

UC Davis

UC Davis Previously Published Works

Title

A proposed framework for the development and qualitative evaluation of West Nile virus models and their application to local public health decision-making

Permalink

<https://escholarship.org/uc/item/7vr815b1>

Journal

PLOS Neglected Tropical Diseases, 15(9)

ISSN

1935-2727

Authors

Keyel, Alexander C

Gorris, Morgan E

Rochlin, Ilia

et al.

Publication Date

2021

DOI

10.1371/journal.pntd.0009653

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

REVIEW

A proposed framework for the development and qualitative evaluation of West Nile virus models and their application to local public health decision-making

Alexander C. Keyel^{1,2*}, Morgan E. Gorris³, Iliia Rochlin⁴, Johnny A. Uelmen⁵, Luis F. Chaves⁶, Gabriel L. Hamer⁷, Imelda K. Moise⁸, Marta Shocket⁹, A. Marm Kilpatrick¹⁰, Nicholas B. DeFelice^{11,12,13}, Justin K. Davis¹⁴, Eliza Little¹⁵, Patrick Irwin^{16,17}, Andrew J. Tyre¹⁸, Kelly Helm Smith¹⁹, Chris L. Fredregill²⁰, Oliver Elison Timm²¹, Karen M. Holcomb²¹, Michael C. Wimberly¹⁴, Matthew J. Ward^{22,23}, Christopher M. Barker²¹, Charlotte G. Rhodes⁷, Rebecca L. Smith⁵

1 Division of Infectious Diseases, Wadsworth Center, New York State Department of Health, Albany, New York, United States of America, **2** Department of Atmospheric and Environmental Sciences, University at Albany, Albany, New York, United States of America, **3** Information Systems and Modeling & Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, New Mexico, United States of America, **4** Center for Vector Biology, Rutgers University, New Brunswick, New Jersey, United States of America, **5** Department of Pathobiology, College of Veterinary Medicine, University of Illinois at Urbana-Champaign, Urbana, Illinois, United States of America, **6** Instituto Costarricense de Investigación y Enseñanza en Nutrición y Salud (INCIENSA), Tres Rios, Cartago, Costa Rica, **7** Department of Entomology, Texas A&M University, College Station, Texas, United States of America, **8** Department of Geography & Regional Studies, University of Miami, Coral Gables, Florida, United States of America, **9** Department of Ecology and Evolutionary Biology, University of California, Los Angeles, California, United States of America, **10** Department of Ecology and Evolutionary Biology, University of California, Santa Cruz, California, United States of America, **11** Department of Environmental Medicine and Public Health, Icahn School of Medicine at Mount Sinai, New York, New York, United States of America, **12** Institute for Exposomic Research, Icahn School of Medicine at Mount Sinai, New York, New York, United States of America, **13** Department of Environmental Health Sciences, Mailman School of Public Health, Columbia University, New York, New York, United States of America, **14** Department of Geography and Environmental Sustainability, University of Oklahoma, Norman, Oklahoma, United States of America, **15** Connecticut Agricultural Experimental Station, New Haven, Connecticut, United States of America, **16** Northwest Mosquito Abatement District, Wheeling, Illinois, United States of America, **17** Department of Entomology, University of Wisconsin-Madison, Madison, Wisconsin, United States of America, **18** School of Natural Resources, University of Nebraska-Lincoln, Lincoln, Nebraska, United States of America, **19** National Drought Mitigation Center, University of Nebraska-Lincoln, Lincoln, Nebraska, United States of America, **20** Mosquito and Vector Control Division, Harris County Public Health, Houston, Texas, United States of America, **21** Department of Pathology, Microbiology, and Immunology, University of California Davis, California, United States of America, **22** Environmental Analytics Group, Universities Space Research Association, NASA Ames Research Center, Moffett Field, California, United States of America, **23** Department of Tropical Medicine, Tulane University School of Public Health & Tropical Medicine, New Orleans, Louisiana, United States of America

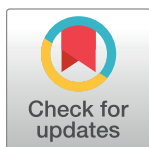
* These authors contributed equally to this work.

‡ OET, KMH, MCW, MJW, CMB, and CGR also contributed equally to this work.

