (**C,D**) are reduced relative to models without telemetry (**A,B**), but, even the telemetry models underestimate AIPs over several areas in the upper Midwestern and Northwestern U.S., as well as in central Mexico. Residual error indicates that additional factors or variables are needed to fully explain the AIPs at these locations.

|	Mean	SD	Q025	Q975
Intercept1 (α_1, Eq. 3)	1.66	0.24	1.18	2.13
Intercept2 (α_2, Eq. 4)	−8.06	1.2	−10.42	−5.71
Soils	0.19	0.07	0.06	0.32
Topography	−0.10	0.18	−0.47	0.27
Distance to wetland	−0.02	0.01	−0.04	−0.01
Poultry abundance	0.12	0.03	0.06	0.18
Human population	0.05	0.01	0.02	0.07
Gamma precision	2.17	0.27	1.70	2.73
Range Z_1	0.09	0.03	0.05	0.15
SD Z_1	0.89	0.15	0.64	1.21
Range Z_2	0.25	0.10	0.12	0.50
SD Z_2	1.39	0.39	0.81	2.32
f_1 Cluster	0.10	0.04	0.05	0.19
f_2 Displacement	0.76	0.27	0.36	1.41
Z_{shared}	0.56	0.25	0.05	1.04

Table 5. Joint model (Model5, residence time) effects and hyperparameters. Mean, standard deviation (SD) and 95% Credible Interval. Coefficients are on the logit scale (logarithm of odds ratio).

Of particular interest in Tables 5 and 6 are the shared spatial fields (Z_{shared}) that represent the relationship of waterfowl telemetry to AIP occurrence probability. The shared fields for both the waterfowl residence time and occurrence probability joint models are estimated to be positive, with 95% credible intervals that exclude zero. In

	Mean	SD	Q025	Q975
Intercept1 (α_1, Eq. 3)	−6.15	0.64	−7.40	−4.90
Intercept2 (α_2, Eq. 4)	−8.02	1.20	−10.37	−5.67
Soils	0.36	0.14	0.09	0.65
Topography	0.23	0.30	0.36	0.81
Distance to wetland	−0.03	0.01	−0.05	−0.01
Poultry abundance	0.12	0.03	0.06	0.19
Human population	0.05	0.01	0.02	0.07
Range Z_1	0.15	0.04	0.09	0.25
SD Z_1	3.98	0.66	2.86	5.45
Range Z_2	0.25	0.10	0.12	0.51
SD Z_2	1.35	0.38	0.79	2.26
f_1 Cluster	0.10	0.04	0.04	0.19
f_2 Displacement	0.77	0.27	0.36	1.42
Z_{shared}	0.13	0.08	0.04	0.30

