

Assessing the Impact of Streamflow Antecedents on West Nile Virus Outbreaks in California (2006–2023)

Abstract

West Nile virus (WNV) has been the leading cause of domestically acquired arboviral disease in California since its arrival in 2003, and the mosquito vectors that sustain its transmission depend on aquatic larval habitat whose availability is governed by hydrology. Because streamflow integrates upstream precipitation, snowmelt, and water management into a measurable signal of landscape wetness, we tested whether antecedent streamflow conditions improve prediction of county-level summer–autumn WNV outbreaks beyond a parsimonious climate baseline. We assembled human WNV case reports for California’s 58 counties over 18 transmission seasons (2006–2023) from the California Department of Public Health, defined a county-season outbreak as at least five reported human cases between July and December, and linked each county to United States Geological Survey daily mean discharge records and Daymet gridded meteorology. After a rigorous gauge audit that retained 703 of 2,397 candidate gauges meeting an 80% period-of-record coverage threshold, we standardized discharge into per-gauge, day-of-year anomalies and engineered two families of hydrological predictors: a spring low-flow proportion capturing drought-like conditions and a dry-to-wet pulse pair capturing rapid rewetting. We compared a baseline mixed-effects logistic regression using only growing degree days and spring precipitation against an augmented model that added the three streamflow features, evaluating both under leave-one-year-out cross-validation. The baseline model was already strongly discriminating (area under the precision–recall curve, AUPRC, = 0.813; ROC-AUC = 0.940), and adding hydrology changed out-of-sample AUPRC by only +0.0007, failing a pre-registered improvement trigger of +0.02. Spring growing degree days dominated the signal (odds ratio $\approx$ 8.4 per standard deviation), whereas streamflow terms were individually small and inconsistent in direction. We conclude that, at the county-season scale and given a temperature-aware baseline, antecedent streamflow adds negligible predictive value for California WNV outbreaks. We nonetheless provide a validated 10-county prioritization for surveillance and discuss the ecological and methodological reasons—including anthropogenic water management that decouples vector habitat from natural streamflow—why a hydrological signal may be masked at this resolution.

Assessing the Impact of Streamflow Antecedents on West Nile Virus Outbreaks in California (2006-2023)

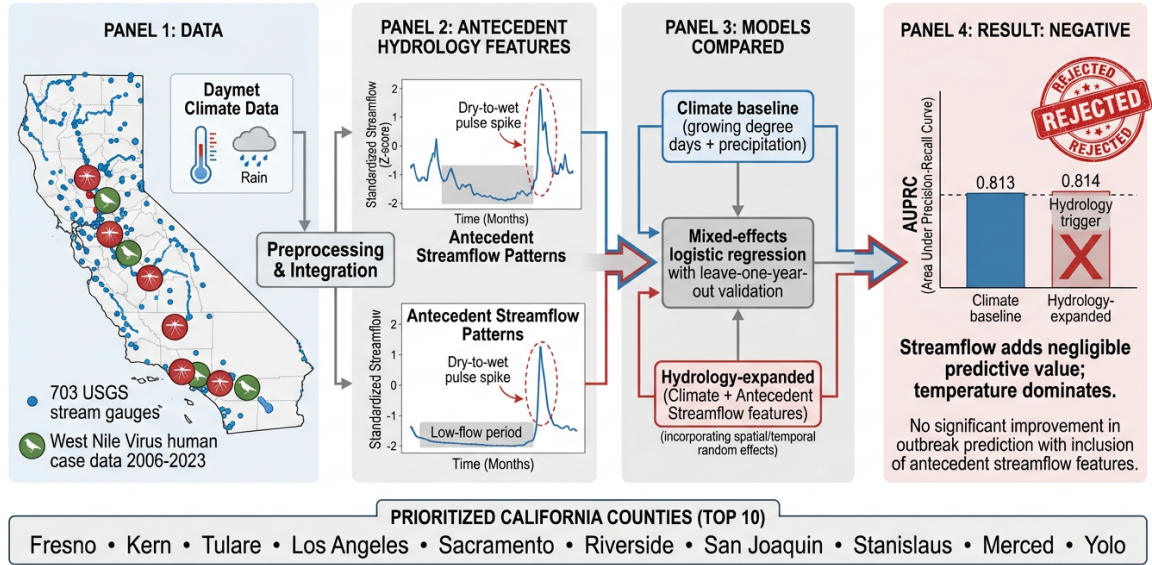


Figure 1: Graphical abstract. Antecedent streamflow anomalies derived from an audited network of USGS gauges were combined with growing-degree-day and precipitation climate covariates to predict county-level West Nile virus outbreaks across California (2006–2023). Under leave-one-year-out validation, a temperature-driven baseline model was already highly discriminating, and the addition of hydrological predictors did not meaningfully improve out-of-sample performance.

1 Introduction

West Nile virus (WNV) is a mosquito-borne flavivirus maintained in an enzootic cycle between ornithophilic *Culex* mosquitoes and passerine birds, with humans and horses serving as incidental, dead-end hosts (Reisen, 2013; Kilpatrick, 2011). Since its introduction into North America in 1999 and its rapid transcontinental spread, WNV has become the most common cause of domestically acquired arboviral disease in the United States, producing recurrent summer epidemics of febrile illness and, in a small but consequential fraction of infections, severe neuroinvasive disease (Petersen et al., 2013). California, which reported its first human cases in 2003, has since experienced sustained transmission across its agricultural valleys, urban basins, and periurban interfaces, and now regularly ranks among the states with the highest annual case burdens (Snyder et al., 2020; Danforth et al., 2022). Understanding the environmental drivers that convert enzootic circulation into human outbreaks is therefore central to targeting the limited resources of vector-control and public-health agencies.

The transmission of WNV is exquisitely sensitive to the environmental conditions that govern both mosquito population dynamics and viral replication within the vector. In California, two *Culex* species dominate the epidemiological landscape: *Culex tarsalis*, the principal rural and agricultural bridge vector associated with irrigated pasture, ricelands, and floodwater habitats of the Central Valley, and members of the *Culex pipiens* complex, which exploit nutrient-rich urban and periurban containers, storm-drain systems, and underground catch basins (Goddard et al., 2002; Reisen, 2013). The precise mix of vector species, including cryptic members such as *Culex torrentium* that can differ in vector competence, further shapes local transmission potential (Jansen et al., 2019). Both

species require standing or slow-moving water for larval and pupal development, so the spatial and temporal availability of aquatic habitat is a first-order constraint on the abundance of adult mosquitoes that can acquire and transmit the virus (Campbell et al., 2013). Avian host community composition further modulates amplification, with competent passerine reservoirs concentrating transmission where mosquito and bird populations overlap (LaDeau et al., 2007; Kilpatrick, 2011).

