Linking hosts, landscapes, and climate to advance zoonotic arbovirus forecasting

V.A. Akshay [1,2], J. Alex Baecher [2,3], Nathan Burkett-Cadena [3,4], James Thorson [5], Yesenia Sánchez [3], Yasmin Tavares [6], Amely Bauer [7,8], Robert Guralnick [2], Lindsay Campbell [3,4]

[1] School of Natural Resources and Environment, IFAS, University of Florida, Gainesville, Florida 32611, United States of America

[2] Florida Natural History Museum, University of Florida, Gainesville, Florida 32611, United States of America

[3] Florida Medical Entomology Laboratory, IFAS, University of Florida, Vero Beach, Florida 32962, United States of America

[4] Department of Entomology and Nematology, IFAS, University of Florida, Gainesville, Florida 32611, United States of America

[5] Resource Ecology and Fisheries Management, Alaska Fisheries Science Center, Seattle, WA 98115, United States of America

[6] Department of Wildlife Ecology and Conservation, IFAS, University of Florida, Gainesville, Florida 32611

[7] Center for Advanced Systems Understanding - CASUS, Helmholtz-Zentrum Dresden-Rossendorf, D-02826 Görlitz, Germany

[8] Department of Community Ecology, Helmholtz Centre for Environmental Research - UFZ, D-06120 Halle (Saale), Germany

Corresponding authors Lindsay Campbell (email: lcampbell2@ufl.edu) and Robert Guralnick (email: rguralnick@flmnh.ufl.edu)

Abstract

Forecasting zoonotic mosquito-borne viruses remains a critical challenge because transmission depends on dynamic, multitrophic interactions among vectors, hosts, pathogens, and the environment. Here, we integrate long-term sentinel chicken surveillance

across much of Florida with environmental data to build a predictive framework for eastern equine encephalitis virus (EEEV), a zoonotic mosquito borne disease of concern to human and equine health. Our models captured both environmental drivers and latent spatiotemporal structure, achieving strong predictive accuracy. Models revealed strong nonlinear effects of moderate precipitation a year prior to sampling and higher minimum temperature 1 month prior to sampling, as well as moderate and high percentages of forest and wetland cover on increased EEEV seroconversion. Retrospective predictions showed shifting virus activity across regions, consistent with *Culiseta melanura* mosquito vector ecology. We also calculated associations between EEEV and abundance estimates for key bird species that are suspected virus hosts using eBird data. Seasonal shifts among migratory and resident birds with predicted virus activity for key species suspected of being important EEEV hosts suggests spring migrants play a role in amplification, residents in summer persistence, and overwintering groups as potential reservoirs. These results demonstrate ecological forecasting of arboviruses is feasible at management-relevant scales, with broad potential to extend to other arbovirus systems. By integrating traditional surveillance with community science, our framework advances both predictive capacity and ecological understanding of zoonotic arboviruses.

Zoonotic mosquito-borne diseases pose a major global health threat, with capacity to rapidly spread to new areas (1). Transmission in these systems is the result of multitrophic interactions among hosts, vectors, and pathogens, shaped by both intrinsic factors (e.g., demographics, vector competence) and extrinsic environmental conditions (e.g., climate, habitat, and biotic interactions) (2, 3). In vector-borne diseases (VBDs) maintained between avian hosts and mosquito vectors, host movement and migration, seasonal phenology of mosquito populations, and their overlap determine pathogen maintenance, amplification, and spillover (4). These systems often include multiple hosts and vectors, and anticipating and predicting transmission risks to humans remains a wicked problem impacting disease prevention (5–7). Rather than proactively assessing where risks might be highest, most disease prevention efforts are reactive, attempting to mitigate outbreaks already in progress (8, 9). Because disease system components are difficult to measure fully across space and time, developing predictive capacity requires intensive monitoring and modeling frameworks that can integrate and harmonize data from multiple sources, while capturing the spatiotemporal structure of unobserved processes driving virus activity (10). Such an integrated approach allows incorporation of new data streams while maintaining the potential to scale prediction and forecasting across broader areas.

Eastern equine encephalitis virus (EEEV; *Alphavirus*, *Togaviridae*) is a zoonotic mosquito borne disease system maintained in an enzootic transmission cycle between the primary mosquito vector, *Culiseta melanura*, and multiple avian host species with occasional spillover transmission to humans and horses (11) (12–15). EEEV in humans is relatively uncommon, but consequences can be severe, including long lasting neurological problems, and EEEV has the highest mortality rate of any arbovirus in the U.S. at ~30% in those developing neuroinvasive disease (16, 17). The virus has a broad geographic distribution spanning eastern North America, the Caribbean, Central America, and portions of South America (18–20) with some distributional overlap with closely related Madariaga virus in South and Central America, which primarily causes disease in equines (21). In the northern part of its distribution, the virus continues to expand its range (13), leading to its designation

as an emerging infectious disease, as well as a select agent due to its potential use as a bioweapon (22). Despite the growing frequency and intensity of EEEV outbreaks (13), anticipating the distribution and dynamics of EEEV activity remains a persistent challenge, and the ability to proactively forecast risk is non-existent.

In Florida, EEEV is enzootic with seasonal spillover to equines, and occasional cases in humans and other animals, including emus (23). Peninsular Florida, in particular, may impact EEEV transmission ecology more broadly, given that it is part of a major intercontinental migratory bird flyway with extensive stopover habitats, which could serve as an important source of virus dispersal (23, 24). To monitor for EEEV, the Florida Department of Health (FDOH), in partnership with mosquito control programs, maintains one of the longest-running arbovirus surveillance programs in the U.S., monitoring sentinel chickens across hundreds of coops statewide (25). This spatially extensive, long-term dataset provides a data basis for integrating long-term surveillance with biotic and abiotic information to predict virus activity and transmission hazard. Building a more predictive framework will not only strengthen early warning systems, but also deepen fundamental insights into multitrophic disease dynamics and enhance ability to proactively forecast zoonotic vector-borne disease risk.

Our approach in this work leverages new advances in spatiotemporal modeling that can capture both measured environmental correlations and spatiotemporal structure in unmeasured variables to improve model predictions (26, 27). Beyond improved prediction, model outputs may reveal patterns that can be compared with complementary data streams, including community science observations (i.e. eBird), to correlate suspected host dynamics with patterns of virus activity. These factors are particularly relevant for EEEV where long-standing questions about avian host migration and overwintering dynamics in Florida and their connection to the broader spatiotemporal ecology of the system remain unresolved (23).

Here we: 1) develop spatiotemporal predictive models of monthly EEEV sentinel chicken seroconversion across Florida between 2005 and 2019 and 2) examine correlations

between estimated values of avian abundances from community science eBird data and predicted EEEV activity between 2005 and 2019 for 12 resident and migratory species that are suspected avian hosts. Although primarily interested in predicting virus activity with statistical tools, our multitrophic approach brings needed realism about host distributions and abundances, which are most often unmeasured, particularly at this scale. We expect higher seroconversion rates at locations surrounded by greater forest and wetland areas, given known habitat preferences of *Culiseta melanura*, and we also expect greater seroconversion with higher lagged precipitation values based on previous studies in the northeastern U.S. We also expect lower seroconversion with very high temperatures based on models of West Nile virus in Florida (10). Finally, we expect greater abundances of spring migratory avian hosts preceding elevated predictions of EEEV seroconversion by 2 - 3 months, while resident species will show peak correlations with predicted seroconversion during the breeding summer season. Lastly, we expect Fall migrants and overwintering species will show strong contemporaneous associations over the winter as possible reservoirs of the virus.

Results

Initial model results: We compiled and analyzed Florida Department of Health EEEV sentinel chicken seroconversion surveillance data spanning two decades, developed spatiotemporal models to predict transmission hazard across Florida's diverse landscapes and identified associations between estimated avian abundances and phenology and predicted virus activity. We start first with our abiotic-focused models. After systematically removing collinear variables evaluated using variance inflation factors (VIF > 5), and performing a stepwise backward selection, our final model incorporated 8 key abiotic predictors: lagged cumulative precipitation (1, 5, and 12 months), lagged maximum temperature (6 and 12 months), minimum temperature (1 month), and percentage forest and wetland land cover (Table 1).

Table 1: Description of the predictor variables used to model spatiotemporal EEEV seroconversion proportions in the state of Florida (USA). Terms included in this table result from a model selection and variable reduction procedure using spatial models (see methods). Descriptions and parameterization details are provided for each variable.

Effect	Term	Description	Variable type	Parameterization
Fixed	*prcp_lag1*	Cumulative precipitation (1 mo. lag)	Continuous	2nd order orthogonal polynomial
	prcp_lag5	Cumulative precipitation (5 mo. lag)	Continuous	2nd order orthogonal polynomial
	prcp_lag12	Cumulative precipitation (12 mo. lag)	Continuous	2nd order orthogonal polynomial
	tmax_lag6	Max temperature (6 mo. lag)	Continuous	2nd order orthogonal polynomial
	tmax_lag12	Max temperature (12 mo. lag)	Continuous	2nd order orthogonal polynomial
	tmin_lag1	Min temperature (1 mo. lag)	Continuous	2nd order orthogonal polynomial
	forest	Proportion forest cover	Continuous	2nd order orthogonal polynomial
	wetlands	Proportion wetlands	Continuous	2nd order orthogonal polynomial
	county	County name	39-level factor	Intercept nested above site_id

Random	*site_id*	Monitoring site name	476-level factor	Intercept nested below county

The spatiotemporal framework incorporating GMRFs substantially outperformed all alternative model formulations. This model showed strong temporal autocorrelation (AR1 ρ = 0.73), with spatial correlation extending to 82.28 km (Matérn range), and substantial spatiotemporal variability (marginal SD = 1.75), indicating spatiotemporal structure in EEEV seroprevalence beyond that explained by covariates. The best GMRF model explained 37.7% of deviance with an AIC of 6164.1, compared to 29.9% and an AIC of 6920.7 for the best performing non-spatial model (Table S1.1).