Table 6. Joint model (Model6, occurrence probability) effects and hyperparameters. Mean, standard deviation (SD) and 95% Credible Interval. Coefficients are on the logit scale (logarithm of odds ratio).
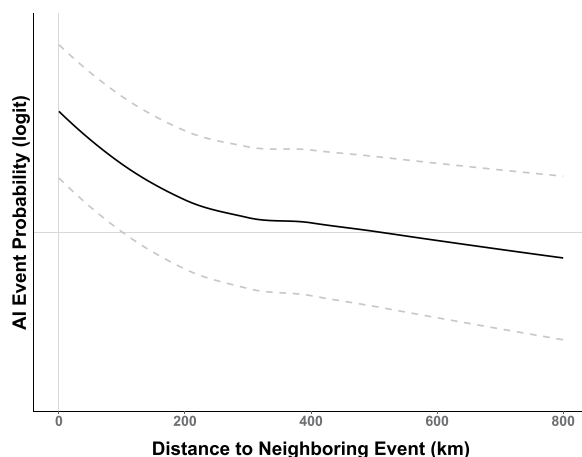
Figure 5. Model estimated clustering covariate. Non-linear, smooth line describes AIP probability as a function of proximity to other, adjacently located AIP occurrences. Vertical axis provides probability (logit scale) and the horizontal axis gives distance in kilometers. Dashed gray lines display 95% Credible Interval. Note that the credible interval includes zero at distances greater than approximately 170 km.
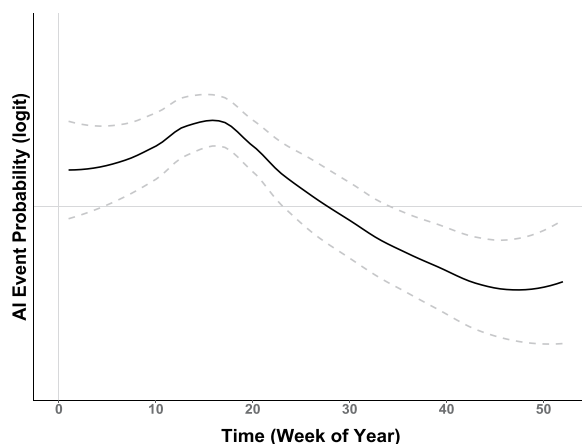
Figure 6. Model estimated temporal effect. Non-linear response of AIP probability as a function of time-structured bird latitudinal displacement. Vertical axis provides probability (logit scale) and the horizontal axis gives time in weeks. Dashed gray lines display 95% Credible Interval. Note that the credible interval includes zero between Weeks 20–35, a time period corresponding with the waterfowl breeding season.
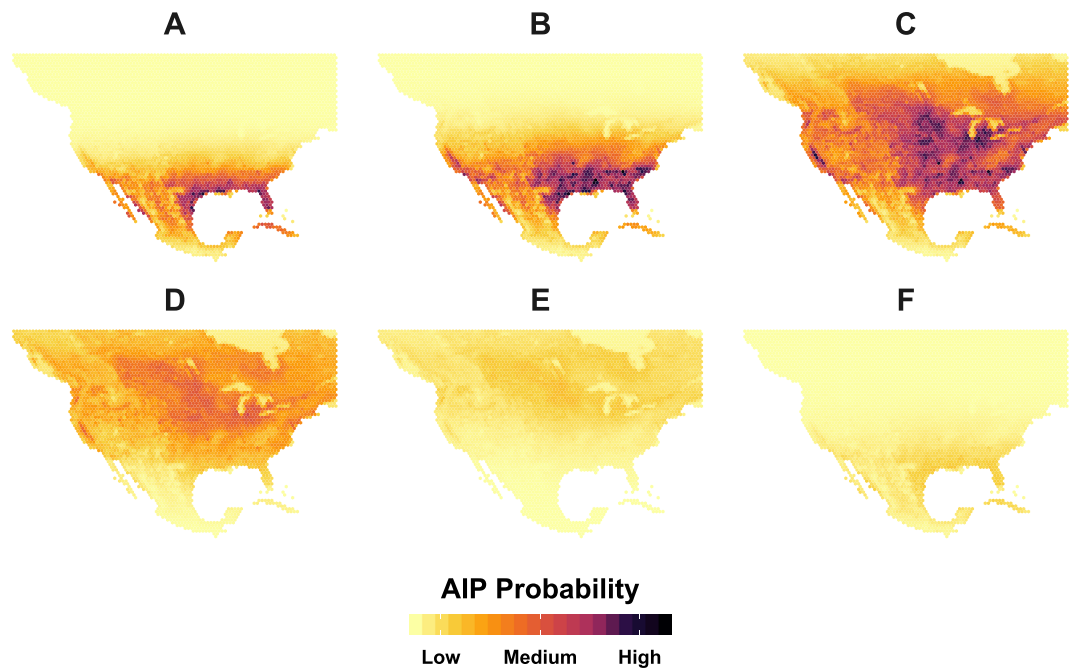
Figure 7. Mapped AIP probabilities predicted by joint-model (Model5). Weeks 7, 15, 18, 23, 33, and 51 of the year are labeled as A through F respectively. Color coding and legend reflect relative probability (0.00–1.00) of H5 or H7 occurrence in North American Poultry with darker colors signifying increased likelihood. Note that C representing week 18 of the year (late April – early May) was predicted to exhibit the greatest AIP probability.

Threshold	PCC	Sensitivity	Specificity	AUC	TSS
0.50	0.81	0.95	0.67	0.96	0.61
0.74	0.90	0.89	0.91	0.96	0.80
0.76	0.90	0.88	0.91	0.96	0.79

Table 7. Model validation summary. Percent Correctly Classified (PCC) at given Threshold. Model sensitivity, specificity, and the AUC are scaled from 0 to 1. The TSS score is ranged from −1 to 1. Higher scores signify improved performance for all metrics.

the instance of Model6, the Z_{shared} estimate is interpreted to indicate that if the probability of waterfowl occurrence at a location is 0.5, then the probability of an AIP is greater than 0.5 while controlling for all other model covariates. More precisely, at locations having a 0.5 waterfowl occurrence probability, the chance of an AIP is 53.2% $[(1/(1 + \exp(0.13))) \times 100\%]$ while controlling for all other covariates. As the occurrence probability of waterfowl increases above 0.5, the chance of an AIP likewise increases above 53.2%. In the case of Model5, the Z_{shared} estimate translates to a 63.6% $[(1/(1 + \exp(0.56))) \times 100\%]$ chance of an AIP at locations exhibiting the domain-wide mean waterfowl residence time of 8.6 days while holding all else constant. A greater AIP likelihood is expected for locations with longer waterfowl residence time. Combining all covariates, AIP probability can be estimated for any week of the year and any location within our study area; Fig. 7 displays the relative chance of an AIP for North America for select weeks of the year.

Model validation results are summarized in Table 7, which provides the PCC, Sensitivity, Specificity, AUC, and TSS for the hold-out data randomly partitioned during preliminary analysis. Statistics are provided for the three AIP presence thresholds determined to maximize Sensitivity, Specificity, and the TSS. To supplement Table 7, Table 8 further describes the validation dataset by providing the count of each virus subtype retained for evaluation as well as the average estimate and probability of cluster membership at each location.