Layered on top of habitat availability is a strong and well-characterized dependence on temperature. Warmer conditions accelerate mosquito larval development, shorten the gonotrophic cycle, increase biting rates, and—critically—compress the extrinsic incubation period during which ingested virus disseminates to the salivary glands and renders the mosquito infectious (Reisen et al., 2006; Kilpatrick et al., 2008). Laboratory and field studies have consistently identified a thermal optimum for WNV transmission in the high twenties Celsius, with mechanistic trait-based models predicting peak transmission near 24 °C–25 °C (Shocket et al., 2020; Hartley et al., 2012). These thermal dependencies also underpin projections that ongoing climate warming will reshape the global distribution and seasonality of arboviral transmission (Ryan et al., 2019). Because these thermal effects accumulate over the transmission season, degree-day formulations that integrate daily temperatures above a developmental threshold provide a biologically grounded summary of the warmth available to drive amplification, and they have repeatedly emerged as among the strongest predictors of interannual and spatial variation in WNV activity (Paz, 2015; Hartley et al., 2012).

The role of water is more ambiguous. Naively, one might expect greater precipitation and wetter landscapes to increase mosquito habitat and hence outbreak risk. Yet the empirical literature paints a more nuanced picture in which the relationship between moisture and WNV depends on region, vector species, habitat type, and the temporal sequencing of wet and dry episodes (Shaman et al., 2010; Paull et al., 2017). Early national analyses established that weather anomalies measurably influenced WNV incidence across the United States (Soverow et al., 2009). In a landmark synthesis, Paull et al. (2017) found that *drought*, rather than precipitation, was the dominant climatic driver of WNV epidemic intensity across the United States, plausibly because drought concentrates birds and mosquitoes around shrinking water sources, intensifies avian–mosquito contact, and elevates infection prevalence in the vector. Hydrological modeling in Colorado likewise showed that wetter spring conditions followed by drier summers best predicted human cases, implicating the seasonal transition between wet and dry states rather than any single moisture extreme (Shaman et al., 2010, 2011). In arid California specifically, irrigation has been shown to increase and stabilize *Culex* abundance by orders of magnitude and to explain a substantial share of county-level variation in human incidence, demonstrating that anthropogenic water delivery can sustain vector habitat independently of natural rainfall (Kovach and Kilpatrick, 2024). Conversely, mandatory outdoor water-use restrictions during California’s 2012–2016 drought were associated with reduced *Culex quinquefasciatus* abundance, underscoring that human water management, not just hydrometeorology, shapes larval habitat (Bhattachan et al., 2021).

Streamflow—the volumetric discharge of rivers and streams measured continuously at gauging stations—offers an attractive, spatially explicit, and physically integrative proxy for landscape wetness. Unlike point precipitation, streamflow aggregates upstream rainfall, snowmelt, soil-moisture storage, and reservoir operations into a single measured signal, and it responds to both drought (through sustained low flows) and rapid rewetting (through discharge pulses that can flood riparian and floodplain habitats favorable to *Culex tarsalis*) (Tall and Gattton, 2019; Walsh, 2012). Antecedent streamflow conditions in late winter and spring might therefore encode information about the aquatic habitat available to mosquitoes as the transmission season begins. The United States Geological Survey (USGS) operates a dense national network of stream gauges with long, quality-controlled records, making streamflow an operationally appealing candidate predictor for outbreak-forecasting

systems, several of which have begun to incorporate hydrometeorological covariates (Holcomb et al., 2023; Tonks et al., 2024; Beeman et al., 2021). Such efforts sit within a broader movement to deliver actionable climate services for infectious-disease early warning (Thomson and Stanberry, 2022), and they build on spatiotemporal statistical frameworks developed specifically for WNV surveillance (Holcomb et al., 2021).

Despite this appeal, whether antecedent streamflow adds predictive value *beyond* an already temperature-aware climate baseline has not been rigorously established for California at the county-season scale relevant to public-health decision-making. The dominance of temperature in WNV transmission raises the possibility that any hydrological signal is either redundant with, or subordinate to, the thermal forcing captured by degree-day and precipitation covariates. Moreover, the pervasive influence of irrigation, reservoir regulation, and urban water infrastructure in California may decouple the mosquito habitat that actually drives transmission from the natural streamflow recorded at gauges. Resolving this question matters practically: incorporating streamflow into forecasting pipelines imposes real costs in data acquisition, gauge auditing, and model complexity, and those costs are justified only if the added signal materially improves prediction.

Here we conduct a pre-specified, decision-oriented evaluation of antecedent streamflow as a predictor of county-level WNV outbreaks in California over the 18 transmission seasons from 2006 through 2023. We define an outbreak as a county-season with at least five reported human cases between July and December, link each county to an audited network of USGS stream gauges and to Daymet gridded meteorology, and engineer hydrological predictors that capture both spring low-flow (drought-like) conditions and dry-to-wet discharge pulses. We then compare a parsimonious climate baseline against a hydrology-augmented model under leave-one-year-out cross-validation, using a pre-registered improvement threshold to decide whether streamflow earns inclusion in an operational forecast. We further translate the validated model into a prioritized list of counties to guide surveillance. Our central finding is a carefully documented *negative* result: given a growing-degree-day-dominated baseline, antecedent streamflow adds negligible predictive value. We regard this as a scientifically useful outcome that clarifies where forecasting effort is best invested and delineates the conditions under which hydrological signals might yet prove informative.

2 Methods

2.1 Study design and outbreak definition

We conducted a retrospective, observational analysis at the resolution of county-transmission-season, treating each of California’s 58 counties in each of the 18 years from 2006 through 2023 as a candidate observation (a maximum of 1,044 county-seasons). The outcome was a binary indicator of a human WNV outbreak, defined *a priori* as at least five reported human WNV cases (of any clinical classification) occurring within a county between 1 July and 31 December of a given year. This July–December window captures the overwhelming majority of California’s clinically apparent transmission, which is concentrated in late summer and early autumn, while the five-case threshold distinguishes county-seasons with sustained local transmission from sporadic single-case detections that are more vulnerable to reporting noise. Across the study period, 240 of 1,044 county-seasons (23.0%) met the outbreak definition.

2.2 West Nile virus case data

Human WNV case data were obtained from the California Department of Public Health (CDPH), which compiles statewide surveillance through the California Arbovirus Surveillance program and publishes case counts by county and year (California Department of Public Health, 2024; Danforth et al., 2022). Records were aggregated to county-season case totals and screened for internal consistency against published annual state summaries. Over 2006–2023 the state recorded 6,360 human cases, with pronounced interannual variability—from fewer than 300 cases in low-transmission years to a peak exceeding 800 cases in 2014 and 2015—and a strong geographic concentration in the southern coastal basin and Central Valley (Figure 2). We interpret the reported case counts as a noisy but epidemiologically meaningful index of transmission intensity, acknowledging the reporting and testing biases inherent in passive surveillance (McDonald et al., 2021; Yeung et al., 2017).