* akeyel@albany.edu

Abstract

West Nile virus (WNV) is a globally distributed mosquito-borne virus of great public health concern. The number of WNV human cases and mosquito infection patterns vary in space and time. Many statistical models have been developed to understand and predict WNV geographic and temporal dynamics. However, these modeling efforts have been disjointed



OPEN ACCESS

Citation: Keyel AC, Gorris ME, Rochlin I, Uelmen JA, Chaves LF, Hamer GL, et al. (2021) A proposed framework for the development and qualitative evaluation of West Nile virus models and their application to local public health decision-making. *PLoS Negl Trop Dis* 15(9): e0009653. <https://doi.org/10.1371/journal.pntd.0009653>

Editor: Elvina Viennet, Australian Red Cross Lifelong, AUSTRALIA

Published: September 9, 2021

Copyright: © 2021 Keyel et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study was funded by cooperative agreements 1U01CK000509-01 (Northeast), U01CK000512 (Gulf [GLH]), 1U01CK000510-03 (Southeast), 1U01CK000516 (Pacific Southwest [CMB], including the Training Grant Program), and U01CK000505 (Midwest [RLS]) funded by the Centers for Disease Control and Prevention (CDC). Contents are solely the responsibility of the authors and do not necessarily represent the official views of the CDC or the Department of Health and Human Services. The CDC had no role in the design of the study, the collection, analysis, and interpretation of data, or in writing the manuscript. This work was also supported by the National Socio-

Environmental Synthesis Center (SESYNC, [ACK, RLS]) under funding received from the National Science Foundation DBI-1639145. Additional funding was provided by NASA Grant ECOSTRES18-0046 [NBD], Coachella Valley Mosquito and Vector Control District 141316E1-CFDC-4B63-9418-DEDB9AE8BBDB [NBD], and NASA Grant NSSC19K1233 [MCW, JKD].

Competing interests: The authors have declared that no competing interests exist.

with little model comparison and inconsistent validation. In this paper, we describe a framework to unify and standardize WNV modeling efforts nationwide. WNV risk, detection, or warning models for this review were solicited from active research groups working in different regions of the United States. A total of 13 models were selected and described. The spatial and temporal scales of each model were compared to guide the timing and the locations for mosquito and virus surveillance, to support mosquito vector control decisions, and to assist in conducting public health outreach campaigns at multiple scales of decision-making. Our overarching goal is to bridge the existing gap between model development, which is usually conducted as an academic exercise, and practical model applications, which occur at state, tribal, local, or territorial public health and mosquito control agency levels. The proposed model assessment and comparison framework helps clarify the value of individual models for decision-making and identifies the appropriate temporal and spatial scope of each model. This qualitative evaluation clearly identifies gaps in linking models to applied decisions and sets the stage for a quantitative comparison of models. Specifically, whereas many coarse-grained models (county resolution or greater) have been developed, the greatest need is for fine-grained, short-term planning models (m–km, days–weeks) that remain scarce. We further recommend quantifying the value of information for each decision to identify decisions that would benefit most from model input.

Introduction

West Nile virus (WNV) is one of the most widely distributed mosquito-borne viruses and represents a global public health threat [1,2]. In the United States, WNV is the most common vector-borne virus with at least 51,801 human cases and 2,390 fatalities reported between its introduction in 1999 and 2019 [3]. WNV has had substantial negative economic impacts through healthcare costs (about \$368 million to \$2.4 billion in Texas in 2012) [4] and in equine-related veterinary financial burdens (e.g., \$1.9 million in 2002 in North Dakota prior to the vaccine in 2004) [5]. In addition to human and veterinary disease, WNV has impacted avian populations being reported in over 300 species of birds in the US [6]. The virus killed millions of songbirds [7] and led to population declines in some species, in particular American crows (*Corvus brachyrhynchos*) [8,9], ruffed grouse (*Bonasa umbellus*) [10], and yellow-billed magpies (*Pica nuttalli*) [9,11]. Effective preparedness and prevention are vital to reduce the direct and indirect impacts of WNV on human health, the environment, and the economy.

The majority of human cases of WNV are asymptomatic (approximately 80%) [12]. Symptomatic cases are classified as either non-neuroinvasive or as neuroinvasive, with neuroinvasive cases representing <1% of cases [12,13]. Many non-neuroinvasive cases of WNV go unreported. However, due to the severity of symptoms, neuroinvasive case reports are expected to be less biased [14]. Mosquito infection rates are typically expressed as either minimum infection rates (MIRs) or maximum likelihood estimates of infection rate (MLE). MIR is in common usage; however, MLE is more accurate and conveys more information [15,16].