Discussion

Considering forty-two marked blue-winged teal as representative natural host waterfowl, we applied advanced temporal modeling techniques to evaluate the location and timing of waterfowl movement and space-use (1) relative to avian influenza events (AIP) across North America between 2004 and 2019 and (2) confirmed as resulting from viruses having the H5 or H7 hemagglutinin surface protein. While our primary goal was to utilize multi-year animal movement telemetry and disease event data to characterize intra-annual patterns, we also emphasized methodological facets to guard against attributing avian influenza detections to waterfowl when they are better explained by mere spatial proximity or membership to the AIP clusters in our sample. To accomplish these goals, we developed a dynamic joint (waterfowl-AIP) temporal model with shared components. Our

Subtype	Estimate (95% CI)	Cluster (95% CI)	Count
H5N1 HPAI	1.00 (1.00, 1.00)	0.99 (0.99, 0.99)	1
H5N1 LPAI	0.98 (0.98, 0.98)	0.79 (0.79, 0.79)	1
H5N2 HPAI	0.90 (0.85, 0.94)	0.61 (0.52, 0.69)	46
H5N2 LPAI	0.72 (−0.34, 1.79)	0.67 (−0.10, 1.44)	3
H7N1 LPAI	0.97 (0.97, 0.97)	0.79 (0.79, 0.79)	1
H7N3 HPAI	0.87 (0.80, 0.95)	0.76 (0.66, 0.85)	38
H7N3 LPAI	0.78 (−1.98, 3.54)	0.80 (−1.67, 3.26)	2
H7N8 LPAI	0.80 (0.14, 1.46)	0.51 (−0.76, 1.77)	2
H7N9 LPAI	0.72 (0.29, 1.15)	0.52 (−0.08, 1.11)	3

Table 8. Predictive accuracy by virus subtype. Model estimated event probability (Estimate), probability that the event belonged to a cluster (Cluster), and the count of testing samples randomly selected from original dataset.

combined animal movement and disease distribution framework was sufficiently flexible to accommodate a large study domain and enabled us to account for cumulative error across what might typically be approached as individual or separate analyses.

Results indicate that the presence of migratory waterfowl is positively related to AIP occurrence within North America. As waterfowl occurrence probability or residence time increase at a given location, so too does the chance of an AIP. This conclusion is supported by estimates for the shared spatial fields (see, Z_{shared}, Tables 5 and 6), which were specifically implemented to account for spatial interaction between waterfowl and AIP locations. The 95% credible intervals for the shared fields included with both the waterfowl residence time (Model5) and occurrence probability (Model6) models indicate that waterfowl presence improves the model's ability to estimate AIPs and that inclusion of movement telemetry increases parsimony (Table 4). Importantly, our inferred support for waterfowl improving AIP estimation is made after accounting for the week-specific, latitudinal displacement of migratory birds (Model3), poultry abundance (Model4), and AIP clustering (Model2). While our approach indicates that waterfowl presence and residence time were associated with North American AIPs during 2004–2019, we cannot conclude that any individual AIP was the direct result of wild bird to poultry viral spillover. Given that HPAI phenotypes are recognized as only evolving in domestic poultry, it is probable that associations between waterfowl presence and AIPs are a function of LPAI H5 and H7 spillover from wild birds to domestic poultry, or that the spatiotemporal associations resulted from other unidentified and unmeasured factors.

Regarding AIP clusters, we speculate that not all[69] of the clustered AIPs visibly apparent (Fig. 2) over portions of Mexico (e.g., Jalisco) and the upper Midwestern U.S. (e.g., Minnesota) represent independent introductions of avian influenza by wild birds. Given the compact spatial arrangement of the events, relative incident timing (Fig. 3)[70,71], and the commonality of virus subtypes at each location (e.g., H7N3 in Jalisco, H5N2 in Upper Midwest), we conjecture that AIP clusters likely stemmed from some combination of mechanical (i.e., shared farm equipment, staff movement between farms, etc.) and environmental (i.e., common water, soil, and air) transmission[28,72]. This interpretation is consistent with recent genetic analyses showing viral relatedness among H5N2 HPAI outbreaks in the Midwest[26,69] and among H7N3 HPAI outbreaks in Mexico[17]. In the absence of extensive genetic sampling and lacking information detailing farm-specific operational practices, the current study is insufficient in itself to assess the tenability of a "farm-to-farm" or "lateral spread" disease cluster hypothesis directly, but, from a modeling context we propose Tobler's Law[73] as an accessible heuristic to account for the spatial patterns that sometimes arise from disease and epidemiological processes. The constructed covariate we designed to quantify fine-scale spatial structure is a statistical implementation of Tobler's Law in that it assigns a cluster membership probability to all locations in the study area based on proximity to a neighboring AIP. To paraphrase Tobler from an epidemiological perspective, we assumed that "nearer outbreaks are more related than distant outbreaks". If a detection is better explained by being in close proximity to another AIP than it is by waterfowl occurrence, then the risk of inferring an independent introduction by waterfowl is inherently reduced.