Model validation: We withheld the last 2 years of surveillance data (2018-2019) for model validation. That validation showed strong predictive performance accurately capturing the magnitude and timing of EEEV activity (overall RMSE = 0.007457; Table S1.2) in both out years. Errors remained low across months (1.88×10^{-5} to 0.015352), with predictions consistently within empirical confidence intervals across most months (Figure 1). Performance was slightly better in 2019 (RMSE = 0.007241) than in 2018 (0.007668), but both out years demonstrated potential for operational forecasting.

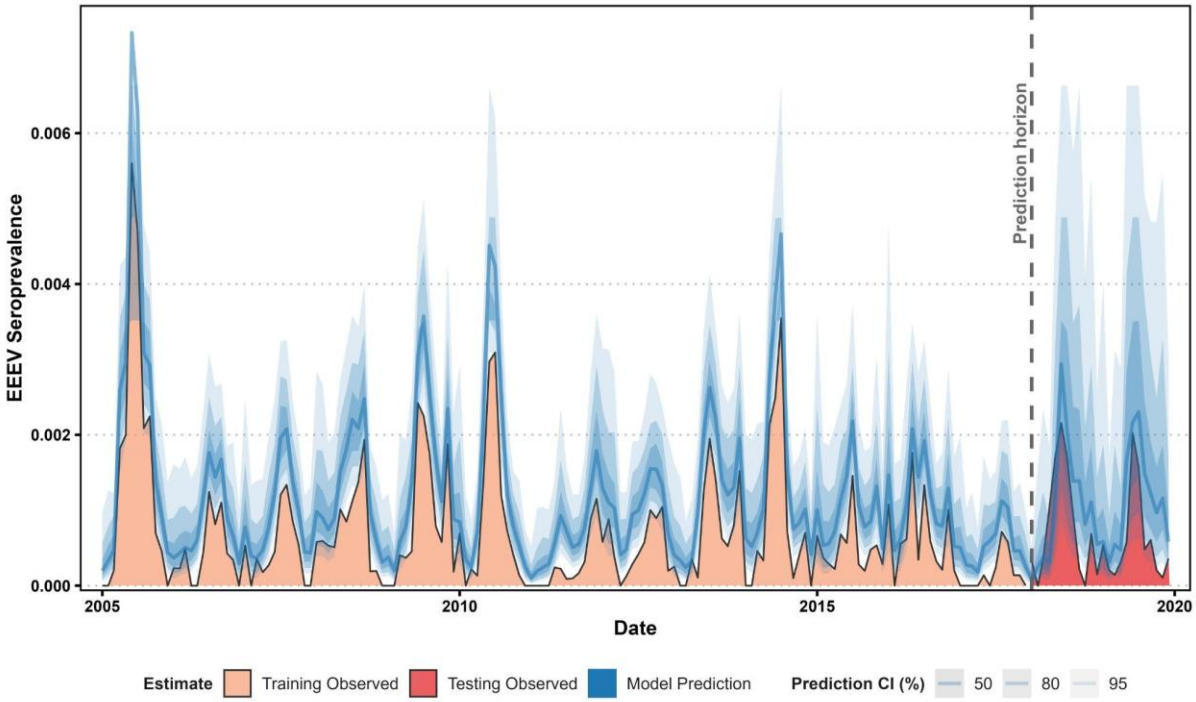


Figure 1: Temporal trends of EEEV seroprevalence in Florida (2005–2019) from empirical data aggregated at the monthly level (orange = training, red = testing) and model predictions (blue). The model was trained on 2005–2017 data and evaluated using 2018–2019 out-of-sample predictions (dotted vertical line). Model predictions are presented as point estimates (blue line) and 50, 80, 90, and 95 % confidence intervals (opaque blue ribbons).

Environmental determinants of and spatiotemporal trends in EEEV activity: Environmental response curves (Figure 2) revealed complex, non-linear relationships with virus activity. The model identified strong precipitation effects, with intermediate prior-year rainfall (12-month lag) predicting elevated seroconversion. Moderate forest cover was the strongest landscape predictor, showing positive effects that decreased at higher coverage, while wetlands also had strong positive associations and plateaued at highest coverage, which are both consistent with known EEEV vector and host ecology (13, 28). For the full set of predictors and parameter estimates see Table S1.3. Parallel analyses conducted on the second candidate model set yielded qualitatively similar results, with detailed results provided in Supplementary Materials (Table S1.4).

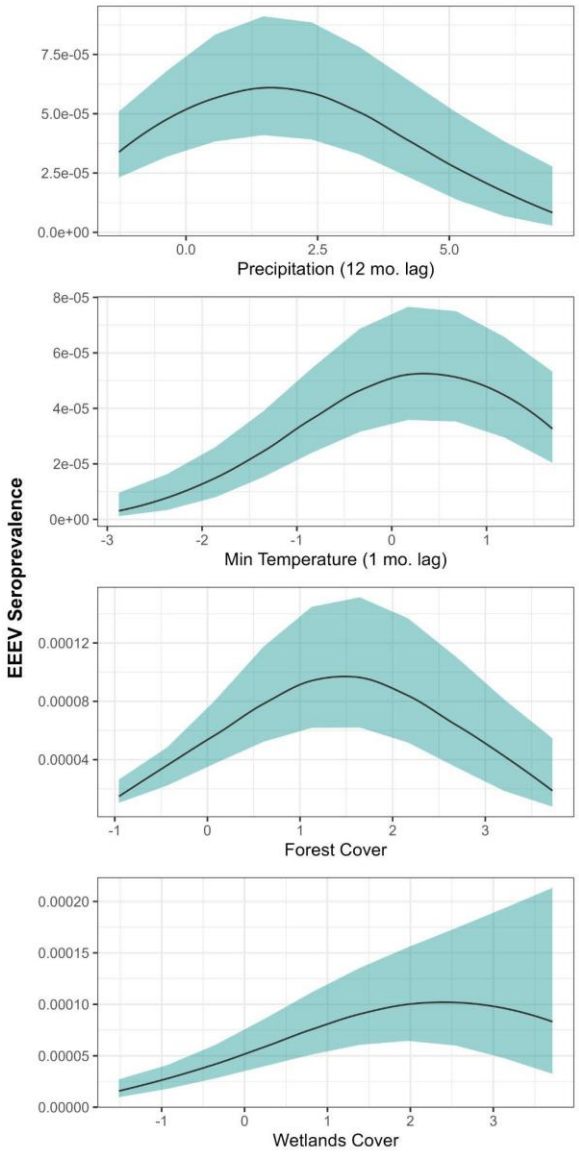


Figure 2: Marginal response plots of predicted Eastern Equine Encephalitis Virus (EEEV) seroprevalence from a spatiotemporal model of Florida sentinel chicken surveillance (2005–2019) aggregated at the monthly-level. Environmental covariates were measured contemporaneously or with 1-, 5-, 6-, or 12-month lags, as indicated by variable names.

Statewide retrospective predictions revealed distinct temporal patterns in EEEV transmission dynamics over the study period in Florida (Figure 3), with peak statewide activity in 2005, spanning the Panhandle, Gulf Coast, and north-central regions. Predicted virus activity declined through 2007 before increasing from 2008-2010. A second decline during 2011-2012 was followed by renewed increases from 2013-2015, with activity shifting toward central Florida. Monthly predictions provided finer resolution insights into seasonal patterns, revealing summer transmission peaks from June through August alongside persistent low-

level activity year-round (Figure S1.1). The model's highest predicted virus activity occurred during June 2005 across the Gulf Coast and extensive areas of central and north-central Florida. Prediction uncertainty was greatest in regions with sparse sentinel chicken surveillance coverage and mean seroprevalence was predicted highest along forested wetland areas across the Panhandle and northern Florida (Figure S1.2).