Numerous statistical [17] and mathematical [18] models have been developed to understand and predict WNV geographic and temporal dynamics that could be potentially used to guide vector control and public health activities. Models can be used to understand the relationships between the spatial distribution of human pathogens, vectors, the prevalence of vector-borne diseases, including social demographic and environmental predictors [19]. Notably, the model requirements and desired end result may vary depending on the decision-making

mechanisms of stakeholders, and it is unlikely that there will be a single model that is suitable for all decisions. A recent review classified 48 WNV models as risk, detection, or warning models [17]. Risk models provide spatial information about relative risk, but do not contain temporal information. Detection models make estimates for the current season but do not integrate the current year's WNV surveillance data. Early warning models include current-year WNV surveillance data. Most models make predictions at a single broad or narrow spatial or temporal scale, whereas decisions occur at multiple scales. Decisions on where and when to apply larvicide (to control immature aquatic stages) or adulticide (to control vagile adults) are usually made on a weekly timescale and a relatively fine spatial scale such as city blocks or neighborhoods. In contrast, decisions on hiring and staffing for control need to be made on a monthly or seasonal timescale at the level of the local mosquito control agency and often in advance of a transmission season.

In this review, we developed a framework for applying models to decisions and tested it on 13 representative models that have been developed to understand and predict WNV geographic and temporal dynamics. These models include descriptive statistical models, a mathematical mechanistic model, and data assimilation-based models. Specifically, we asked the question "what information does each model provide about WNV?" It is critical that state, tribal, local, or territorial public health and mosquito control agencies understand WNV model outputs and find them useful in operational support. We compared model properties, inputs, and outputs in the context of these public health decisions. Specifically, we examined the capacity and suitability of the models with respect to spatial and temporal scales to guide the timing and the locations for mosquito and virus surveillance, to support mosquito vector control decisions, and to assist in conducting public health outreach campaigns.

Materials and methods

We aimed to review models that were in active development or currently being applied to the West Nile virus system by decision-makers. Rather than identify individual models, we identified research teams studying spatiotemporal dynamics of WNV based on academic conference presentations, recent publications, involvement in one of the 5 Centers for Disease Control and Prevention (CDC) Regional Centers of Excellence, and referral to participate in a workshop hosted by the National Socio-Ecological Synthesis Center (SESYNC). We aimed to identify and include a diverse audience of participants to minimize various biases. We were initially capped at 25 participants but expanded the West Nile Virus Model Comparison Project participant list to 35 when the workshop was moved to a virtual format. With respect to primary affiliation, 74.3% were from academia, 11.4% from a department of health, and 14.3% in vector control. The 13 models included in this review cover much of the regional variation within the US (Fig 1).

By taking a research team-based approach, we were able to take an in-depth look at each of the selected models. This in-depth analysis would not have been possible with in a traditional review format. The models selected here are broadly representative of the WNV models that have been developed (e.g., statistical [20–24], data assimilation [25,26], mathematical trait-based [27], machine learning [28–30], threshold-based risk [31–33], and distributed lag approaches [34–36]). We also include a probabilistic historical null model in our comparison [37]. Our framework is reproducible providing the templates for model description (S1 Text) and instructions (S2 Text), the template questions for decision-makers (S3 Text), and detailed descriptions of each model (S4 Text), so any omitted or future models could be evaluated and compared following the framework outlined here.