The timing of seasonal waterfowl migration was also found to be an AIP indicator. The time-structured covariate reflecting bird net latitudinal displacement was important as shown by 95% credible intervals (Fig. 6) and improved model parsimony (Model3, Table 4). Model results suggest that AIP occurrence probability is greatest during the months of April and May, a period coincident with peak blue-winged teal northward movement, and less during the months of September and October, which is the peak of observed southward movement. Stated differently, the chance of an AIP occurring in North America is greatest during the spring migration, less during the fall migration, and generally insignificant during the waterfowl breeding and overwintering seasons. More generally, the temporal relationship of waterfowl to past AIPs suggests that the chance of avian influenza being introduced into a U.S. poultry facility steadily increases between February and May and then rapidly decreases for the remainder of the year. Although increased hatchling abundances and bird aggregation at pre-migration staging areas through the late-breeding and fall seasons (August and September) are thought to contribute to a generally higher AI prevalence in dabbling ducks broadly, this spatiotemporal pattern may not be as consistent for blue-winged teal specifically. Having a reputation for early departure and long flight distances, vast numbers of blue-winged teal have already vacated pre-migration staging areas and arrived at migration stop-over sites or overwintering habitats by August and September[40,74]. Having experienced reduced rates of virus exposure prior to autumn migration, immunologically naïve blue-winged teal may help maintain AI at lower latitude locations

throughout the non-breeding period and potentially redistribute virus during the spring migration[40,41,75]. As it relates to both the timing and degree of clustering shown in our AIP sample, it is important to note that models underestimate AIP probability over several areas in the upper Midwest and northwestern U.S., as well as in central Mexico (Fig. 4(C,D)). This indicates that additional factors or variables are needed to fully explain the epidemiological processes at these locations. Remaining error may be due to delayed surveillance or event reporting, environmental persistence, post-introduction redistribution by wild birds, farm-specific operational practices, species-specific virus relationships, or other unmodeled factors.

Geographically, several regions appear particularly vulnerable to an AIP. In the U.S., the states of Mississippi, Alabama, and Georgia in the Southeast exhibit elevated risk from February through April (Fig. 7). The Delmarva Peninsula on the East Coast shows an increased chance of a poultry influenza event during the second half of March through April as does the southern Great Lakes Region and central California on the West Coast. By comparison to the U.S., model estimates indicate relatively lower AIP probabilities for southern Canada and Mexico. Although Southwestern Ontario and the Montreal vicinity show increased risk at the height of waterfowl spring migration (Fig. 7(C)), AIP suitability in Canada falls below 15% for the remainder of the year. In Mexico, the highest AIP probability was estimated for the second-half of January and first part of February when conditions were approximately 30% suitable for an AIP along the western-most edge of the Mexican Plateau between Jalisco in the north and Puebla to the south. This estimated period of highest AIP probability in Mexico coincides with the timing of blue-winged teal staging for the spring migration.

Although validation evidenced overall good model performance (Tables 7 and 8), comparing model estimated AIP probabilities to the full disease occurrence dataset used as input (Fig. 2) suggests that the model's ability to capture past AIPs varied by virus subtype. More specifically, waterfowl space use and migration timing show greater correspondence to past H7 AIPs than to those of H5 origin. The majority of H5 events (e.g., H5N2 and H5N8) incorporated into our analyses were sampled from the outbreaks in 2014–2015 that affected 21 States across the Northwestern and Midwestern U.S.[27], but, as discussed with respect to possible farm-to-farm transmission above, estimated AIP occurrence across these regions is generally very low (Fig. 7) and considerable residual error remains (Fig. 4) at these locations. By comparison, AIP probabilities in the vicinity of past H7 events (e.g., H7N3 and H7N9) are considerably higher. This tendency may be a reflection of blue-winged teal being more commonly infected with H7 subtype viruses during spring as compared to H5 viruses[18]. In short, model results provide additional support that blue-winged teal were unlikely acting as a vector of HPAI during the 2014–2015 H5N2 and H5N8 outbreaks across the U.S.[71]; however, findings may add support to the notion that this species could introduce H7 subtype viruses to domestic poultry during the spring migration.

Our study could be improved in at least a couple regards. First, despite having intended our marked blue-winged teal as representative of the larger North American dabbling duck population, we acknowledge that migratory movement behavior may vary even within a species and that our study likely exhibits species-specific sample bias. As discussed above, the tendency for blue-winged teal to be early and long-distance flyers may have important implications for AI disease ecology and the representativeness of the species as an archetypal dabbling duck. We likewise recognize that bird-virus associations are often species-specific due to co-adaptation and variable physiological and immunological tolerances. Studies drawing on a larger sample size or incorporating different duck species under a comparative framework may be better suited to infer waterfowl-AIP relationships over such a heterogeneous study area. Studies aimed at investigating waterfowl-AIP dynamics by either choosing a wider range of focal bird species or being more selective in the virus subtypes examined would certainly be worthwhile. Secondly, the analysis of a large spatial domain may offer insight into long-distance animal movement and virus transmission potential, but, it also restricts data precision and availability. For example, the Gridded Livestock database that was used to estimate poultry abundance was essential to our study, but, had focused on a smaller geographic area in the U.S., we would have gained access to more recent and detailed information to characterize poultry locations and abundances.