2.3 Streamflow data acquisition and gauge audit

Daily streamflow records were retrieved from the USGS National Water Information System (NWIS), using daily mean discharge (parameter code 00060, statistic code 00003) reported in cubic feet per second (U.S. Geological Survey, 2024). We began by identifying all NWIS surface-water sites in California with discharge records, yielding 2,397 candidate gauges. Because a predictor is only useful if it is consistently observed, we imposed a stringent period-of-record coverage requirement: a gauge was retained only if it reported valid daily discharge on at least 80% of the 6,574 days spanning the study period. Of the 2,397 candidate gauges, 1,584 were screened out for insufficient period-of-record overlap before download, and of the 804 gauges whose data were downloaded and evaluated, 703 met the 80% coverage threshold and were retained. The retained network comprised 4,545,375 daily discharge observations with a mean per-gauge coverage of 98.3% and a minimum of 80.0% (Figure 3a). Each county was assigned the gauges falling within its boundary; the four counties lacking any qualifying in-county gauge (San Francisco, Glenn, Lassen, and Kings) were assigned the record of their nearest qualifying neighbor, with nearest-neighbor distances of 42–61 km computed between county centroids obtained from the U.S. Census Bureau gazetteer (U.S. Census Bureau, 2023).

2.4 Hydrological standardization

Raw discharge is not comparable across gauges—a large lowland river and a small ephemeral creek differ by orders of magnitude in absolute flow—nor across seasons, given the strong Mediterranean seasonality of California hydrology. To render discharge comparable and to isolate anomalies from the expected seasonal cycle, we standardized each daily value into a per-gauge, day-of-year z -score. For each gauge and each calendar day-of-year, we computed the mean and standard deviation of log-transformed discharge across all years in the record, and expressed each observation as its deviation from that day-of-year climatology in standard-deviation units. This standardization removes both the between-gauge magnitude differences and the deterministic seasonal signal, so that positive values indicate a gauge running wetter than its own seasonal norm and negative values indicate drier-than-normal conditions (Figure 3b). County-level daily anomalies were obtained by averaging the standardized series across the gauges assigned to each county.

2.5 Feature engineering

From the standardized county series we engineered two families of antecedent hydrological predictors designed to capture ecologically distinct pathways by which water availability might influence early-season mosquito production. The first, a *spring low-flow proportion*, quantified drought-like conditions as the fraction of spring days (defined over the pre-transmission window) on which the standardized discharge fell below a low-flow threshold; higher values indicate a persistently dry antecedent spring, the condition hypothesized by [Paull et al. \(2017\)](#) to intensify transmission through host and vector concentration. The second family captured *dry-to-wet pulses*—episodes in which discharge transitions rapidly from below-normal to substantially above-normal, potentially flooding riparian and floodplain habitats favorable to *Culex tarsalis*. We summarized these pulses with two predictors: a binary *pulse occurrence* indicating whether at least one qualifying dry-to-wet transition took place during the antecedent window, and a continuous *pulse magnitude* quantifying the largest such transition. All hydrological predictors were computed over the antecedent (winter–spring) window preceding each transmission season, ensuring that they used only information available before the outbreak period and were therefore admissible as genuine forecast predictors.

2.6 Climate covariates

To construct a biologically grounded, non-hydrologic baseline, we obtained daily maximum and minimum temperature and daily precipitation from the Daymet Version 4 gridded surface-meteorology product, queried at the single 1-km pixel containing each county centroid ([Thornton et al., 2021](#)). From these we derived two climate predictors over the antecedent spring window. The first was accumulated *growing degree days* with a base temperature of 10 °C (GDD_{10}), a standard entomological summary of the thermal energy available for mosquito development and viral incubation, computed as the sum over spring days of the daily mean temperature above the 10 °C threshold ([Reisen et al., 2006](#); [Shocket et al., 2020](#)). The second was accumulated *spring precipitation*, the total rainfall over the same window. These two covariates constitute a deliberately parsimonious baseline that encodes the dominant temperature dependence of WNV transmission together with a simple direct-rainfall term, against which the incremental value of streamflow could be judged.

2.7 Statistical model

We modeled the binary outbreak outcome with a mixed-effects (hierarchical) logistic regression ([Bates et al., 2015](#)) that included a random intercept for county to absorb persistent between-county differences in baseline risk arising from unmodeled factors such as land use, avian community composition, vector-control intensity, and surveillance effort. All fixed-effect predictors were standardized to zero mean and unit variance so that coefficients are interpretable as log-odds changes per standard deviation and are directly comparable in magnitude. Models were fit using the variational-Bayes binomial mixed generalized linear model implementation in `statsmodels` ([Seabold and Perktold, 2010](#)). We specified two nested models: a *baseline* model containing only the two climate covariates (GDD_{10} and spring precipitation) plus the county random intercept, and a *hydrology* model that added the three streamflow predictors (low-flow proportion, pulse occurrence, and pulse magnitude). The contrast between these two models isolates the incremental predictive contribution of antecedent streamflow beyond the climate baseline.

2.8 Cross-validation and evaluation

To estimate genuine out-of-sample forecasting skill and to guard against overfitting in a temporally autocorrelated dataset, we evaluated both models under leave-one-year-out (LOYO) cross-validation (Roberts et al., 2017). In each of the 18 folds, the model was trained on all county-seasons from 17 years and used to predict outbreak probabilities for the held-out year, so that every prediction was made for a year absent from training. We summarized discrimination using the area under the precision–recall curve (AUPRC), which is the most informative metric under the moderate class imbalance of our outcome (Saito and Rehmsmeier, 2015), alongside the area under the receiver-operating-characteristic curve (ROC-AUC). We assessed probabilistic accuracy with the Brier score and evaluated calibration through the slope and intercept of a logistic recalibration of observed outcomes on predicted log-odds, where a slope near one and intercept near zero indicate well-calibrated probabilities (Vehtari et al., 2017). Fixed-effect coefficients from the full-data fit were reported as odds ratios per standard deviation with 95% intervals, and the statistical significance of the streamflow terms was assessed with Benjamini–Hochberg control of the false-discovery rate across the three hydrological predictors (Danforth et al., 2022).

2.9 Decision rule and county prioritization

We pre-specified a decision trigger to translate the model comparison into an operational recommendation: antecedent streamflow would be judged to add practically meaningful value, and therefore be recommended for inclusion in a forecasting pipeline, only if the hydrology model improved out-of-sample AUPRC by at least +0.02 relative to the baseline. Independently of that trigger, we used the fitted model to generate a surveillance-prioritization ranking. For each county we averaged the predicted outbreak probability across all study years and ranked counties accordingly, comparing the ranking against the observed historical outbreak rate as a validation check and reporting the average values of the key climate and hydrological drivers for the highest-priority counties.

3 Results

3.1 Case burden and gauge network

Human WNV activity in California over 2006–2023 was substantial and highly variable from year to year (Figure 2). Statewide case totals ranged from a low of 272 in 2006 to peaks of 801 and 782 in the epidemic years of 2014 and 2015, respectively, before declining to 433 by 2023, for a cumulative 6,360 reported human cases. The burden was strongly concentrated geographically: Los Angeles County alone accounted for 1,877 cases, followed by Orange (622), Kern (457), Riverside (391), Stanislaus (348), Fresno (327), San Bernardino (290), Butte (250), Sacramento (243), and Tulare (204) counties, together comprising the great majority of the state total. This concentration in the southern coastal basin and the Central Valley motivates a county-resolved modeling approach and foreshadows the prioritization presented below.