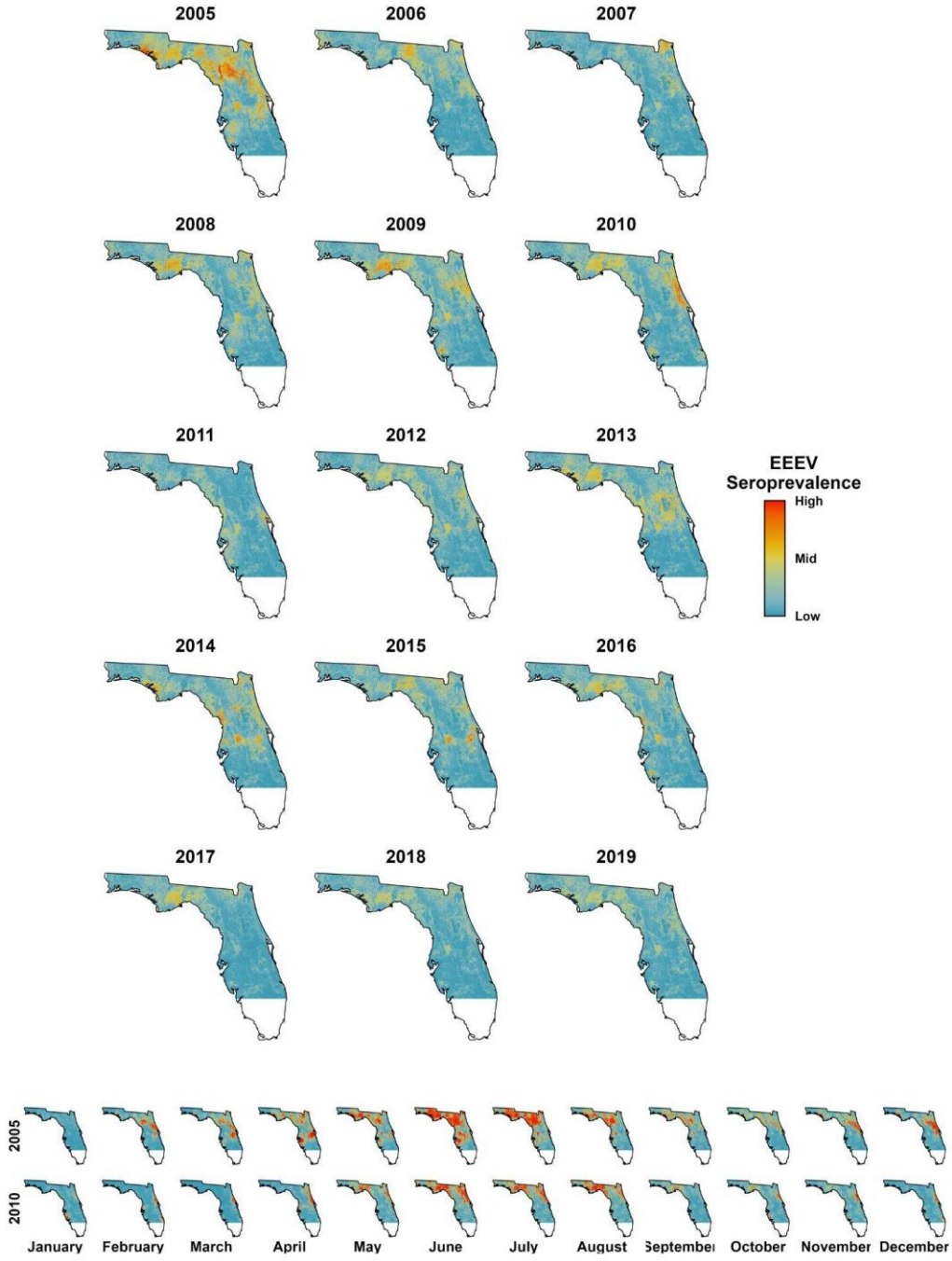


Figure 3: Statewide predictions of EEEV seroprevalence from Florida sentinel chicken surveillance system aggregated to year from 2005 - 2019 and monthly predictions for 2005

and 2010 All predictions are provided with a quantile truncation for ease of visualization at the 98th percentile of values.

Associations of viral activity with migratory and resident birds: Associational analyses between predicted seroconversion rates and abundances of suspected avian hosts revealed marked seasonal patterns across the avian community, in directions expected by migratory status. Spring migrants showed particularly strong associations with predictions of elevated transmission hazard, led by the Red-eyed Vireo (r = 0.58, July across a 3-month lag, Figure 4). Year-round residents exhibited varied responses, with Pine Warbler showing peak associations with EEEV seropositive rates during spring breeding season (r = 0.58, May) while Northern Cardinal peaked during summer months (r = 0.36, July). Common Yellowthroat, another resident species, demonstrated moderate spring associations (r = 0.38, May), suggesting shared seasonal patterns among certain resident taxa (Figure S1.3). Winter migrants presented a more complex picture, with American Robin showing strong late-season correlations (r = 0.57, November) and Yellow-rumped Warbler peaking in December (r = 0.45). Notably, several species exhibited negative or very weak correlations with predicted virus activity, including Green Heron (r = -0.37, August), Black-crowned Night Heron (r = -0.31, July), and Wood Thrush (r = -0.12, December across a 3-month lag).

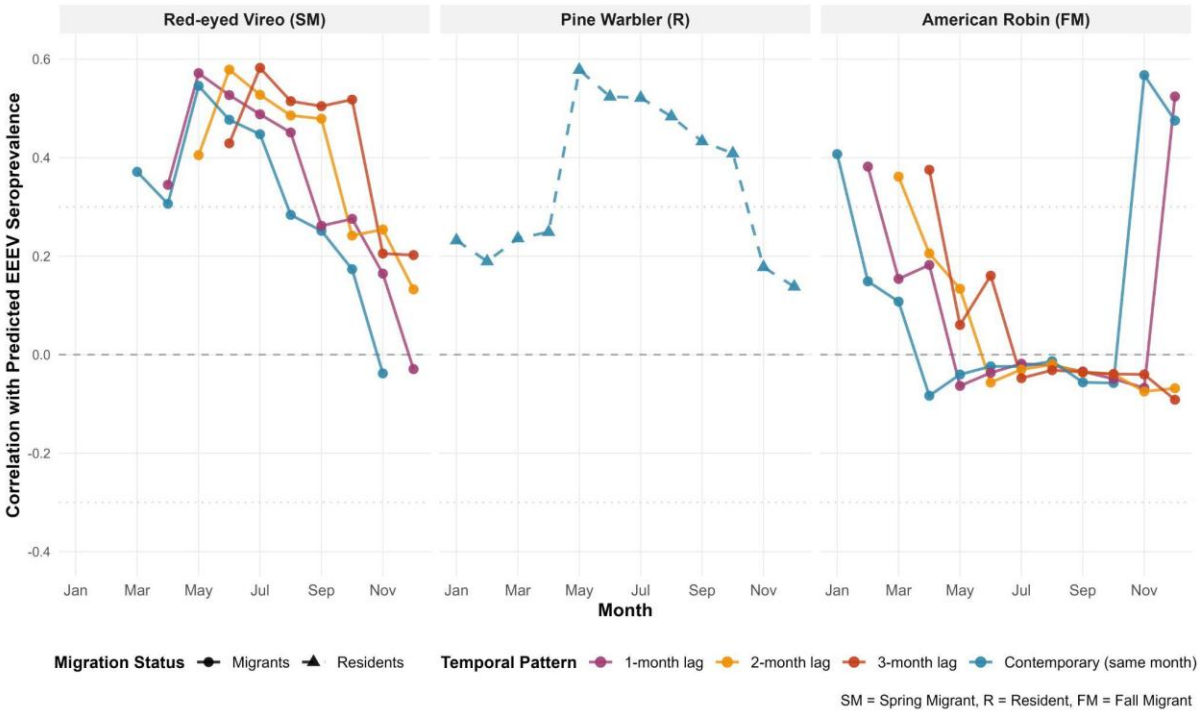


Figure 4: Monthly pixelwise Pearson correlations (y-axis) between predicted EEEV seroprevalence (2005–2019) and bird abundance across seasons (x-axis) for three species with the strongest associations (r ~ 0.60). Monthly predicted seroconversion was averaged across years, and eBird Status and Trends data provided monthly species-specific relative abundance for a "typical year". For resident species, we assessed contemporaneous correlations; for migratory species, we also calculated 1–3 month lags to observe potential delays in virus amplification. Correlations for all 12 species are shown in Figure S1.3.

Discussion

Here we demonstrate the capacity to forecast near-term EEEV activity across a regional scale, while more fully accounting for abiotic and biotic factors underlying transmission dynamics. This step forward is enabled by integrating four separate key elements: 1) A broadly sampled multi-site, multi-year time-series of EEEV activity measured across most of Florida; 2) A set of statistical models that can appropriately account for spatiotemporal structure in unmeasured variables, thus providing needed information about lag effects that underlie effective near term forecasting; 3) Knowledge of the habitat preferences of the main vector species that transmit EEEV that can be incorporated into models; 4) Knowledge of the key bird species that likely serve as EEEV hosts, whose relative abundances are now available at scale through harmonized citizen science data collected by an avid volunteer

234 birding community. We bring all these resources and advancements into one framework that
235 showcases underlying processes and the capacity to build operational forecasts of
236 transmission hazard, and we establish a foundation to further examine both migratory and
237 resident host species in disease dynamics across this broadscale system.
238
239 *Landscape and climate determinants of viral activity:* As expected, moderate forest cover
240 and extensive wetlands were strongly associated with elevated seroconversion, consistent
241 with the known hardwood swamp habitat of *Culiseta melanura* vectors (28) and patterns
242 reported for equine cases and other landscape studies in Florida and Connecticut (28–32).
243 This finding further underscores forest and wetland as important habitats in the EEEV
244 system, linking mosquito habitat preferences with virus activity. As well, these findings have
245 potential implications for human exposure through ongoing development, as well as
246 restoration efforts, which have been suggested in EEEV emergence in the northeastern U.S.
247 (33). Consistent findings of these landscape predictors across disparate regions highlights
248 their potential to scale beyond regional study areas and generalize across the full
249 geographic range of EEEV. This finding is in sharp contrast to West Nile virus, another
250 mosquito vector and avian host system, where consistent landscape predictors have
251 remained elusive in Florida and other regions (10, 31, 34, 35). This work points to the
252 potential for targeted mosquito abatement and control activities near forest and wetland
253 habitats as a means to reduce disease risk.
254 We also found elevated seroconversion was associated with mild minimum
255 temperatures 1 month prior to sampling and intermediate precipitation levels 12 months prior
256 to sampling. Warmer minimum temperatures in the time period immediately preceding
257 seroconversion likely promotes increased mosquito survival and activity, which may sustain
258 virus circulation and enhance transmission in the active season (3, 36) because mosquito
259 development, metabolism, and viral replication all accelerate with warmer, but not extreme,
260 temperatures. Intermediate levels of precipitation 12 months prior to sampling can also
261 impact mosquito abundance and survival needed to maintain virus circulation during the

following winter or spring months. Although (28) found that extremely wet hydrological conditions 8 months prior to the EEEV season was associated with a greater number of positive mosquito pools in the northeastern U.S., extreme precipitation can also result in larval and egg flushing of mosquitoes (37) and too dry conditions can reduce overall habitat availability, and impact mosquito activity, survival, and abundances (38).

Overall, climate variables at distinct time scales, combined with landscape data, offer complementary pathways for anticipating and ultimately forecasting elevated EEEV activity. We show here that predictions for out years generally match the magnitude and timing of EEEV activity for hold out data. Of particular note is the ability of these models to performantly predict EEEV based on calibration and validation data, despite a relatively low seroconversion rate. A 12-month precipitation lag provides long lead times to guide strategic planning and resource allocation, while contemporaneous minimum temperature aligns with the short-term horizons of existing climate forecasts to refine local response. Together, these predictors bridge seasonal to near-real-time scales, creating a foundation for operational early-warning systems that can target surveillance and control across high-risk habitats.