In closing, we believe that the spatial and temporal relationship of blue-winged teal to past AIPs offers important insights into migratory waterfowl ecology and avian influenza disease dynamics that can aid in better preparing for future AIPs. Being able to anticipate the geographic locations and times most vulnerable to avian influenza outbreaks would be an asset to the domestic poultry industry and those oversight agencies responsible for disease surveillance and response. Additionally, we hope that the methodological approach presented in this study will motivate future research into modeling disparate facets of disease systems at the wildlife-agriculture interface.

Data availability

All analysis and modeling were performed using the open-source R language for statistical computing[52] with freely-available data. Wild bird telemetry, environmental data, and avian influenza detections can be accessed using web addresses provided in the Materials and Methods Section.

References

1. Alexander, D. J. An overview of the epidemiology of avian influenza. *Vaccine* **25**, 5637–5644, https://doi.org/10.1016/j.vaccine.2006.10.051 (2007).
2. Jones, K. E. *et al.* Global trends in emerging infectious diseases. *Nature* **451**, 990–993, https://doi.org/10.1038/nature06536 (2008).
3. Prosser, D. J. *et al.* Wild bird migration across the Qinghai-Tibetan Plateau: A transmission route for highly pathogenic H5N1. *PLoS One* **6**, https://doi.org/10.1371/journal.pone.0017622 (2011).
4. Prosser, D. J. *et al.* Mapping Avian Influenza Transmission Risk at the Interface of Domestic Poultry and Wild Birds. *Front. Public Heal.* **1**, 1–11, https://doi.org/10.3389/fpubh.2013.00028 (2013).