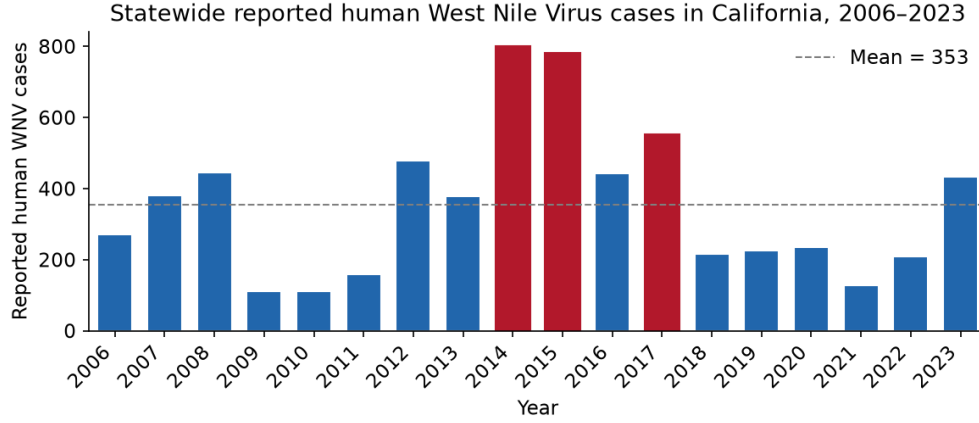


Figure 2: Statewide human West Nile virus case counts in California by year, 2006–2023. Case burden was highly variable, with epidemic peaks in 2014 and 2015 exceeding 780 cases, and a cumulative total of 6,360 reported human cases over the study period.

The gauge audit yielded a high-quality streamflow network (Figure 3a). Of 2,397 candidate USGS discharge sites in California, 1,584 were screened out for insufficient temporal overlap with the study period before download, leaving 804 gauges whose records were downloaded and evaluated. Applying the 80% period-of-record coverage threshold retained 703 gauges, discarding a further 101 for sparse records. The retained network contributed 4,545,375 daily discharge observations, with per-gauge coverage averaging 98.3% and never falling below 80.0%. The per-gauge, day-of-year standardization successfully removed the strong seasonal and between-gauge magnitude structure, producing anomaly series in which wet and dry departures from each gauge’s own climatology are directly comparable (Figure 3b).

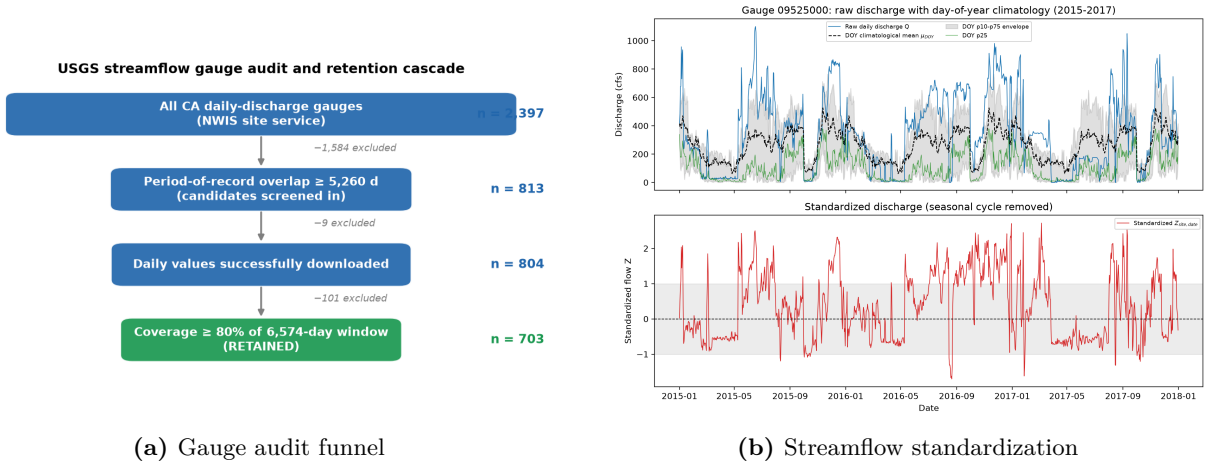


Figure 3: Streamflow data pipeline. (a) The gauge audit reduced 2,397 candidate USGS discharge sites to 703 gauges meeting the 80% period-of-record coverage threshold. (b) Example of per-gauge, day-of-year standardization converting raw discharge into seasonally adjusted anomaly z -scores, removing seasonal and magnitude structure so that wet and dry departures are comparable across gauges.

3.2 Model comparison: baseline versus hydrology

The parsimonious climate baseline was already a strong predictor of county-season outbreaks under leave-one-year-out validation (Table 1). It achieved an out-of-sample AUPRC of 0.813 against a background outbreak prevalence of 0.230—indicating discrimination far above the no-skill level—together with a ROC-AUC of 0.940 and a Brier score of 0.083. Its calibration slope of 0.741 indicated mild overconfidence, consistent with the known tendency of variational-Bayes inference to underestimate posterior variance, while the calibration intercept of -0.041 indicated negligible systematic bias.

Adding the three antecedent streamflow predictors produced essentially no change in out-of-sample performance (Table 1). The hydrology model attained an AUPRC of 0.814, a ROC-AUC of 0.940, and a Brier score of 0.083—statistically and practically indistinguishable from the baseline. The incremental AUPRC was $+0.0007$, more than an order of magnitude below the pre-registered improvement threshold of $+0.02$. The calibration slope in fact declined slightly (0.728 versus 0.741), indicating no calibration benefit from the added complexity. The near-perfect agreement between the two models is visually apparent in the precision–recall curves (Figure 4a), the calibration curves (Figure 4b), and a scatter of the per-observation predicted probabilities, which fall almost exactly on the identity line (Figure 4c).