Retrospective predictions

Retrospective monthly predictions and annual summaries of EEEV activity provide a comprehensive view across the state and provided the first tool to observe dynamic distributions of this system in this region. A key time period of elevated EEEV activity was in 2005 when 50 counties reported virus activity, including 5 human cases and 3 deaths (39). Our models predicted elevated EEEV activity in June, July, and August 2005 in Gadsden and Leon Counties in the Florida Panhandle, in Suwannee County in north central Florida, as well as Pasco County in central Florida, where human cases occurred in July and August (39). Additional retrospective predictions demonstrate variation in monthly spatiotemporal dynamics of virus activity across the state (Figure S1.1). Outputs have the potential to provide new insights into the ecology of disease system dynamics, including the role of early

season transmission in later season amplification and how predicted virus activity links to downstream human and equine transmission risk.

A framework for integrating host dynamics and for process-oriented forecasts

A second element of our framework was integrating eBird community science data to explore associations between avian phenology and abundances and predicted EEEV seroconversion. The results reveal striking seasonal patterns that align with hypothesized host roles where spring migrants may introduce or amplify low-level circulation, resident species appear to sustain summer transmission, and overwintering migrants may maintain virus activity during cooler months (23). Red-eyed Vireos, which arrive in March, and to a lesser extent early season Hermit Thrush, were both correlated with peak transmission in May–July, consistent with the possibility of low level early season amplification followed by amplification during the breeding season. Pine Warbler and other resident songbirds showed positive correlations only during the breeding season, supporting their suspected role as amplifying hosts. Strong correlations between American Robins and predicted EEEV activity in fall and winter were also consistent with their implication as important hosts (40). Eastern Phoebe and Yellow-rumped Warbler showed weaker but similar patterns, suggesting a potential role in sustaining virus activity during the non-breeding season. By contrast, Wood Thrush showed unexpectedly weak correlations despite its prominence as a host of EEEV in northeastern systems (13, 41), and wading birds such as Green Heron and Black-crowned Night Heron exhibited weak negative correlations, raising the possibility of dilution effects, though their host competence remains unknown. Overall, the alignment between migration phenology, abundances, and correlation peaks show compelling patterns supporting avian movement in shaping EEEV landscape-scale dynamics. These species-specific patterns highlight the promise of integrating community science with traditional surveillance to better disentangle host contributions and strengthen early-warning systems for zoonotic arboviruses. Still, much needs to be done to better integrate bird abundance data into modeling frameworks, as we discuss below.

Caveats, Conclusions and Future Directions

While our models and predictions are robust and the most comprehensive to date, we also recognize some key limitations and caveats. First, historical records of EEEV sentinel chickens indicate whether susceptible chickens were placed in coops in a mosquito program for the sampling week but do not indicate the exact number of birds within each coop. The sentinel chicken surveillance program follows a standardized protocol with ~6 chickens tested weekly per coop. However, this number can vary slightly, which may introduce some uncertainty in our weighting term. The number of chickens tested in each coop is now recorded electronically, and future analyses will benefit from this more precise number.

A second key limitation that precluded a joint analysis of abiotic and biotic factors in the same predictive modeling framework was the nature of the eBird data we used. Briefly, eBird abundance data was available for only a single representative year. This limitation means that we could not account for interannual variation in bird abundance, and while it may be possible to back calculate a proxy for this from raw eBird datasets, it was out of scope of the intended effort here. Despite this challenge, eBird data representing both migratory and resident species of suspected hosts provided insights into key associations between the timing and distribution of avian species abundances and predicted EEEV.

A next step is to expand this framework from retrospective predictions to forecasts of EEEV seroconversion and to incorporate eBird data and near-real-time and forecasted bird migration data from platforms such as BirdCast (42), which could provide additional information toward improving early-warning capacity. More broadly, linking avian movement, abundances, and phenology with virus activity creates a foundation for scalable ecological forecasting of mosquito–bird vector-borne disease systems in Florida and beyond. Emerging data streams from community science and cross-sector monitoring will allow forecasts to capture macroscale drivers of host and vector dynamics, while informing local-scale hazard assessments under accelerating global change. Essential to this effort is continued integration of biological and environmental data with modeling approaches that account for

unmeasured latent variables, thereby improving the ability to predict when and where elevated transmission risk is most likely to occur.

Materials and Methods

Study Area

Florida spans subtropical to tropical climates, characterized by year-round warmth and pronounced wet (May to October) and dry (November to April) seasons. The state has experienced significant environmental change, including rising minimum temperatures, altered precipitation (43), and ongoing land use transformation, primarily from natural to agricultural habitats and urban development (44). Additionally, the Florida peninsula is a critical stopover site for migratory bird species along the Atlantic flyway and also serves as an overwintering ground for numerous North American bird species (45).

Sentinel Chicken Data and Preparation

We utilized EEEV surveillance data from the Florida Department of Health's sentinel chicken monitoring program, which operates through partnerships with local mosquito control programs across the state (25). Following standardized protocols, alphavirus-susceptible chickens are housed at monitoring sites and sampled weekly throughout large portions of the year (25). Blood samples undergo initial screening with hemagglutination inhibition tests, followed by IgM enzyme-linked immunosorbent assays (ELISA) to identify antibody-positive samples. If ELISA results are negative or unequivocal, a Plaque Reduction Neutralization Test (PRNT) is performed to differentiate between EEEV and Highlands J virus.

Following the approach outlined in (10), we created a bioinformatics pipeline to digitize and quality-control nearly two decades of surveillance records from 2001-2019. Paper reports were processed using Amazon Web Services Textract optical character recognition, then cleaned and formatted using R tidyverse functions (46) and OpenRefine software (47). Our initial dataset comprised 116,179 weekly records from 526 sites across 42 counties. Data were then aggregated to monthly intervals and spatially filtered using a

non-convex hull around sites with at least one EEEV detection, excluding sentinel sites in southern Florida where seroconversion was never recorded. Records prior to 2005 were removed due to diagnostic limitations between EEEV and Highlands J virus. For each coop, we calculated the monthly proportion of positive chickens, weighted by the number tested. Similar to (10) and general FDOH guidelines, we assumed six chickens were sampled per coop per week, though a gap is present in historical records reporting the exact number of chickens in individual coops each week, which may introduce some uncertainty in derived weights. The dataset was partitioned into training (2005–2017) and testing (2018–2019) subsets, yielding 84,719 monthly records from 476 sites in 39 counties, providing broad coverage of EEEV transmission across Florida.

Environmental Variables

We compiled climate and landscape covariates for each sentinel chicken site. Daymet daily precipitation and minimum and maximum temperatures from 2000 to 2020 at a 1-km² resolution (48) were downloaded using the 'climateR' R package (49). We then aggregated values to monthly means, before extracting values to site locations using available functions in the 'terra' and 'sf' R packages (50, 51). To identify temporal lags, we used cross-correlation analyses between statewide monthly EEEV seroconversions and each monthly climate variable at a maximum of 12 monthly lags using the 'forecast' R package (52). Significant lags with coefficients exceeding ±0.2 identified 1-, 5-, and 12-month lags for precipitation and 1-, 6-, and 12-month lags for both temperature variables.

Land use/land cover (LULC) data were extracted from 30-m² resolution National Land Cover Database data (i.e., 2001, 2004, 2006, 2008, 2011, 2016, 2019) (53) using the 'fedData' R package (54). Land cover data was reclassified into five categories (i.e., developed, cropland, natural, forest, wetland), and then assigned to EEEV records by nearest year. We then calculated the proportion of each LULC type within 2.5-km² buffers around sentinel sites. For statewide predictions, we generated 1-km² environmental covariate grids by computing land cover proportions within 2.5-km² windows and prepared

climate variables with corresponding temporal lags at the same 1-km² resolution. All environmental covariates were standardized using z-score normalization based on training data statistics (2005-2017), with scaling parameters applied consistently to testing data (2018-2019) to maintain temporal integrity. For statewide spatial predictions, environmental variables at 1-km² resolution were standardized using training data means and standard deviations.

To assess multicollinearity between environmental predictors, we first calculated binomial generalized linear mixed effects models (GLMMs) with a complementary log-log (cloglog) link function with monthly proportions of EEEV seroconversion weighted by sampling effort as the response variable, landscape and lagged climate variables as predictor variables, and nested site-by-country random intercepts. Predictors were parameterized as second-order polynomials to capture non-linear responses. We then calculated variance inflation factors (VIFs) using a threshold >5 to identify candidate predictors sets, resulting in 2 candidate sets for subsequent analyses with differing levels of predictor variance inflation (Table S1.5). GLMMs were run in the 'glmmTMB' R package and VIFs were calculated using the 'performance' R package (55).

Model Selection

Our overall modeling framework used GLMMs (56) that explicitly account for spatiotemporal autocorrelation using Gaussian Markov random fields (GMRFs) (57) approximated using stochastic partial differential equations (SPDE) with Matérn covariance functions and a first order autoregressive (AR1) term (58, 59). For full model specification see Methods S.1. Models were implemented in the 'sdmTMB' R package, which combines utility from TMB and INLA (60).

For computational efficiency, we first performed model selection using spatial binomial GLMMs without a temporal term, using a cloglog link, nested random intercepts for sites within each county, and corresponding environmental predictor variables. Model selection followed a backward stepwise approach, systematically removing predictors and retaining those whose removal most improved Akaike Information Criterion (AIC) scores

(61). To construct spatial meshes required to calculate the SPDE, we followed the approach detailed in (62) using the 'fmesher' R package (63), using a coarse mesh with 250 vertices to balance model complexity with computational efficiency during the model selection process. We then systematically tested mesh resolutions from 250 to 1,000 vertices by adjusting cutoff parameters (i.e., minimum triangle size), max edge (i.e., maximum triangle edge length), and offset (i.e., spatial extensions) values scaled by factors of the spatial extent of the study area. Model convergence was assessed using 'sdmTMB' diagnostic functions, with a 750-vertex mesh (cutoff = 8.5 km), which provided optimal convergence properties without excessive computational costs.

After identifying best performing models from both candidate sets we fitted full spatiotemporal models (Table 1 & Table S1.6), and performance was evaluated using sanity check functions in 'sdmTMB' that examined parameter convergence, extreme eigenvalues, standard errors, and random field variances. Residual diagnostics were assessed by generating 500 simulated datasets using maximum likelihood estimates with multivariate normal sampling to assess model adequacy, residual patterns, and potential violations of distributional assumptions, using the 'DHARMa' R package (64).