5. Belkhiria, J., Hijmans, R. J., Boyce, W., Crossley, B. M. & Martínez-López, B. Identification of high risk areas for avian influenza outbreaks in California using disease distribution models. *PLoS One* **13**, e0190824, https://doi.org/10.1371/journal.pone.0190824 (2018).
6. La Sala, L. F. *et al.* Spatial modelling for low pathogenicity avian influenza virus at the interface of wild birds and backyard poultry. *Transboundary Emerg. Dis.* **66**, 1493–1505, https://doi.org/10.1111/tbed.13136 (2019).
7. Campitelli, L. *et al.* Interspecies transmission of an H7N3 influenza virus from wild birds to intensively reared domestic poultry in Italy. *Virology* **323**, 24–36, https://doi.org/10.1016/j.virol.2004.02.015 (2004).
8. Li, L. *et al.* Genetic Evidence Supports Sporadic and Independent Introductions of Subtype H5 Low-Pathogenic Avian Influenza A Viruses from Wild Birds to Domestic Poultry in North America. *J. Virol.* **92**, 1–16, https://doi.org/10.1128/jvi.00913-18 (2018).
9. Li, L., DeLiberto, T. J., Killian, M. L., Torchetti, M. K. & Wan, X. F. Evolutionary pathway for the 2017 emergence of a novel highly pathogenic avian influenza A(H7N9) virus among domestic poultry in Tennessee, United States. *Virology* **525**, 32–39, https://doi.org/10.1016/j.virol.2018.09.003 (2018).
10. Yu, Z., Cheng, K. & Gao, Y. Poultry Infection with Influenza Viruses of Wild Bird Origin, China, 2016. *Emerg. Infect. Dis.* **24**, 1375–1377, https://doi.org/10.1007/s13312-018-1234-3 (2018).
11. Kawaoka, Y., Nestorowicz, A., Alexander, D. J. & Webster, R. G. Molecular analyses of the hemagglutinin genes of H5 influenza viruses: origin of a virulent turkey strain. *Virology* **158**, 218–227, https://doi.org/10.1016/0042-6822(87)90256-X (1987).
12. Munster, V. J. & Fouchier, R. A. Avian influenza virus: Of virus and bird ecology. *Vaccine* **27**, 6340–6344, https://doi.org/10.1016/j.vaccine.2009.02.082 (2009).
13. Claas, E. C. *et al.* Human influenza A H5N1 virus related to a highly pathogenic avian influenza virus. *The Lancet* **351**, 472–477, https://doi.org/10.1016/s0140-6736(97)11212-0 (1998).
14. Chen, Y. *et al.* Human infections with the emerging avian influenza A H7N9 virus from wet market poultry: Clinical analysis and characterisation of viral genome. *The Lancet* **381**, 1916–1925, https://doi.org/10.1016/S0140-6736(13)60903-4 (2013).
15. Olsen, B. *et al.* Global patterns of influenza A virus in wild birds. *Science* **312**, 384–388, https://doi.org/10.1126/science.1122438, arXiv:1011.1669v3 (2006).
16. González-Reiche, A. S. *et al.* Influenza a viruses from wild birds in Guatemala belong to the North American lineage. *PLoS One* **7**, 1–11, https://doi.org/10.1371/journal.pone.0032873 (2012).
17. Lu, L., Lycett, S. J. & Brown, A. J. Determining the phylogenetic and phylogeographic origin of highly pathogenic avian influenza (H7N3) in Mexico. *PLoS One* **9**, https://doi.org/10.1371/journal.pone.0107330 (2014).
18. Ramey, A. M. *et al.* Evidence for Seasonal Patterns in the Relative Abundance of Avian Influenza Virus Subtypes in Blue-Winged Teal (Anas discors). *J. Wildl. Dis.* **50**, 916–922, https://doi.org/10.7589/2013-09-232 (2014).
19. Nallar, R. *et al.* Ecological Determinants of Avian Influenza Virus, West Nile Virus, and Avian Paramyxovirus Infection and Antibody Status in Blue-Winged Teal (Anas Discors) in the Canadian Prairies. *J. Wildl. Dis.* **52**, 33–46 (2016).
20. Lee, D. *et al.* Intercontinental Spread of Asian-Origin H5N8 to North America through Beringia by Migratory Birds. *J. Virol.* **89**, 6521–6524, https://doi.org/10.1128/JVI.00728-15 (2015).
21. Ip, H. *et al.* Novel Eurasian highly pathogenic avian influenza A H5 viruses in wild birds, Washington, USA. *Emerg. Infect. Dis.* **21**, 886–890 (2014).
22. Ramey, A. M. *et al.* Evidence for common ancestry among viruses isolated from wild birds in beringia and highly pathogenic intercontinental reassortant H5N1 and H5N2 influenza a viruses. *Infect. Genet. Evol.* **40**, 176–185, https://doi.org/10.1016/j.meegid.2016.02.035 (2016).
23. Lee, D. *et al.* Highly pathogenic avian influenza viruses and generation of novel reassortants, United States, 2014-2015. *Emerg. Infect. Dis.* **22**, 1283–1285, https://doi.org/10.3201/eid2207.160048 (2016).
24. Bevins, S. N. *et al.* Widespread detection of highly pathogenic H5 influenza viruses in wild birds from the Pacific Flyway of the United States. *Sci. Reports* **6**, 1–9, https://doi.org/10.1038/srep28980 (2016).
25. Ramey, A. M. *et al.* Surveillance for highly pathogenic influenza A viruses in California during 2014-2015 provides insights into viral evolutionary pathways and the spatiotemporal extent of viruses in the Pacific Americas Flyway. *Emerg. Microbes Infect.* **6**, https://doi.org/10.1038/emi.2017.66 (2017).
26. Grear, D. A., Hall, J. S., Dusek, R. J. & Ip, H. S. Inferring epidemiologic dynamics from viral evolution: 2014–2015 Eurasian/North American highly pathogenic avian influenza viruses exceed transmission threshold, R0 = 1, in wild birds and poultry in North America. *Evol. Appl.* **11**, 547–557, https://doi.org/10.1111/eva.12576 (2018).
27. Greene, J. L. Update on the highly pathogenic avian influenza outbreak of 2014–2015. *Congr. Res. Serv. Rep. R44114* (2015).
28. Dargatz, D., Beam, A., Wainwright, S. & McCluskey, B. Case Series of Turkey Farms from the H5N2 Highly Pathogenic Avian Influenza Outbreak in the United States During 2015. *Avian Dis.* **60**, 467–472, https://doi.org/10.1637/11350-121715-reg (2016).
29. Prosser, D. J. *et al.* Satellite-marked waterfowl reveal migratory connection between H5N1 outbreak areas in China and Mongolia. *Ibis* **151**, 568–576, https://doi.org/10.1111/j.1474-919X.2009.00932.x (2009).
30. Takekawa, J. Y. *et al.* Migration of Waterfowl in the East Asian Flyway and Spatial Relationship to HPAI H5N1 Outbreaks. *Avian Dis. Dig.* **5**, e101–e102, https://doi.org/10.1637/9187-891409-DIGEST.1, 15334406 (2010).
31. Yamaguchi, N., Hupp, J. W., Higuchi, H., Flint, P. L. & Pearce, J. M. Satellite-tracking of Northern Pintail *Anas acuta* during outbreaks of the H5N1 virus in Japan: implications for virus spread. *Ibis* **152**, 262–271 (2010).
32. Takekawa, J. Y. *et al.* Movements of wild ruddy shelducks in the Central Asian flyway and their spatial relationship to outbreaks of highly pathogenic avian influenza H5N1. *Viruses* **5**, 2129–2152, https://doi.org/10.3390/v5092129 (2013).
33. Sullivan, J. D. *et al.* Waterfowl Spring Migratory Behavior and Avian Influenza Transmission Risk in the Changing Landscape of the East Asian-Australasian Flyway. *Front. Ecol. Evol.* **6**, 1–14, https://doi.org/10.3389/fevo.2018.00206 (2018).
34. Pepin, K. M. *et al.* Using quantitative disease dynamics as a tool for guiding response to avian influenza in poultry in the United States of America. *Prev. Vet. Medicine* **113**, 376–397, https://doi.org/10.1016/j.prevetmed.2013.11.011 (2014).
35. Nallar, R. *et al.* Demographic and spatiotemporal patterns of avian influenza infection at the continental scale, and in relation to annual life cycle of a migratory host. *PLoS One* **10**, 1–14 (2015).
36. Papp, Z. *et al.* The ecology of avian influenza viruses in wild dabbling ducks (Anas spp.) in Canada. *PLoS One* **12**, 1–18, https://doi.org/10.1371/journal.pone.0176297 (2017).
37. Gonzalez-Reiche, A. S. & Perez, D. R. Where do avian influenza viruses meet in the americas? *Avian Dis.* **56**, 1025–1033–9, https://doi.org/10.1637/10203-041412-Reg.1 (2012).
38. Spackman, E., McCracken, K. G., Winker, K. & Swayne, D. E. H7n3 avian influenza virus found in a south american wild duck is related to the chilean 2002 poultry outbreak, contains genes from equine and north american wild bird lineages, and is adapted to domestic turkeys. *J. virology* **80**, 7760–7764, https://doi.org/10.1128/JVI.00445-06 (2006).
39. Humphreys, J. M., Murrow, J. L., Sullivan, J. D. & Prosser, D. J. Seasonal occurrence and abundance of dabbling ducks across the continental united states: Joint spatio-temporal modelling for the genus anas. *Divers. Distributions* **25**, 1497–1508, https://doi.org/10.1111/ddi.12960 (2019).
40. Hanson, B. A. *et al.* Avian influenza viruses and paramyxoviruses in wintering and resident ducks in Texas. *J. Wildl. Dis.* **41**, 624–628, https://doi.org/10.7589/0090-3558-41.3.624 (2005).
41. Stallknecht, D. E. & Brown, J. D. Wild birds and the epidemiology of avian influenza. *J. Wildl. Dis.* (2007).
42. Kulldorff, M. & Nagarwalla, N. Spatial Disease Clusters: Detection and INference. *Stat. Medicine* **14**, 799–810 (1995).