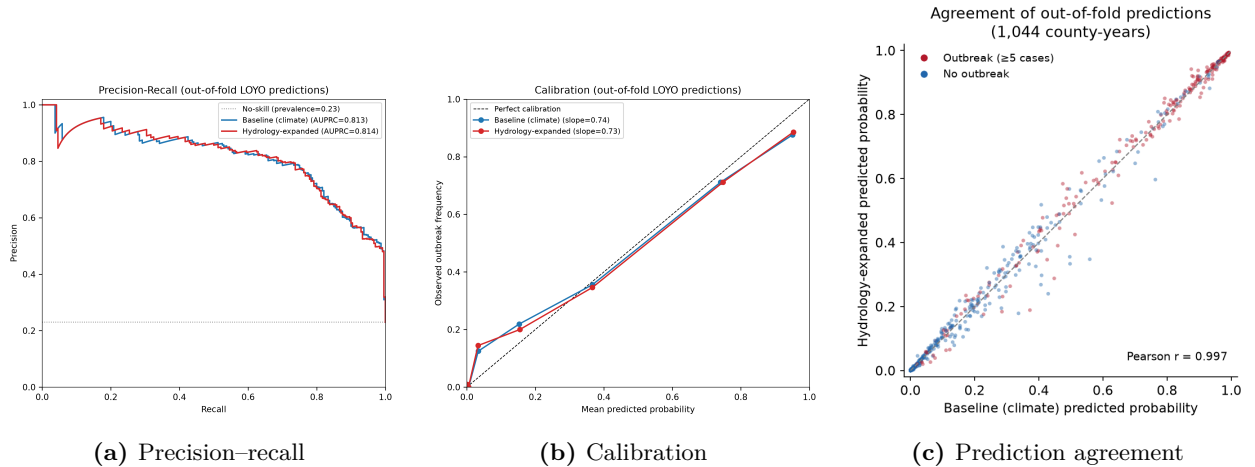


Figure 4: Baseline (climate-only) and hydrology-augmented models are essentially indistinguishable out-of-sample. (a) Leave-one-year-out precision–recall curves nearly overlap (AUPRC 0.813 vs. 0.814). (b) Calibration curves show both models are similarly, mildly overconfident. (c) Per-county-season predicted outbreak probabilities from the two models fall almost exactly on the identity line.

Table 1: Out-of-sample performance of the baseline (climate-only) and hydrology-augmented models under leave-one-year-out cross-validation. The incremental change in AUPRC (+0.0007) falls far short of the pre-registered improvement trigger of +0.02, so streamflow is not recommended for inclusion. Outbreak prevalence was 0.230 (240 of 1,044 county-seasons).

Metric	Baseline	Hydrology	Δ (Hydro – Base)	Trigger
AUPRC	0.8130	0.8138	+0.0007	FAIL (< 0.02)
ROC-AUC	0.9405	0.9401	−0.0004	—
Brier score	0.0827	0.0826	−0.0001	—
Calibration slope	0.7410	0.7275	−0.0135	—
Calibration intercept	−0.041	−0.055	−0.014	—

3.3 Coefficient estimates

The fixed-effect estimates clarify why streamflow contributed so little (Table 2; Figure 5). In both models the accumulated spring growing degree days were the overwhelmingly dominant predictor, with an odds ratio of approximately 8.4–8.5 per standard deviation ($p \approx 6 \times 10^{-65}$), meaning that a one-standard-deviation increase in antecedent warmth multiplied the odds of an outbreak more than eightfold. Spring precipitation carried a modest protective association (odds ratio ≈ 0.62 – 0.65 per standard deviation; $p < 0.01$), consistent with wetter springs corresponding to cooler conditions less favorable to transmission. Against this backdrop the streamflow terms were small. The low-flow proportion had an odds ratio of 0.797 per standard deviation but did not reach significance (raw $p = 0.085$). The two pulse predictors were nominally significant after false-discovery-rate correction (both BH-adjusted $p = 0.004$) but pointed in opposing directions—pulse occurrence was associated with lower outbreak odds (odds ratio 0.653) while pulse magnitude was associated with higher odds (odds ratio 1.540)—an internally offsetting pattern that limits any coherent predictive contribution and is consistent with their negligible effect on out-of-sample skill.

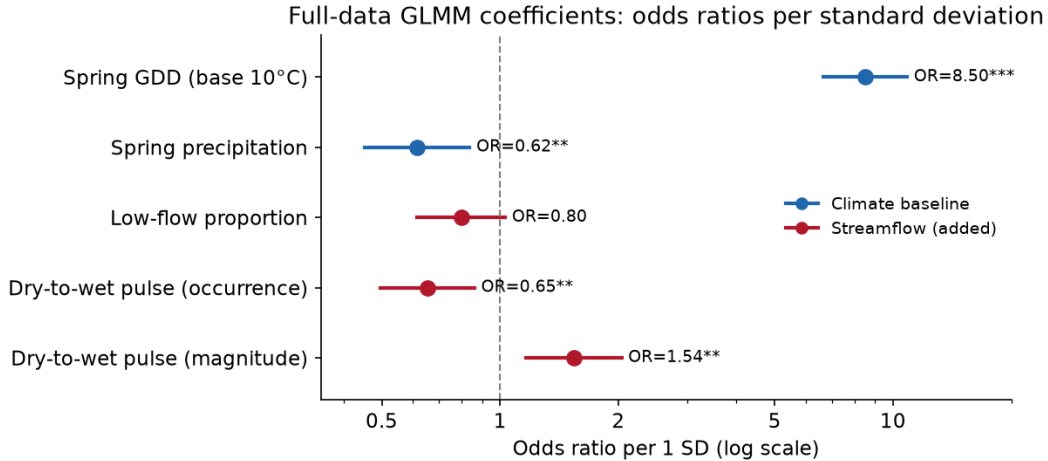


Figure 5: Forest plot of fixed-effect odds ratios per standard deviation (95% CI) from the hydrology-augmented mixed-effects logistic regression. Spring growing degree days (base 10 °C) dominate with an odds ratio near 8.5, dwarfing the streamflow terms. The dry-to-wet pulse predictors are mutually offsetting (occurrence protective, magnitude harmful), and the low-flow proportion is not statistically significant.

Table 2: Fixed-effect odds ratios per standard deviation from the hydrology-augmented mixed-effects logistic regression (full-data fit), with 95% confidence intervals and false-discovery-rate-adjusted p -values for the streamflow terms. Growing degree days dominate; the streamflow predictors are small and, for the pulse pair, mutually offsetting.

Predictor	Odds ratio / SD	95% CI	p -value	BH-FDR p
Spring GDD (base 10 °C)	8.50	6.64–10.87	6×10^{-65}	—
Spring precipitation	0.62	0.45–0.84	0.002	—
Low-flow proportion	0.80	0.62–1.03	0.085	0.085
Pulse occurrence	0.65	0.50–0.86	0.003	0.004
Pulse magnitude	1.54	1.16–2.04	0.003	0.004

3.4 Trigger decision

Because the observed incremental AUPRC of +0.0007 fell far below the pre-registered +0.02 threshold, the decision rule returned a clear **FAIL**: antecedent streamflow is not recommended for inclusion in an operational county-season WNV forecasting model for California. The baseline climate model, driven principally by growing degree days, should be preferred on grounds of parsimony, data availability, and equivalent performance.

3.5 County prioritization

Although streamflow did not improve prediction, the fitted model provides a well-calibrated basis for prioritizing surveillance. Ranking counties by their average predicted outbreak probability produced a top-10 list that closely tracked observed historical outbreak frequency, validating the model’s spatial fidelity (Table 3; Figure 6). The highest-priority counties were Stanislaus, Kern, Fresno, Los Angeles, Tulare, Orange, San Joaquin, Sacramento, Riverside, and Butte—a set dominated by the warm Central Valley and southern basin counties that also carried the greatest historical case burden. For every county in the top 10, the model-predicted probability and the observed historical outbreak rate agreed to within a few percentage points, and the baseline and hydrology models produced nearly identical rankings, reinforcing that the prioritization is driven by climate rather than hydrology. The average spring growing degree days for these counties were uniformly high (typically 700–1,080 degree-days), whereas the average streamflow drivers varied idiosyncratically and showed no systematic relationship with priority rank.