To determine whether fixed-effect environmental predictors, spatiotemporal structures, and random intercepts improved model performance, we observed marginal AIC values to compare our full spatiotemporal models against seven alternative models including different combinations of these terms (Table S1.1). We calculated conditional percent deviance explained using log-likelihood ratios as *1 - (model deviance / null deviance)*, where *deviance* = *-2* × *log-likelihood* and the null model was the intercept-only non-spatial model. AIC weights were computed using the 'qpcR' R package (65) to quantify relative model support.

Predictive accuracy was evaluated using root mean square error (RMSE) calculations on out-of-sample predictions (2018-2019) and across the full training period (2005 - 2017) at annual and monthly aggregation levels. Following Thorson & Kristensen (2016) (66), we applied epsilon bias-correction estimators to obtain accurate temporal

predictions of statewide EEEV seroprevalence and associated uncertainty. This approach accounts for bias accumulation when aggregating non-linear model predictions across sites and time periods. We generated 500 simulations using multivariate normal sampling, converted counts to proportions using the binomial size weights, then calculated weighted averages across surveillance sites by month. We then computed temporal indices with multiple confidence intervals (50%, 80%, 95%) to characterize prediction uncertainty over the 15-year time series using the *get_index_sims()* function.

Models predicted monthly EEEV seroprevalence across Florida from 2005-2019 using 100 simulations per prediction to quantify uncertainty. Predictions were converted from cloglog link space to the response scale and truncated at various quantiles (90th, 95th, 98th percentiles) for visualization. Annual summaries were derived by averaging monthly predictions, and overall summaries were aggregated across years to map persistent risk. Epsilon bias correction was applied to temporal indices but omitted from full statewide predictions due to computational intensity.

Avian Host-Pathogen Association Analysis

We tested whether there are strong correlations between relative abundance of 12 common bird species that are suspected hosts of EEEV and model-predicted EEEV seroprevalence. In particular, we selected migratory and resident Florida species with differing expected correlations to virus activity, based on prior *Cs. melanura* blood meal analyses in Florida and Alabama (23). Our migratory species were: American Robin, Eastern Phoebe, Hermit Thrush, Red-eyed Vireo, Wood Thrush, and Yellow-rumped Warbler. Our year-round residents were: Northern Cardinal, Common Yellowthroat, White-eyed Vireo, Green Heron, Pine Warbler, and Black-crowned Night Heron. We obtained weekly species-specific relative abundance estimates from the eBird Status and Trends dataset (67), modeled as the number of individuals a citizen scientist may observe during a 1- km² traveling checklist during the optimal time of day. Raster data are available at 3, 9, and 27 km² resolutions and

represent species-specific relative abundance across a “typical year.” Here, we used median weekly relative abundances at 3-km² resolution for each species.

We processed EEEV model predictions by converting spatiotemporal estimates to monthly rasters at a 1- km² resolution, then computed mean monthly seroprevalence across the full study period (2005-2019) to characterize average seasonal patterns. eBird weekly abundance data were aggregated to monthly scales using a standard week-to-month mapping (52 weeks to 12 months).

All spatial data were reprojected to UTM Zone 17N (EPSG:32617) and resampled to eBird's 3- km² resolution using bilinear interpolation, ensuring spatial alignment for correlation analyses. We then calculated pixel-wise Pearson correlations between estimated bird abundance and EEEV seroprevalence for each month, incorporating temporal lags of 0-3 months for migratory species to capture delayed associations between bird migration and EEEV virus activity, and for resident species, only contemporary correlations (0-month lag) were examined due to year-round local populations.

Acknowledgements

This study was funded by the University of Florida Research Opportunity Seed Fund. We thank Dr. Jonathan Day, Bradley Eastmond, Florida Department of Health, and all mosquito control district personnel for thousands of hours of maintenance, sample collection, and testing of sentinel chicken flocks, as well as providing access to these data. We also thank Dr. Sean C. Anderson for analytical support.

Data availability

All code necessary to conduct these analyses are stored in the following Github repository: https://github.com/Campbell-Lab-FMEL/EEEV_forecasting. Georeferenced sentinel chicken seroconversion data is available upon request through the Florida Department of Health Arbovirus Surveillance program upon agreement from participating Florida mosquito control programs through a memorandum of understanding. The authors

did not receive special privileges in accessing the data that other researchers would not have. Contact information for data requests are available through the FDOH website: https://www.floridahealth.gov/diseases-and-conditions/mosquito-borne-diseases/surveillance.html.

Supplementary Methods

The spatiotemporal model developed in this study is similar to non-spatial GLMMs but includes additional terms to account for spatiotemporal non-independence. Here, we estimated Gaussian Markov random fields (GMRFs) for spatiotemporal variation ε_t using the SPDE method that approximates a Matérn correlation function.

We evaluated the likelihood for each observation $i \in \{1,2,\ldots,N_{observations}\}$ as proportion data following a binomial distribution, where c_i represents the number of EEEV-seropositive birds out of n_i total birds tested, with $y_i = c_i/n_i$ being the observed proportion and using a complementary log-log link function:

$$c_i \sim Binomial(n_i, \mu_i),$$

$$\mu_i = 1 - exp(-exp(p_i)),$$

$$p_i = X_i^T \beta + Z_i^T \alpha + A_i \varepsilon_{t(i)},$$

$$\varepsilon_i = \{$$

$$MVN(0, \Sigma_e) \quad \text{if } t = 1$$

$$MVN(\backslash, \Sigma_e) \quad \text{if } t > 1$$

$$\text{MVN}(\rho\varepsilon_{t-1}, (1 - \rho^2)\Sigma_e) \text{ if } t > 1$$

$$\}$$

where y_i represents the observed proportion of EEEV-seropositive birds at observation i and time t[i]; μ_i represents the true probability of EEEV seropositivity; n_i represents the number of chickens tested (weights = testing × 6 chickens per testing event); X and Z are design matrices for fixed and random effects with coefficient vectors β and α, respectively; ε_t

543 represents the spatiotemporal random field at time t across SPDE vertices; A is the bilinear
544 interpolation matrix such that $A_i\varepsilon_{t(i)}$ interpolates spatiotemporal variation for observation i; Σ_e
545 represents the spatial covariance matrix of the spatiotemporal random field; and ρ
546 represents the first-order autocorrelation over time in the spatiotemporal term ε_t. The
547 complementary log-log link function relates the linear predictor to the probability μ_i, which is
548 appropriate for modeling rare events and asymmetric probability distributions.

549
550 Table S1.1: Description of the model components leading to the development of generalized
551 linear mixed models ("GLMMs") with Gaussian Markov random fields ("GMRF"), including
552 Deviance Explained, marginal AIC and AIC Weight. Models were fitted to seroconversion
553 proportions of Eastern Equine Encephalitis Virus ("EEEV") in Florida at monthly time scales.
554 These models are described in terms of their major components, including intercepts ("~1"),
555 fixed effects, random effects, and Gaussian Markov random fields (GMRF).
556
557

Model	Deviance Explained	AIC	ΔAIC	AIC Weight
EEEV ~ GMRF + fixed effects + random effects	0.377	6164.1	0	1
EEEV ~ GMRF + random effects	0.369	6210.2	46	0
EEEV ~ GMRF + fixed effects	0.34	6518.5	354.4	0
EEEV ~ 1 + GMRF	0.331	6573.7	409.5	0
EEEV ~ fixed effects + random effects	0.299	6920.7	756.6	0
EEEV ~ random effects	0.274	7135.5	971.3	0
EEEV ~ 1	0	9822.8	3658.6	0

558
559 Table S1.2: Summary of model predictive accuracy when predicting out-of-sample Eastern
560 Equine Encephalitis seroprevalence proportions. Predictive accuracy is measured using root
561 mean square error (RMSE) and reported at the monthly scale with yearly totals, as well as
562 the overall study period (Total) for out-of-sample data for the years 2018 and 2019.
563

Year	Month	RMSE - Candidate Set 1	RMSE - Candidate Set 2

2018	1	0.00255	0.00255
	2	0.00002	0.00002
	3	0.00305	0.00305
	4	0.00700	0.00700
	5	0.01416	0.01415
	6	0.01172	0.01172
	7	0.01425	0.01425
	8	0.00726	0.00726
	9	0.00318	0.00318
	10	0.00004	0.00003
	11	0.00555	0.00556
	12	0.00255	0.00255
2019	1	0.00360	0.00360
	2	0.00191	0.00191
	3	0.00153	0.00153
	4	0.00360	0.00360
	5	0.01028	0.01028
	6	0.01535	0.01535
	7	0.01349	0.01349
	8	0.00425	0.00425

	9	0.00565	0.00565
	10	0.00270	0.00270
	11	0.00191	0.00191
	12	0.00360	0.00360
Yearly Totals			
2018		0.00767	0.00767
2019		0.00724	0.00724
Total		0.00746	0.00746

Table S1.3: Parameter estimates of predictors from a spatiotemporal model fitted against monthly aggregated EEEV seroprevalence proportions in Florida's sentinel chicken surveillance system. Predictors with Confidence Intervals (95% CI; Lower, Upper) that do not overlap with zero are highlighted in **bold** and are considered important predictors of EEEV seroprevalence.