This paper is organized starting with an overview of all models (Table 1), model inputs (Table 2), model outputs and predictions (Fig 2 and Table 3), and, finally, model applications

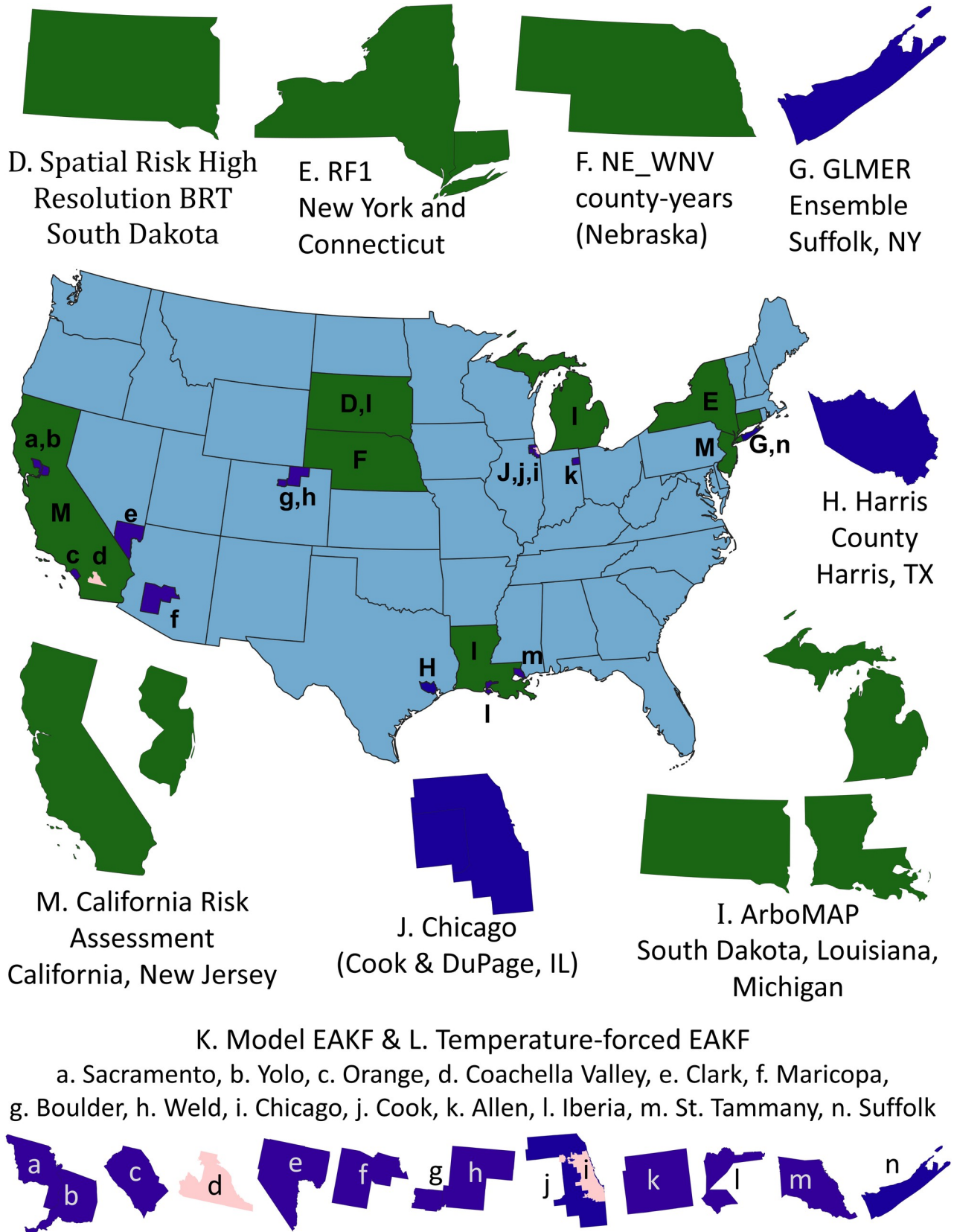


Fig 1. Map of specific locations where WNV models included in this comparison have been applied. Some models (Spatial Risk Random Forest, not shown) have been applied across the entire US. Green corresponds to analyses with state extents, blue to county extents, and pink to subcounty extents. State outlines are from Natural Earth (<https://www.naturalearthdata.com/downloads/50m-cultural-vectors/>). City of Chicago boundary is publicly available from the City of Chicago (<https://data.cityofchicago.org/Facilities-Geographic-Boundaries/Boundaries-City/ewy2-6yfk>), and county boundaries and the outline for Coachella Valley were derived from US Census tract boundaries (<https://www.census.gov/geographies/mapping-files/time-series/geo/carto-boundary-file.html>) dissolved to provide a single outline using the Dissolve algorithm in QGIS (<https://qgis.org/en/site/>). WNV, West Nile virus.