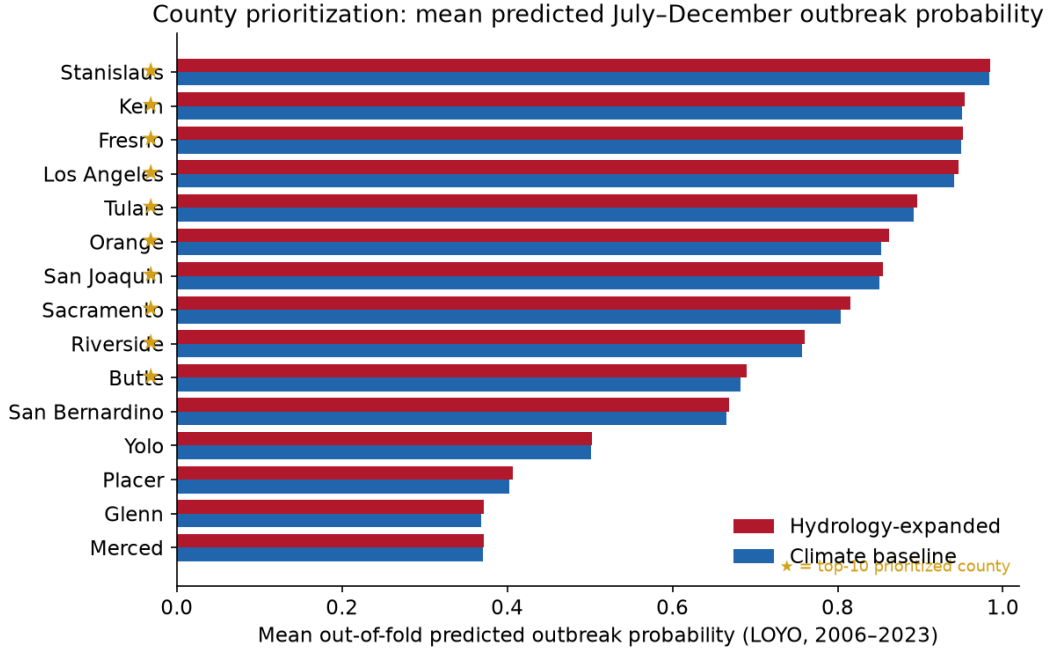


Figure 6: Top prioritized California counties ranked by average model-predicted West Nile virus outbreak probability (2006–2023). Predicted probabilities closely track observed historical outbreak rates, and the warm Central Valley and southern-basin counties dominate the ranking.

Table 3: Top 10 counties prioritized for West Nile virus surveillance by average model-predicted outbreak probability, 2006–2023. Predicted probabilities from the hydrology and baseline models are nearly identical and both closely match the observed historical outbreak rate. Average spring growing degree days (GDD, base 10 °C) and the mean low-flow proportion illustrate that the prioritization is driven by thermal, not hydrological, conditions.

Rank	County	Pred. prob. (hydro)	Hist. rate	Outbreak yrs (of 18)	Avg spring GDD (base 10)	Avg low-flow prop.
1	Stanislaus	0.985	1.000	18	923	0.093
2	Kern	0.954	0.944	17	1008	0.085
3	Fresno	0.952	0.944	17	1039	0.089
4	Los Angeles	0.946	0.944	17	715	0.066
5	Tulare	0.896	0.889	16	769	0.084
6	Orange	0.862	0.833	15	820	0.048
7	San Joaquin	0.855	0.833	15	898	0.101
8	Sacramento	0.815	0.778	14	871	0.065
9	Riverside	0.760	0.722	13	1077	0.055
10	Butte	0.690	0.667	12	888	0.089

4 Discussion

Our central result is a clear and carefully validated negative finding: at the county-transmission-season scale, and relative to a parsimonious climate baseline dominated by growing degree days,

antecedent streamflow adds negligible predictive value for West Nile virus outbreaks in California. The baseline model already discriminated outbreak from non-outbreak county-seasons with an out-of-sample AUPRC of 0.813, and augmenting it with three ecologically motivated streamflow predictors changed that figure by only +0.0007—roughly one part in a thousand and more than an order of magnitude short of the pre-registered threshold for practical relevance. This is not a failure of the analysis but an informative answer to a well-posed operational question, and it carries direct implications for how forecasting resources should be allocated.

Several complementary explanations account for why streamflow contributed so little. The most immediate is statistical redundancy: growing degree days are such a powerful predictor—an odds ratio exceeding eight per standard deviation—that they leave little residual variance for other covariates to explain. The strong thermal dependence of WNV transmission is among the best-established findings in the field, reflecting temperature’s simultaneous acceleration of larval development, biting frequency, and, above all, the extrinsic incubation period that governs whether an infected mosquito survives long enough to become infectious (Reisen et al., 2006; Kilpatrick et al., 2008; Shocket et al., 2020). When a single mechanism dominates the biology this strongly, additional environmental covariates must carry genuinely orthogonal information to improve prediction, and antecedent streamflow evidently does not at this scale.

A second, ecologically deeper explanation concerns the decoupling of natural streamflow from the aquatic habitat that actually produces California’s WNV vectors. The state’s mosquito populations are sustained to a remarkable degree by anthropogenic water: irrigated agriculture in the Central Valley, urban storm-drain and catch-basin systems, neglected swimming pools, and managed wetlands. Kovach and Kilpatrick (2024) showed that irrigation increases *Culex* abundance by 17- to 21-fold and stabilizes it against natural precipitation variability, explaining roughly a third of county-level variation in human incidence, while Bhattachan et al. (2021) demonstrated that residential water-use restrictions—not rainfall—suppressed vector abundance during the 2012–2016 drought. In such a heavily managed hydrological system, discharge measured at USGS river gauges may simply not register the water that matters for larval production, which resides in irrigation infrastructure and urban containers rather than in gauged stream channels. This interpretation aligns with the drought-centric findings of Paull et al. (2017), whose signal operated through host and vector concentration and infection prevalence rather than through streamflow-detectable habitat expansion.

A third explanation lies in the spatial and temporal mismatch between our streamflow features and the processes they were meant to capture. County-averaged, day-of-year-standardized discharge is a coarse summary that may blur the fine-scale flooding and pooling events most relevant to *Culex tarsalis* floodwater habitat (Tall and Gatton, 2019). The internally offsetting behavior of our two pulse predictors—occurrence associated with lower odds, magnitude with higher odds—suggests that dry-to-wet transitions have context-dependent effects that a county-season aggregation cannot resolve. It remains possible that streamflow carries genuine signal at finer spatial resolution, at sub-seasonal timescales, or in specific floodwater-dominated landscapes, even though it is uninformative in the aggregate.