Variable	Type	Estimate	Std. Error	95% CI (Lower)	95% CI (Upper)
Intercept	*NA*	-10.5257	0.3436	-11.199	-9.8523
Precipitation Lag 1 Month	*Linear*	10.831	19.0896	-26.5838	48.2459
Precipitation Lag 1 Month	*Quadratic*	-3.2013	15.3243	-33.2363	26.8338
Precipitation Lag 5 Months	*Linear*	-18.8877	20.4343	-58.9382	21.1628

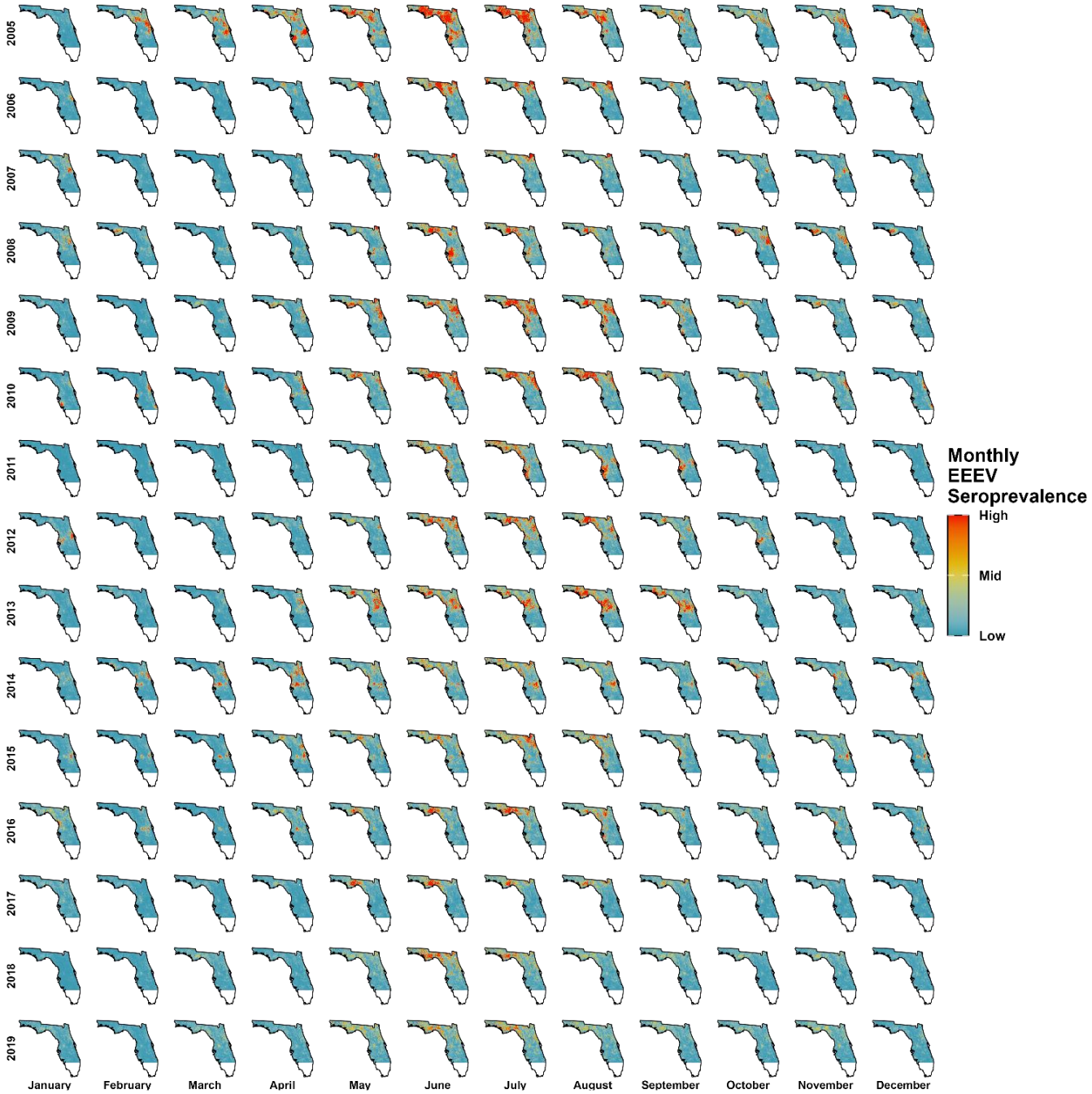
Precipitation Lag 5 Months	*Quadratic*	-24.9001	19.0994	-62.3343	12.5341
Precipitation Lag 12 Months	*Linear*	36.3576	18.3542	**0.384**	**72.3312**
Precipitation Lag 12 Months	*Quadratic*	-30.315	14.1791	**-58.1056**	**-2.5244**
Tmax Lag 6 Months	*Linear*	-37.9701	56.2275	-148.174	72.2337
Tmax Lag 6 Months	*Quadratic*	38.4038	21.9422	-4.6022	81.4097
Tmax Lag 12 Months	*Linear*	-9.8866	54.723	-117.142	97.3686
Tmax Lag 12 Months	*Quadratic*	11.7183	26.0875	-39.4122	62.8489
Tmin Lag 1 Month	*Linear*	76.811	40.4647	-2.4983	156.1204
Tmin Lag 1 Month	*Quadratic*	-67.3957	28.0519	**-122.376**	**-12.4151**
Forest Cover	*Linear*	142.3936	45.2068	**53.7899**	**230.9973**
Forest Cover	*Quadratic*	-87.0073	30.0387	**-145.882**	**-28.1325**
Wetland Cover	*Linear*	128.5891	32.5892	**64.7155**	**192.4627**
Wetland Cover	*Quadratic*	-44.5205	30.6417	-104.577	15.5361

Table S1.4: Parameter estimates of predictors from a spatiotemporal model fitted against monthly aggregated EEEV seroprevalence proportions in Florida's sentinel chicken surveillance system. The variables included in this model were those that were removed based on initial VIF screening and fitted as an alternative candidate set. Predictors with Confidence Intervals (95% CI; Lower, Upper) that do not overlap with zero are highlighted in **bold** and are considered important predictors of EEEV seroprevalence.

Variable	Type	Estimate	Std. Error	95% CI (Lower)	95% CI (Upper)
Intercept	*NA*	-10.5292	0.346	-11.2075	-9.851
Tmin (Current)	*Linear*	136.385	66.3949	6.2534	266.5165
Tmin (Current)	*Quadratic*	22.1909	34.736	-45.8903	90.2721
Tmin Lag 6 Months	*Linear*	-59.6911	57.672	-172.726	53.3439
Tmin Lag 6 Months	*Quadratic*	22.9416	20.5329	-17.3022	63.1853
Tmin Lag 12 Months	*Linear*	-99.4721	70.6447	-237.933	38.9891
Tmin Lag 12 Months	*Quadratic*	61.264	34.09	-5.5512	128.0792
Tmax Lag 1 Month	*Linear*	46.5541	51.4576	-54.301	147.4091
Tmax Lag 1 Month	*Quadratic*	-96.7682	30.8724	**-157.277**	**-36.2595**
Forest Cover	*Linear*	143.511	45.6413	**54.0557**	**232.9663**

Forest Cover	*Quadratic*	-86.9773	30.1115	**-145.995**	**-27.9599**
Wetland Cover	*Linear*	129.2191	32.5968	**65.3305**	**193.1077**
Wetland Cover	*Quadratic*	-44.8197	30.6592	-104.911	15.2712

Figure S1.1: Statewide predictions of the spatiotemporal dynamics of EEEV seroprevalence from the Florida sentinel chicken system during the years of 2001 – 2019. Predictions were obtained from a spatiotemporal model fitted to data aggregated to monthly proportions of EEEV seroprevalence in a sentinel chicken system. All predictions are provided with a quantile truncation for ease of visualization (98th percentile).


2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
January
February
March
April
May
June
July
August
September
October
November
December
Monthly
EEEV
Seroprevalence
High
Mid
Low

590

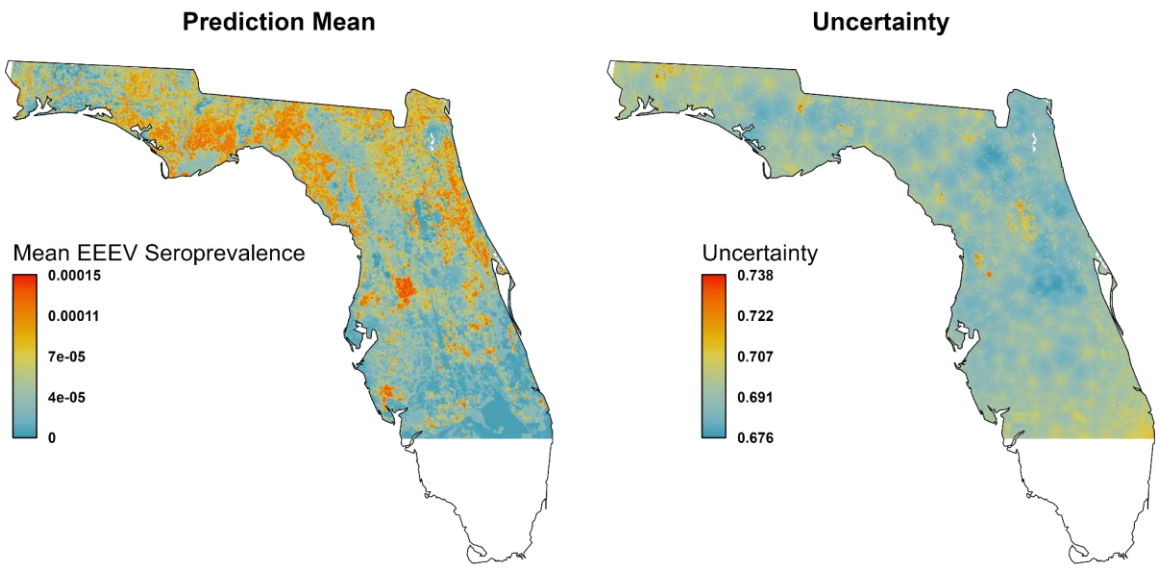


Figure S1.2: Summary of statewide prediction means and associated uncertainty of EEEV seroprevalence during the years of 2005 - 2019 in the Florida Sentinel Chicken Surveillance system. Predictions were obtained from a spatiotemporal model fitted against monthly aggregated seroprevalence proportions. Predictions are mapped using 98th percentile quantile truncation for ease of visualization.