<https://doi.org/10.1371/journal.pntd.0009653.g001>

(Table 4). Decisions and models are compared with respect to temporal and spatial resolution (Figs 3–5). The model description template used to collect model information is available as [S1 Text](#), and the detailed description of all fields is provided in [S2 Text](#). This framework lays the foundation for qualitative assessment as a precursor to more rigorous quantitative comparisons for those models that are suitable for the intended purposes. The majority of the models have been published; for those models, additional details can be found in the respective publications (see citations in Table 4). Computation time will depend on the computer used, and no formal benchmarking was performed for the models. However, all models except the

Table 1. Model overview: A comparison of model class, spatial, temporal resolution, software implementation, and code availability.

Model	Class of Model ¹	Spatial Resolution	Temporal Resolution	Software	Code Available
A. Historical Null	Spatial patterns	Flexible	Annual ²	R	www.github.com/akeyel/dfmip
B. Spatial Risk Random Forest	Spatial patterns	County	Mean from 2005–2018	R	No
C. Temperature-trait-based Relative R ₀ Model	Spatial patterns	Flexible	Flexible	R	https://datadryad.org/stash/dataset/doi:10.5068/D1VW96
D. Spatial Risk High Resolution BRT Model	Spatial patterns	300 × 300 m	Mean from (2004–2017)	R	No (in progress)
E. RF1	Early warning	Flexible	Annual	R	www.github.com/akeyel/rf1
F. NE_WNV County-years	Early warning	County	Annual	R, mgcv	www.github.com/khelmsmith/flm_NE_WNV
G. GLMER Ensemble	Early warning	13 × 13 km grid	Monthly	R	No
H. Harris County	Early warning Early detection	Whole Harris County	Month	R	Based on code for SAR models presented by [38]
I. ArboMAP	Early detection	Typically county	Weekly	R	https://github.com/EcoGRAPH/ArboMAP/releases/
J. Chicago Ultra-Fine Scale	Early detection	1-km hexagon	1 week (epi weeks 18–38)	JMP, SAS	No (in progress)
K. Model-EAKF System	Early detection	Mosquito abatement district	Weekly	Matlab/R	Available upon request
L. Temperature-forced Model-EAKF System	Early detection	Mosquito abatement district	Weekly	Matlab/R	Available upon request
M. California Risk Assessment	Early detection	Flexible	Flexible	VectorSurv Gateway (website) ³	Available upon request

¹Spatial patterns: models with predictions that do not vary by year. Early warning: models that do not include current-year surveillance data, may include current-year climate/weather data, and have a model lead time on the order of days to months. Early detection: models that include current-year surveillance data, may include other data streams, and have a lead time on the order of days to months.

²The model itself is flexible with respect to temporal resolution. The GitHub implementation was designed for annual temporal resolution.

³The website is implemented in Javascript, PHP, SQL, Google Maps API, and Mapbox API.

<https://doi.org/10.1371/journal.pntd.0009653.t001>

Table 2. Model inputs.

Model	Human Data	Mosquito Surveillance	Climate/Weather	Land-cover	Sociological	Other
A. Historical Null	Y ¹	Y ¹	N	N	N	N
B. Spatial Risk Random Forest	Y	N	Y	N	N	N
C. Temperature-trait-based Relative R ₀ Model	N	N	Y	N	N	N
D. Spatial Risk High Resolution BRT	Y	N	Y	Y	N	Y
E. RF1	Y	Y	Y	Y	Y	Y
F. NE_WNV County-years	Y	N	Y	N	N	N
G. GLMER Ensemble	N	Y	Y	N	N	N
H. Harris County	N	Y	Y	Y	N	Y
I. ArboMAP	Y	Y	Y	N	N	N
J. Chicago Ultra-Fine Scale	Y	Y	Y	Y	Y	Y
K. Model-EAKF System	Y	Y	N	N	N	Y
L. Temperature-forced Model-EAKF System	Y	Y	Y	N	N	Y
M. California Risk Assessment	Y	Y	Y	Y	N	Y

¹For the Null model, only human data are required to predict human cases, and only mosquito surveillance data are required to predict mosquito infection rates. Mosquito surveillance is not used to predict human cases or vice versa in this model.