43. Aamodt, G., Samuelsen, S. O. & Skrondal, A. A simulation study of three methods for detecting disease clusters. *Int. J. Heal. Geogr.* **5**, 1–11, https://doi.org/10.1186/1476-072X-5-15 (2006).

44. Van Dijk, J. G. *et al*. Juveniles and migrants as drivers for seasonal epizootics of avian influenza virus. *J. Animal Ecol.* **83**, 266–275, https://doi.org/10.1111/1365-2656.12131, NIHMS150003 (2014).

45. Alexander, D. J. A review of avian influenza in different bird species. *Veterinary microbiology. Vet. Microbiol.* **1**, 3–13 (2000).

46. Takekawa, J. Y. *et al*. Migration of Waterfowl in the East Asian Flyway and Spatial Relationship to HPAI H5N1 Outbreaks. *Avian Dis.* **54**, 466–476 (2010).

47. Douglas, D. C. *et al*. Moderating Argos location errors in animal tracking data. *Methods Ecol. Evol.* **3**, 999–1007, https://doi.org/10.1111/j.2041-210X.2012.00245.x (2012).

48. Ramey, A. *et al*. Tracking data for Blue-winged Teal (*Anas discors*), U.S. Geological Survey data release, https://doi.org/10.5066/P9Z9BA9F (2019).

49. Hengl, T. *et al*. Soilgrids250m: Global gridded soil information based on machine learning. *PLoS One* **12**, 1–40, https://doi.org/10.1371/journal.pone.0169748 (2017).

50. Title, P. O. & Bemmels, J. B. Envirem: an expanded set of bioclimatic and topographic variables increases flexibility and improves performance of ecological niche modeling. *Ecography* 291–307, https://doi.org/10.1111/ecog.02880 (2017).

51. Dormann, C. F. *et al*. Collinearity: A review of methods to deal with it and a simulation study evaluating their performance. *Ecography* **36**, 027–046, https://doi.org/10.1111/j.1600-0587.2012.07348.x, NIHMS150003 (2013).

52. R Core Team. *R: A Language and Environment for Statistical Computing*. (R Foundation for Statistical Computing, Vienna, Austria, 2017).

53. Baddeley, A. & Turner, R. spatstat: An R Package for Analyzing Spatial Point Patterns. *J. Stat. Softw.* **12**, 1–42 (2005).

54. Gilbert, M. *et al*. Global distribution data for cattle, buffaloes, horses, sheep, goats, pigs, chickens and ducks in 2010. *Sci. Data* **5**, 1–11, https://doi.org/10.1038/sdata.2018.227 (2018).