Notwithstanding the negative predictive result, the analysis delivered a validated, immediately actionable output: a 10-county surveillance-prioritization ranking whose predicted probabilities closely match observed historical outbreak frequencies. That the ranking is driven by thermal geography—concentrating on the warm Central Valley and southern basin counties that dominate the state’s case burden—reinforces the primacy of temperature and provides vector-control agencies with a defensible basis for targeting effort, including the timing of adulticide interventions whose efficacy in reducing WNV transmission has been documented in California (Carney et al., 2008).

The near-identity of the baseline and hydrology rankings further underscores that this prioritization requires only widely available climate data, not the substantial infrastructure of gauge auditing and hydrological feature engineering.

4.1 Limitations

Five limitations qualify our conclusions. First, the mixed-effects models were fit using variational-Bayes inference, which is known to underestimate posterior variance and can yield slightly overconfident predictions; this is reflected in the calibration slopes near 0.73–0.74, below the ideal of 1.0, and a future sensitivity analysis using Markov chain Monte Carlo (e.g., in PyMC or Stan) would provide fuller posteriors and more honest uncertainty quantification (Morris et al., 2019). Second, four counties without a qualifying in-county stream gauge—San Francisco, Glenn, Lassen, and Kings—were assigned the hydrology of their nearest gauged neighbor at distances of 42–61 km, an approximation that may fail across watershed or climate boundaries. Third, meteorological covariates were drawn from a single Daymet pixel at each county centroid rather than area-weighted across the county polygon, which may not capture climate variation within large or topographically complex counties such as San Bernardino and Riverside (Thornton et al., 2021). Fourth, human case counts are subject to reporting and testing biases—driven by variation in surveillance intensity, health-seeking behavior, and diagnostic capacity across counties and years—that introduce non-random measurement error into the outbreak target (McDonald et al., 2021; Yeung et al., 2017). Fifth, the model represents between-county heterogeneity only through independent random intercepts and does not model spatial adjacency through a conditional autoregressive or similar structure, so residual spatial spillover between neighboring counties is not captured (Riebler et al., 2016; Morris et al., 2019). Each of these limitations bears principally on the precision of uncertainty estimates and the treatment of a handful of counties rather than on the central comparison, which is robust: the incremental value of streamflow is so small that no plausible refinement of these details would lift it above the practical-relevance threshold.

5 Conclusion

We set out to determine whether antecedent streamflow, an attractive and physically integrative proxy for landscape wetness, improves county-scale prediction of West Nile virus outbreaks in California beyond what a simple, temperature-aware climate model already achieves. Across 18 transmission seasons and 58 counties, evaluated under leave-one-year-out cross-validation, the answer is no: a baseline model driven overwhelmingly by spring growing degree days discriminated outbreaks with an AUPRC of 0.81, and adding audited, standardized streamflow features changed out-of-sample skill by less than one part in a thousand, failing our pre-specified improvement trigger. The result is best understood as a consequence of the dominance of temperature in WNV transmission combined with the decoupling of gauged streamflow from the anthropogenic water sources—irrigation, urban infrastructure, and managed wetlands—that actually sustain California’s *Culex* vectors. For operational forecasting, this argues for parsimonious, temperature-centered models that avoid the substantial cost of hydrological data pipelines without sacrificing accuracy. At the same time, the validated 10-county prioritization we derived offers immediate practical utility for surveillance targeting. Future work should test whether streamflow carries signal at finer spatial and temporal resolution or in specific floodwater-dominated landscapes, incorporate explicit spatial structure and full Bayesian uncertainty, and—most promisingly—develop predictors that directly represent the managed and urban water sources that this analysis suggests are the true hydrological

drivers of WNV risk in the arid West.

Data and code availability

Human West Nile virus case data are available from the California Department of Public Health California Arbovirus Surveillance program ([California Department of Public Health, 2024](#)). Daily streamflow data are available from the USGS National Water Information System ([U.S. Geological Survey, 2024](#)). Gridded meteorology is available from the Daymet Version 4 product ([Thornton et al., 2021](#)), and county centroids from the U.S. Census Bureau gazetteer ([U.S. Census Bureau, 2023](#)). Analyses were performed in Python using `statsmodels` ([Seabold and Perktold, 2010](#)).

References

- Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.
- Beeman, S. P., Morrison, A. M., Unnasch, T. R., and Unnasch, R. S. (2021). Ensemble ecological niche modeling of West Nile virus probability in Florida. *PLOS ONE*, 16(10):e0256868.
- Bhattachan, A., Skaff, N. K., Irish, A. M., Vimal, S., Remais, J. V., and Lettenmaier, D. P. (2021). Outdoor residential water use restrictions during recent drought suppressed disease vector abundance in southern California. *Environmental Science & Technology*, 55(1):478–487.
- California Department of Public Health (2024). West Nile virus cases, 2006–present. California Health and Human Services Open Data Portal, dataset UUID 3205b420-3f62-4a02-8d2e-9a9ed34c49f4. Available at <https://data.chhs.ca.gov/>.
- Campbell, R., Thiemann, T. C., Lemenager, D., and Reisen, W. K. (2013). Host-selection patterns of *Culex tarsalis* (Diptera: Culicidae) determine the spatial heterogeneity of West Nile virus enzootic activity in northern California. *Journal of Medical Entomology*, 50(6):1303–1309.
- Carney, R. M., Husted, S., Jean, C., Glaser, C., and Kramer, V. (2008). Efficacy of aerial spraying of mosquito adulticide in reducing incidence of West Nile virus, California, 2005. *Emerging Infectious Diseases*, 14(5):747–754.
- Danforth, M. E., Snyder, R. E., Lonstrup, E. T. N., Barker, C. M., and Kramer, V. L. (2022). Evaluation of the effectiveness of the California mosquito-borne virus surveillance & response plan, 2009–2018. *PLOS Neglected Tropical Diseases*, 16(5):e0010375.
- Goddard, L. B., Roth, A. E., Reisen, W. K., and Scott, T. W. (2002). Vector competence of California mosquitoes for West Nile virus. *Emerging Infectious Diseases*, 8(12):1385–1391.
- Hartley, D. M., Barker, C. M., Le Menach, A., Niu, T., Gaff, H. D., and Reisen, W. K. (2012). Effects of temperature on emergence and seasonality of West Nile virus in California. *The American Journal of Tropical Medicine and Hygiene*, 86(5):884–894.
- Holcomb, K. M., Reiner, R. C., and Barker, C. M. (2021). Spatio-temporal impacts of aerial adulticide applications on populations of West Nile virus vector mosquitoes. *Parasites & Vectors*, 14:120.