<https://doi.org/10.1371/journal.pntd.0009653.t002>

GLMER Ensemble are expected to run in under an hour for a county or smaller extent on a PC with a 2.9 GHz processor and 16 GB RAM.

In order to understand public decision-making processes and goals, we sent a request for information on what decisions are routinely made with respect to WNV (Table 5, form provided as S3 Text), through the Centers for Disease Control Regional Centers of Excellence. We received responses from 4 mosquito abatement districts (St. Tammany Parish Mosquito Abatement, LA; Northwest Mosquito Abatement District, IL; North Shore Mosquito Abatement District, IL; and the Harris County Mosquito and Vector Control Division, TX). We categorized the responses and then compared the suitability of each model with respect to temporal and spatial scale in relation to each decision identified by the public health professionals, until a consensus was reached.

Model classification framework

The models varied in spatial and temporal resolution. Based on the distribution of temporal and spatial resolutions in Fig 4, we propose a 3 × 3 description system for models based on their temporal and spatial scales (Table 6). We define 3 temporal scales: (1) short-term, corresponding to operational decisions made on the scale of weeks to months; (2) medium-term, corresponding to decisions made over months within a year related to planning and preparation; and (3) long-term for planning efforts made across multiple years. With respect to spatial scale, we identify fine-grain models with resolution of meters to kilometers (e.g., Fig 6C), medium-grain models with a resolution of a single management unit (e.g., mosquito abatement districts or county subdivisions; Fig 6B), and coarse-grain models that make predictions for multiple aggregated management units or single management units with large geographic coverage (e.g., county-level models; Fig 6A). Note that these scales apply to the resolution of the models, not the extent of the models. For example, weekly risk estimates with a 30 m × 30 m cell size would be a short-term fine-grain model, regardless of whether it was applied to a single neighborhood or an entire country. The aim of this description system is to better align model descriptions with the scales of application. For example, decisions requiring a short-term fine-grain model, such as where to apply an adulticide, would not be informed by a