55. CIESIN. Gridded Population of the World, Version 4 (GPWv4): Population Density Grid (2005).

56. Krainski, E. T. *et al. Advanced Spatial Modeling with Stochastic Partial Differential Equations Using R and INLA*. 1 edn. (Chapman and Hall/CRC, New York, 2018).

57. Blangiardo, M. & Cameletti, M. *Spatial and Spatio-temporal Bayesian Models with R-INLA*. (John Wiley & Sons, Inc., Chichester, UK, 2015).

58. Lindgren, F. & Rue, H. Bayesian Spatial Modelling with R-INLA. *J. Stat. Softw.* **63**, 1–25 (2015).

59. Lawson, A. B. Bayesian point event modeling in spatial and environmental epidemiology. *Stat. Methods Med. Res.* **21**, 509–529, https://doi.org/10.1177/0962280212446328, 1706.09079 (2012).

60. Sumner, M. D. *trip: Tools for the Analysis of Animal Track Data*, R package version 1.5.0 (2016).

61. Simpson, D., Illian, J., Lindgren, F., Sørbye, S. H. & Rue, H. Going off grid: Computationally efficient inference for log-Gaussian Cox processes. *Biometrika* **103**, 49–70, https://doi.org/10.1093/biomet/asv064, arXiv:1111.0641v1 (2015).

62. Fuglstad, G.-A., Simpson, D., Lindgren, F. & Rue, H. Constructing priors that penalize the complexity of gaussian random fields. *J. Am. Stat. Assoc.* **0**, 1–8, https://doi.org/10.1080/01621459.2017.1415907 (2018).

63. Rue, H., Martino, S. & Chopin, N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *J. Royal Stat. Soc. Ser. B: Stat. Methodol.* **71**, 319–392, https://doi.org/10.1111/j.1467-9868.2008.00700.x (2009).

64. Freeman, E. A. & Moisen, G. Presence Absence: An r package for presence absence analysis. *J. Stat. Softw.* **23**, 1–31, Eafreeman@fs.fed.us (2008).

65. Allouche, O., Tsoar, A. & Kadmon, R. Assessing the accuracy of species distribution models: Prevalence, kappa and the true skill statistic (TSS). *J. Appl. Ecol.* **43**, 1223–1232, https://doi.org/10.1111/j.1365-2664.2006.01214.x (2006).

66. Watanabe, S. A Widely Applicable Bayesian Information Criterion. *Mach. Learn. Res.* **14**, 867–897, 1208.6338 (2013).

67. Gelman, A., Hwang, J. & Vehtari, A. Understanding predictive information criteria for Bayesian models. *Stat. Comput.* **24**, 997–1016, https://doi.org/10.1007/s11222-013-9416-2 (2014).

68. Humphreys, J., Elsner, J., Jagger, T. & Mahjoor, A. Disaggregating the Patchwork: Probabilistic models as tools to predict wetland presence as a continuous gradient. *Wetlands* **37**, 205–219, https://doi.org/10.1007/s13157-016-0859-z (2017).

69. Lee, D.-H. *et al*. Transmission Dynamics of Highly Pathogenic Avian Influenza. *Emerg. Infect. Dis.* **24**, 1840–1848, https://doi.org/10.3201/eid2410.171891 (2018).

70. Jennelle, C. *et al*. Surveillance for highly pathogenic avian influenza virus in wild birds during outbreaks in domestic poultry, Minnesota, 2015. Emerging infectious diseases. *Emerg. Infect. Dis.* **22**, 1278–1282 (2016).

71. Krauss, S. *et al*. The enigma of the apparent disappearance of Eurasian highly pathogenic H5 clade 2.3.4.4 influenza A viruses in North American waterfowl. *Proc. Natl. Acad. Sci.* **113**, 9033–9038, https://doi.org/10.1073/pnas.1608853113 (2016).

72. Wells, S. J. *et al*. Epidemiologic Investigation of Highly Pathogenic H5N2 Avian Influenza Among Upper Midwest U.S. Turkey Farms, 2015. *Avian Dis.* **61**, 198–204, https://doi.org/10.1637/11543-112816-reg.1 (2017).

73. Tobler, W. A computer movie simulating urban growth in the Detroit region. *Econ. Geogr.* **46**, 234–240 (1970).

74. Stallknecht, D. E., Shane, S. M., Zwank, P. J., Senne, D. A. & Kearney, M. T. Avian influenza viruses from migratory and resident ducks of coastal louisiana. *Avian Dis.* **34**, 398–405 (1990).

75. Ferro, P. J. *et al*. Multiyear surveillance for avian influenza virus in waterfowl from wintering grounds, Texas Coast, USA. *Emerg. Infect. Dis.* **16**, 1224–1230, https://doi.org/10.3201/eid1608.091864 (2010).

Acknowledgements

Author contributions

D.P., A.R. and J.M. conceived the study concept; J.H. led and D.D., A.R. and D.P. contributed to design of the methodology and data analysis. A.R., C.S., P.L. and P.W. led wild bird capture and marking efforts. All authors contributed to writing the manuscript and gave final approval for publication.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to J.M.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.