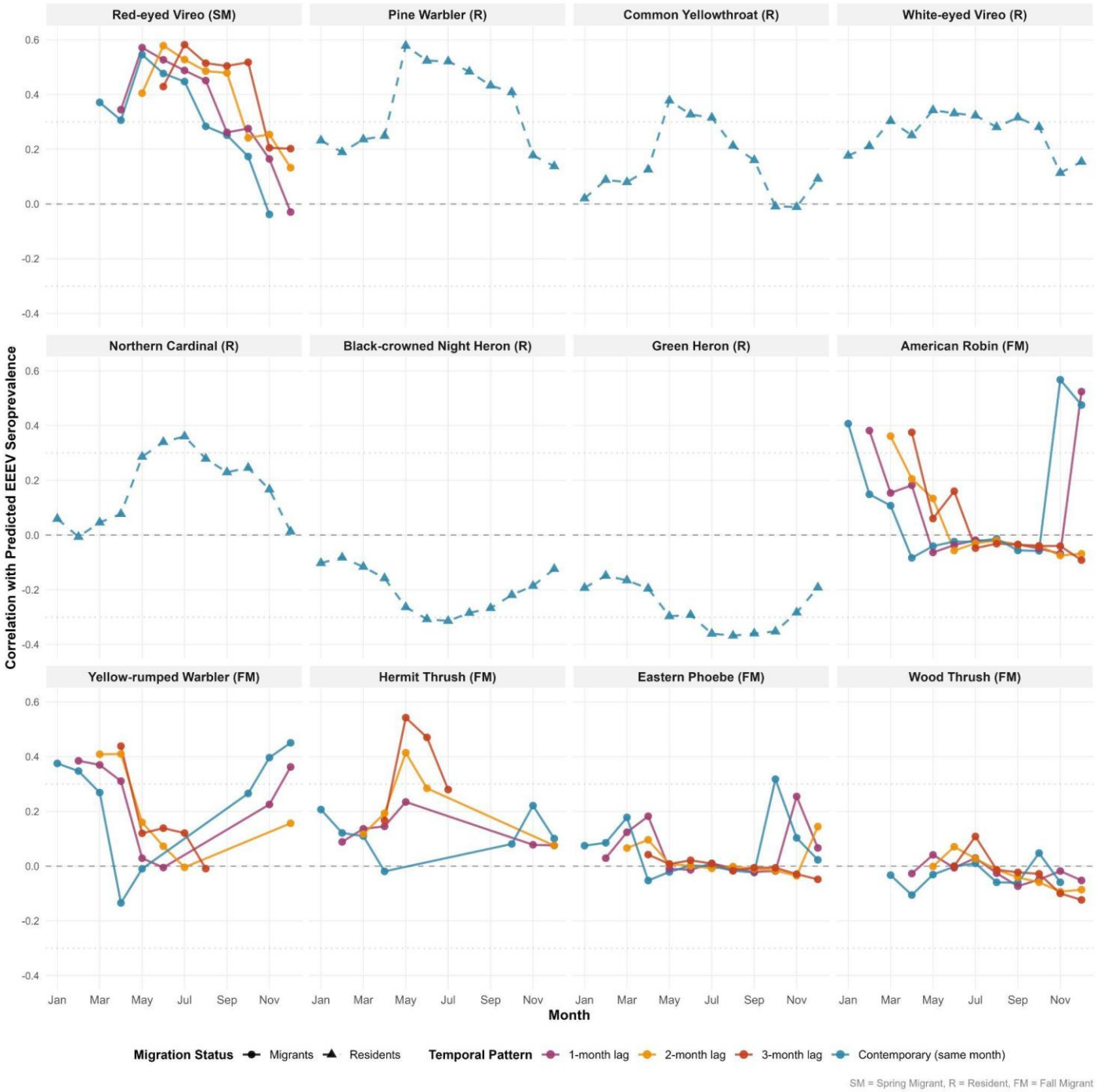


Figure S1.3: Pixelwise Pearson correlations between predicted monthly EEEV seroprevalence and monthly bird abundance across Florida. EEEV predictions were temporally averaged across the study period, while weekly eBird abundance estimates were aggregated to monthly resolution. For resident species, we assessed contemporaneous correlations only. For migratory species, we evaluated both contemporaneous correlations and temporal lags of 1-3 months to account for delayed associations following migration events.

Table S1.5: Results from backward stepwise variable selection to identify the best predictors of monthly Eastern Equine Encephalitis Virus (EEEV) seroprevalence. We first removed highly collinear variables to create two candidate sets, then systematically dropped the least

important variables based on AIC improvement. All models included nested random effects (sites within counties) and were weighted by monthly testing frequency. All predictors were modeled as quadratic terms.

Candidate Set	Step	Variable Removed	AIC	ΔAIC	Remaining Predictors
1	0	(starting model)	6930.3	-	precipitation (current + 1,5,12 month lags), temperature max (current + 6,12 month lags), temperature min (1 month lag), developed land, forest, wetlands
	1	developed cover	6927.8	2.5	precipitation (current + lags), temperature max (6,12 month lags), temperature min (1 month lag), forest, wetlands
	2	current temperature max	6925.6	2.2	precipitation (current + lags), temperature max (6,12 month lags), temperature min (1 month lag), forest, wetlands
	3	current precipitation	6924.7	0.9	precipitation (1,5,12 month lags), temperature max (6,12 month lags), temperature min (1 month lag), forest, wetlands
2	0	(starting model)	6879.4	-	temperature min (current + 6,12 month lags), temperature max (1 month lag), developed land, forest, wetlands
	1	developed cover	6877.9	1.5	temperature min (current + 6,12 month lags), temperature max (1 month lag), forest, wetlands

Table S1.6: Description of the predictor variables used to model spatiotemporal EEEV seroprevalence in the state of Florida (USA). Terms included in this table were those that

were removed during initial VIF screening of variables and fit as an alternative candidate set. Descriptions and parameterization details are provided for each variable.

Effect	Term	Description	Variable type	Parameterization
Fixed	*tmin*	Min temperature	Continuous	2nd order orthogonal polynomial
	tmin_lag6	Min temperature (6 mo. lag)	Continuous	2nd order orthogonal polynomial
	tmin_lag12	Min temperature (12 mo. lag)	Continuous	2nd order orthogonal polynomial
	tmax_lag1	Max temperature (1 mo. lag)	Continuous	2nd order orthogonal polynomial
	forest	Proportion forest cover	Continuous	2nd order orthogonal polynomial
	wetlands	Proportion wetlands	Continuous	2nd order orthogonal polynomial
Random	*county*	County name	39-level factor	Intercept nested above site_id
	site_id	Monitoring site name	476-level factor	Intercept nested below county

References

1. E. Münger, *et al.*, One Health approach uncovers emergence and dynamics of Usutu and West Nile viruses in the Netherlands. *Nat. Commun.* **16**, 7883 (2025).

2. E. N. Pavlovsky, The natural nidus of a disease as a pathobiocenose. *The Natural Nidality of Transmissible Disease. Univ*. (1966).

3. E. A. Mordecai, *et al.*, Thermal biology of mosquito-borne disease. *Ecol. Lett.* **22**, 1690–1708 (2019).

4. W. K. Reisen, Landscape epidemiology of vector-borne diseases. *Annu. Rev. Entomol.* **55**, 461–483 (2010).

5. H. D. Kiryluk, C. B. Beard, K. M. Holcomb, The use of environmental data in descriptive and predictive models of vector-borne disease in North America. *J. Med. Entomol.* **61**, 595–602 (2024).

6. L. P. Molina-Guzmán, L. A. Gutiérrez-Builes, L. A. Ríos-Osorio, Models of spatial analysis for vector-borne diseases studies: A systematic review. *Vet. World* **15**, 1975–1989 (2022).

7. K. M. Holcomb, *et al.*, Evaluation of an open forecasting challenge to assess skill of West Nile virus neuroinvasive disease prediction. *Parasit. Vectors* **16**, 11 (2023).

8. W. M. Chung, C. M. Buseman, S. N. Joyner, S. M. Hughes, The 2012 West Nile encephalitis epidemic in Dallas, Texas. *JAMA* (2013).

9. K. C. Dye-Braumuller, *et al.*, Riding the wave: Reactive vector-borne disease policy renders the United States vulnerable to outbreaks and insecticide resistance. *J. Med. Entomol.* **59**, 401–411 (2022).

10. J. A. Baecher, *et al.*, Toward ecological forecasting of West Nile virus in Florida: Insights from two decades of sentinel chicken surveillance. *Sci. Total Environ.* **1000**, 180308 (2025).

11. L. T. Webster, F. H. Wright, Recovery of eastern equine encephalomyelitis virus from brain tissue of human cases of encephalitis in Massachusetts. *Science* **88**, 305–306 (1938).

12. C. D. Morris, *Eastern equine encephalomyelitis. Monath TP, ed. The Arboviruses: Epidemiology and Ecology* (CRC Press, 1988).

13. P. M. Armstrong, T. G. Andreadis, Ecology and epidemiology of eastern equine encephalitis virus in the northeastern United States: An historical perspective. *J. Med. Entomol.* **59**, 1–13 (2022).

14. D. M. Morens, G. K. Folkers, A. S. Fauci, Eastern equine encephalitis virus - another emergent arbovirus in the United States. *N. Engl. J. Med.* **381**, 1989–1992 (2019).

15. J. E. Staples, C. V. Gould, Eastern equine encephalitis in the US. *JAMA* **334**, 530–531 (2025).

16. CDC, U.S. Centers for Disease Control and Prevention. Eastern Equine Encephalitis. (2022). Available at: https://www.cdc.gov/easternequineencephalitis/index.htm [Accessed 5 October 2025].

17. N. P. Lindsey, S. W. Martin, J. E. Staples, M. Fischer, Notes from the field: Multistate outbreak of eastern equine encephalitis virus - United States, 2019. *MMWR Morb. Mortal. Wkly. Rep.* **69**, 50–51 (2020).

18. T. Corrin, R. Ackford, M. Mascarenhas, J. Greig, L. A. Waddell, Eastern equine encephalitis virus: A scoping review of the global evidence. *Vector Borne Zoonotic Dis.* **21**, 305–320 (2021).