- Holcomb, K. M., Staples, J. E., Nett, R. J., Beard, C. B., Petersen, L. R., Benjamin, S. G., Green, B. W., Jones, H., and Johansson, M. A. (2023). Multi-model prediction of West Nile virus neuroinvasive disease with machine learning for identification of important regional climatic drivers. *GeoHealth*, 7(11):e2023GH000906.
- Jansen, S., Heitmann, A., Lühken, R., Leggewie, M., Helms, M., Badusche, M., Rossini, G., Schmidt-Chanasit, J., and Tannich, E. (2019). *Culex torrentium*: A potent vector for the transmission of West Nile virus in Central Europe. *Viruses*, 11(6):492.
- Kilpatrick, A. M. (2011). Globalization, land use, and the invasion of West Nile virus. *Science*, 334(6054):323–327.
- Kilpatrick, A. M., Meola, M. A., Moudy, R. M., and Kramer, L. D. (2008). Temperature, viral genetics, and the transmission of West Nile virus by *Culex pipiens* mosquitoes. *PLoS Pathogens*, 4(6):e1000092.
- Kovach, T. J. and Kilpatrick, A. M. (2024). Irrigation increases and stabilizes mosquito populations and increases West Nile virus incidence. *Scientific Reports*, 14:19913.
- LaDeau, S. L., Kilpatrick, A. M., and Marra, P. P. (2007). West Nile virus emergence and large-scale declines of North American bird populations. *Nature*, 447(7145):710–713.
- McDonald, E., Mathis, S., Martin, S. W., Staples, J. E., Fischer, M., and Lindsey, N. P. (2021). Surveillance for West Nile virus disease — United States, 2009–2018. *MMWR Surveillance Summaries*, 70(1):1–15.
- Morris, M., Wheeler-Martin, K., Simpson, D., Mooney, S. J., Gelman, A., and DiMaggio, C. (2019). Bayesian hierarchical spatial models: Implementing the Besag York Mollié model in Stan. *Spatial and Spatio-temporal Epidemiology*, 31:100301.
- Paull, S. H., Horton, D. E., Ashfaq, M., Rastogi, D., Kramer, L. D., Diffenbaugh, N. S., and Kilpatrick, A. M. (2017). Drought and immunity determine the intensity of West Nile virus epidemics and climate change impacts. *Proceedings of the Royal Society B: Biological Sciences*, 284(1848):20162078.
- Paz, S. (2015). Climate change impacts on West Nile virus transmission in a global context. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1665):20130561.
- Petersen, L. R., Brault, A. C., and Nasci, R. S. (2013). West Nile virus: Review of the literature. *JAMA*, 310(3):308–315.
- Reisen, W. K. (2013). Ecology of West Nile virus in North America. *Viruses*, 5(9):2079–2105.
- Reisen, W. K., Fang, Y., and Martinez, V. M. (2006). Effects of temperature on the transmission of West Nile virus by *Culex tarsalis* (Diptera: Culicidae). *Journal of Medical Entomology*, 43(2):309–317.
- Riebler, A., Sørbye, S. H., Simpson, D., and Rue, H. (2016). An intuitive Bayesian spatial model for disease mapping that accounts for scaling. *Statistical Methods in Medical Research*, 25(4):1145–1165.

- Roberts, D. R., Bahn, V., Ciuti, S., Boyce, M. S., Elith, J., Guillera-Arroita, G., Hauenstein, S., Lahoz-Monfort, J. J., Schröder, B., Thuiller, W., Warton, D. I., Wintle, B. A., Hartig, F., and Dormann, C. F. (2017). Cross-validation strategies for data with temporal, spatial, hierarchical, or phylogenetic structure. *Ecography*, 40(8):913–929.
- Ryan, S. J., Carlson, C. J., Mordecai, E. A., and Johnson, L. R. (2019). Global expansion and redistribution of Aedes-borne virus transmission risk with climate change. *PLOS Neglected Tropical Diseases*, 13(3):e0007213.
- Saito, T. and Rehmsmeier, M. (2015). The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. *PLOS ONE*, 10(3):e0118432.
- Seabold, S. and Perktold, J. (2010). statsmodels: Econometric and statistical modeling with Python. In *Proceedings of the 9th Python in Science Conference (SciPy 2010)*, pages 92–96.
- Shaman, J., Day, J. F., and Komar, N. (2010). Hydrologic conditions describe West Nile virus risk in Colorado. *International Journal of Environmental Research and Public Health*, 7(2):494–508.
- Shaman, J., Harding, K., and Campbell, S. R. (2011). Meteorological and hydrological influences on the spatial and temporal prevalence of West Nile virus in Culex mosquitoes, Suffolk County, New York. *Journal of Medical Entomology*, 48(4):867–875.
- Shocket, M. S., Verwillow, A. B., Numazu, M. G., Slamani, H., Cohen, J. M., El Moustaid, F., Rohr, J., Johnson, L. R., and Mordecai, E. A. (2020). Transmission of West Nile and five other temperate mosquito-borne viruses peaks at temperatures between 23°C and 26°C. *eLife*, 9:e58511.
- Snyder, R. E., Feiszli, T., Foss, L., Messenger, S., Fang, Y., Barker, C. M., Reisen, W. K., Vugia, D. J., Padgett, K., and Kramer, V. L. (2020). West Nile virus in California, 2003–2018: A persistent threat. *PLOS Neglected Tropical Diseases*, 14(11):e0008841.
- Soverow, J. E., Wellenius, G. A., Fisman, D. N., and Mittleman, M. A. (2009). Infectious disease in a warming world: How weather influenced West Nile virus in the United States (2001–2005). *Environmental Health Perspectives*, 117(7):1049–1052.
- Tall, J. A. and Gatton, M. L. (2019). Flooding and arboviral disease: Predicting Ross River virus disease outbreaks across inland regions of south-eastern Australia. *Journal of Medical Entomology*, 57(1):241–251.
- Thomson, M. C. and Stanberry, L. R. (2022). Climate change and vectorborne diseases. *New England Journal of Medicine*, 387(21):1969–1978.
- Thornton, P. E., Shrestha, R., Thornton, M., Kao, S.-C., Wei, Y., and Wilson, B. E. (2021). Gridded daily weather data for North America with comprehensive uncertainty quantification. *Scientific Data*, 8:190.
- Tonks, A., Harris, T., Li, B., Brown, W., and Smith, R. (2024). Forecasting West Nile virus with graph neural networks: Harnessing spatial dependence in irregularly sampled geospatial data. *GeoHealth*, 8(7):e2023GH000784.
- U.S. Census Bureau (2023). 2023 U.S. gazetteer files: Counties. U.S. Census Bureau Geography Program.

- U.S. Geological Survey (2024). USGS water data for the nation: National water information system (NWIS) daily values, parameter 00060 (discharge), statistic 00003 (mean). U.S. Geological Survey water data. Accessed via <https://waterservices.usgs.gov/>.
- Vehtari, A., Gelman, A., and Gabry, J. (2017). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5):1413–1432.
- Walsh, M. G. (2012). The role of hydrogeography and climate in the landscape epidemiology of West Nile virus in New York State from 2000 to 2010. *PLoS ONE*, 7(2):e30620.
- Yeung, M. W., Shing, E., Nelder, M., and Sander, B. (2017). Epidemiologic and clinical parameters of West Nile virus infections in humans: A scoping review. *BMC Infectious Diseases*, 17:609.