19. A. C. Brault, *et al.*, Genetic and antigenic diversity among eastern equine encephalitis viruses from North, Central, and South America. *Am. J. Trop. Med. Hyg.* **61**, 579–586 (1999).

20. A. Burgueño, *et al.*, Genomic characterization and seroprevalence studies on alphaviruses in Uruguay. *Am. J. Trop. Med. Hyg.* **98**, 1811–1818 (2018).

21. N. C. Arrigo, A. P. Adams, S. C. Weaver, Evolutionary patterns of eastern equine encephalitis virus in North versus South America suggest ecological differences and taxonomic revision. *J. Virol.* **84**, 1014–1025 (2010).

22. U.S. Department of Health and Human Sercives & U.S. Department of Agriculture, HHS and USDA Select Agents and Toxins List (CS-357006-A). *https://www.selectagents.gov/sat/docs/HHS-USDA-Select-Agents-and-Toxins-List_508.pdf*.

23. N. D. Burkett-Cadena, J. F. Day, T. R. Unnasch, Ecology of eastern equine encephalitis virus in the southeastern United States: Incriminating vector and host species responsible for virus amplification, persistence, and dispersal. *J. Med. Entomol.* **59**, 41–48 (2022).

24. V. Hill, *et al.*, Dynamics of eastern equine encephalitis virus during the 2019 outbreak in the Northeast United States. *Curr. Biol.* **33**, 2515-2527.e6 (2023).

25. FDOH, Florida Department of Health. Non-Human Mosquito-Borne Disease Monitoring Activities. (2021). Available at: http://www.floridahealth.gov/diseases-and-conditions/mosquito-borne-diseases/_documents/guidebook-chapter-ten.pdf [Accessed 5 October 2025].

26. P. C. Kyriakidis, A. G. Journel, Geostatistical space–time models: A review. *Math. Geol.* **31**, 651–684 (1999).

27. A. E. Gelfand, S. Banerjee, Bayesian modeling and analysis of geostatistical data. *Annu. Rev. Stat. Appl.* **4**, 245–266 (2017).

28. N. K. Skaff, P. M. Armstrong, T. G. Andreadis, K. S. Cheruvelil, Wetland characteristics linked to broad-scale patterns in Culiseta melanura abundance and eastern equine encephalitis virus infection. *Parasit. Vectors* **10**, 501 (2017).

29. P. T. V. Kelen, *et al.*, Habitat associations of eastern equine encephalitis transmission in Walton County Florida. *J. Med. Entomol.* **49**, 746–756 (2012).

30. P. T. Vander Kelen, *et al.*, Spatial epidemiology of eastern equine encephalitis in Florida. *Int. J. Health Geogr.* **11**, 47 (2012).

31. L. P. Campbell, *et al.*, Spatiotemporal modeling of zoonotic arbovirus transmission in northeastern Florida using sentinel chicken surveillance and Earth observation data. *Remote Sensing* **14**, 3388 (2022).

32. S. J. Mundis, S. Harrison, D. Pelley, S. Durand, S. J. Ryan, Spatiotemporal environmental drivers of eastern equine encephalitis virus in central Florida: Towards a predictive model for a lethal disease. *J. Med. Entomol.* **59**, 1805–1816 (2022).

33. P. M. Armstrong, T. G. Andreadis, Eastern equine encephalitis virus--old enemy, new threat. *N. Engl. J. Med.* **368**, 1670–1673 (2013).

34. Y. Tavares, *et al.*, Regional variation in the landscape ecology of West Nile virus sentinel chicken seroconversion in Florida. *PLoS One* **19**, e0305510 (2024).

35. J. P. DeGroote, R. Sugumaran, National and regional associations between human West Nile virus incidence and demographic, landscape, and land use conditions in the coterminous United States. *Vector Borne Zoonotic Dis.* **12**, 657–665 (2012).

36. G. R. Mullen, L. A. Durden, *Medical and Veterinary Entomology* (Academic Press, 2009).

37. C. J. M. Koenraadt, L. C. Harrington, Flushing effect of rain on container-inhabiting mosquitoes *Aedes aegypti* and *Culex pipiens* (Diptera: Culicidae). *J. Med. Entomol.* **45**, 28–35 (2008).

38. J. J. Brown, M. Pascual, M. C. Wimberly, L. R. Johnson, C. C. Murdock, Humidity - The overlooked variable in the thermal biology of mosquito-borne disease. *Ecol. Lett.* **26**, 1029–1049 (2023).

39. Shultz R, Collins C, Roberts D, DeSouza C, Blackmore C, Florida Department of Health., Florida 2005 Arbovirus Activity by County: 2005 Annual Report. Tallahassee, FL: Florida Department of Health; 2005. *Available from: https://www.floridahealth.gov/diseases-and-conditions/mosquito-borne-diseases/_documents/2005annual-report.pdf*.

40. G. Molaei, P. M. Armstrong, A. C. Graham, L. D. Kramer, T. G. Andreadis, Insights into the recent emergence and expansion of eastern equine encephalitis virus in a new focus in the Northern New England USA. *Parasit. Vectors* **8**, 516 (2015).

41. G. Molaei, *et al.*, Dynamics of vector-host interactions in avian communities in four eastern equine encephalitis virus foci in the northeastern U.s. *PLoS Negl. Trop. Dis.* **10**, e0004347 (2016).

42. B. M. Van Doren, K. G. Horton, A continental system for forecasting bird migration. *Science* **361**, 1115–1118 (2018).

43. USGCRP, “Fifth national climate assessment,” A. R. Crimmins, *et al.*, Eds. (U.S. Global Change Research Program, 2023).

44. M. Volk, T. Hoctor, B. Nettles, R. Hilsenbeck, F. Putz, “Florida land use and land cover change in the past 100 years” in *Florida’s Climate: Changes, Variations, & Impacts*, (Florida Climate Institute, 2017).

45. S. Feng, Q. Yang, H. Qiao, L. E. Escobar, X. Yan, Stopover hotspots for migratory birds in North and Central America. *Biodivers. Inf.* **19** (2025).

46. H. Wickham, *et al.*, Welcome to the tidyverse. *J. Open Source Softw.* **4**, 1686 (2019).

47. K. Ham, OpenRefine (version 2.5). http://openrefine.org. Free, open-source tool for cleaning and transforming data. *J. Med. Libr. Assoc.* **101**, 233–234 (2013).

48. Thornton, M.M., R. Shrestha, Y. Wei, P.E. Thornton, S-C. Kao, and B.E. Wilson, Daymet: Daily Surface Weather Data on a 1-km Grid for North America, Version 4 R1. ORNL DAAC, Oak Ridge, Tennessee, USA. (2022). Available at: https://doi.org/10.3334/ORNLDAAC/2129.

49. M. Johnson, *climateR: climateR*. R package version 0.3.7, https://github.com/mikejohnson51/climater.

50. R. J. Hijmans, terra: Spatial Data Analysis. R package version 1.7-39. *The R Foundation for Statistical Computing* (2023).

51. E. J. Pebesma, R. Bivand, *Spatial Data Science: With Applications in R* (CRC Press, Taylor & Francis Group, 2023).

52. O. Rodriguez, D. Jiménez, ForecasteR: Time series forecast system. The R Foundation. https://doi.org/10.32614/cran.package.forecaster. Deposited 2 March 2022.

53. Dewitz, J., and U.S. Geological Survey, National Land Cover Database (NLCD) 2019 Products (ver. 2. 0, June 2021): U.S. Geological Survey data release (ver. 2. 0, June 2021): U.S. Geological Survey data release.

54. R. K. Bocinsky, D. Beaudette, S. Chamberlain, J. Hollister, J. Gustavsen, *FedData: Download Geospatial Data Federated Data Sources* (2025).

55. D. Lüdecke, M. Ben-Shachar, I. Patil, P. Waggoner, D. Makowski, Performance: An R package for assessment, comparison and testing of statistical models. *J. Open Source Softw.* **6**, 3139 (2021).

56. P. Diggle, E. Gabriel, “Spatio-Temporal Point Processes” in *Chapman & Hall/CRC Handbooks of Modern Statistical Methods*, (CRC Press, 2010), pp. 449–461.

57. H. Rue, L. Held, *Gaussian Markov random fields: Theory and applications* (Chapman and Hall, 2004).

58. F. Lindgren, H. Rue, J. Lindström, An explicit link between Gaussian fields and Gaussian Markov random fields: The stochastic partial differential equation approach. *J. R. Stat. Soc. Series B Stat. Methodol.* **73**, 423–498 (2011).

59. E. Krainski, *et al.*, *Advanced spatial modeling with stochastic partial differential equations using R and INLA* (CRC Press, 2020).

60. S. C. Anderson, E. J. Ward, P. A. English, L. A. K. Barnett, J. T. Thorson, sdmTMB: An R package for fast, flexible, and user-friendly generalized linear mixed effects models with spatial and spatiotemporal random fields. *bioRxiv* (2022).

61. H. Akaike, Likelihood of a model and information criteria. *J. Econom.* **16**, 3–14 (1981).

62. H. Bakka, *et al.*, Spatial modeling with R-INLA: A review. *Wiley Interdiscip. Rev. Comput. Stat.* **10**, e1443 (2018).

63. F. Lindgren, fmesher: Triangle Meshes and Related Geometry Tools (p. 0.3.0). https://cran.r-project.org/package=fmesher.

64. F. Hartig, DHARMa: Residual Diagnostics for Hierarchical (Multi-Level/Mixed)

Regression Models, R Package Version 0.4. 5. (2022).

65. C. Ritz, A.-N. Spiess, qpcR: an R package for sigmoidal model selection in quantitative real-time polymerase chain reaction analysis. *Bioinformatics* **24**, 1549–1551 (2008).

66. J. T. Thorson, K. Kristensen, Implementing a generic method for bias correction in statistical models using random effects, with spatial and population dynamics examples. *Fish. Res.* **175**, 66–74 (2016).

67. D. Fink, *et al.*, Modeling avian full annual cycle distribution and population trends with citizen science data. *Ecol. Appl.* **30**, e02056 (2020).