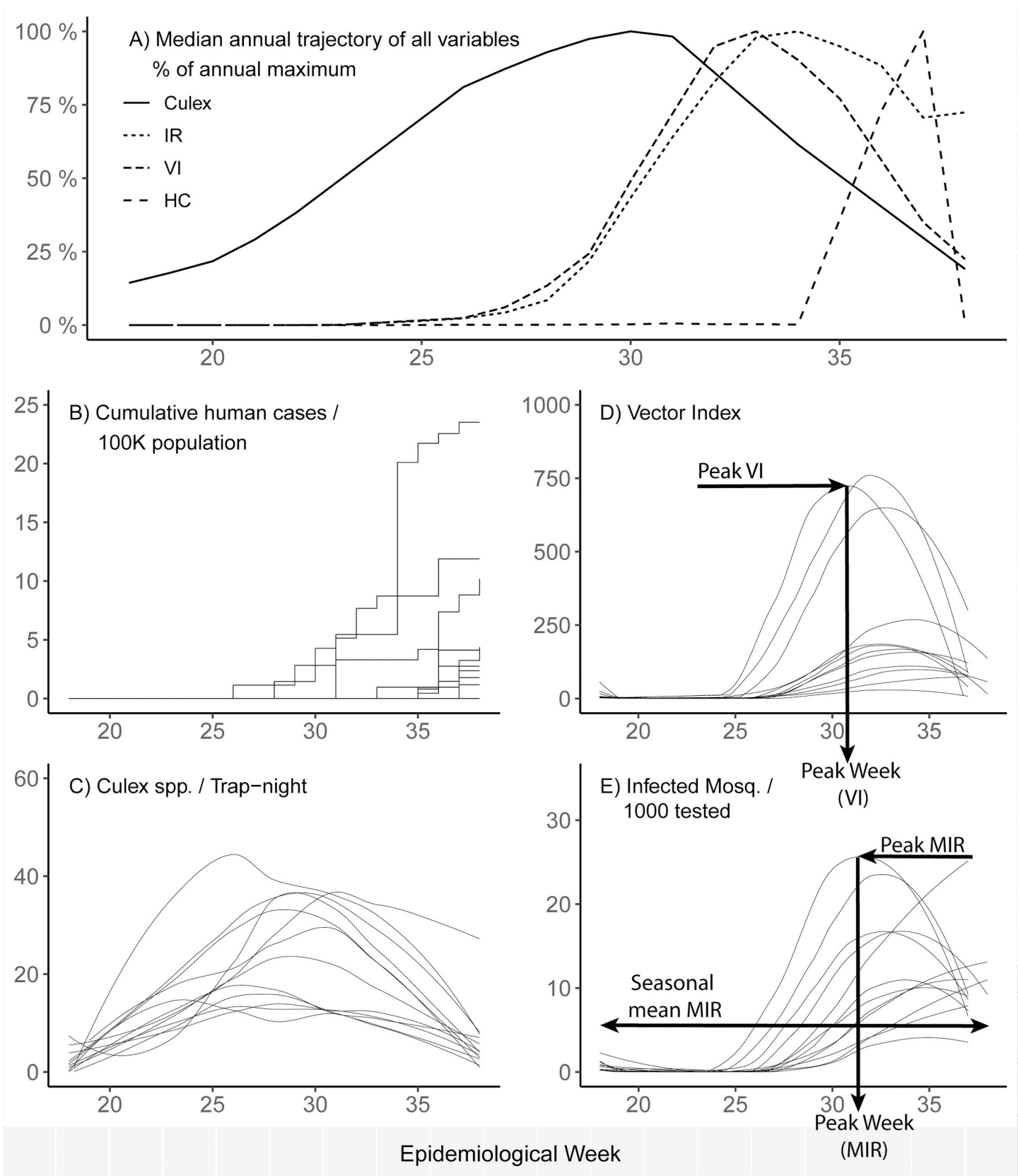


Fig 2. Examples of key model outputs. (A) A summary of key outputs for 1 year. (B) Cumulative human cases (annual human cases), (C) *Culex* mosquito abundance per trap night, (D) vector index (*Culex* abundance times infection rate by week), and (E) MIR per 1,000 mosquitoes. Peak MLE/IR is the mosquito infection rate in the peak week, Peak week for MLE/IR is the week in which the peak is reached, while Seasonal MLE/MIR is the infection rate over the season when the mosquitoes are active (using either MLEs or MIRs). *Culex*, *Culex* abundance; IR, mosquito infection rate, either as MIR or MLE; HC, human cases; MIR, minimum infection rate; MLE, maximum likelihood estimate of infection rate; VI, vector index.

<https://doi.org/10.1371/journal.pntd.0009653.g002>

medium-term coarse-grain planning model. We suggest that models could also be classified based on lead time and accuracy, sensitivity, and specificity, but these classifications may be region and/or scale dependent and, therefore, require a rigorous quantitative comparison to be developed.

Overview of models

The individual models are described in detail in [S4 Text](#). Here, we summarize the models with respect to the purpose for which they were developed, the statistical basis of each model, the models' use of surveillance data and climate data, and, finally, on their model selection approaches.

Table 3. Model output/predictions. Prediction targets included human case counts, mosquito infection rates as either MIRs or as MLEs. Probabilistic models are those that generate predictions as probability distributions rather than single mean values. The additional prediction targets column indicates whether the model generates additional outputs not otherwise included in the table.

Model	Annual Human Cases	Seasonal MLE/MIR	Peak MLE/MIR	Peak Week for MLE/MIR ¹	Vector Index (weekly)	Probabilistic?	Additional Prediction Targets
A. Historical Null	Y	Y	N	N	N	Y	N
B. Spatial Risk Random Forest	N	N	N	N	N	N	Y
C. Temperature-trait-based Relative R ₀ Model	N	N	N	N	N	Y	Y
D. Spatial Risk High Resolution BRT	N	N	N	N	N	Y	Y
E. RF1	Y	Y	N	N	N	Y ²	N
F. NE_WNV County-years	Y	N	N	N	N	Y	Y ³
G. GLMER Ensemble	N	Y	N	N	N	N	N
H. Harris County	N	Y	Y	Y	N	N ⁴	Y
I. ArboMAP	Y	Y	N	N	N	Y	Y
J. Chicago Ultra-Fine Scale	Y	N	N	N	N ⁵	Y	Y ⁶
K. Model-EAKF System	Y	Y	Y	Y	N ⁷	Y	Y ⁸
L. Temperature-forced Model-EAKF System	Y	Y	Y	Y	N ⁷	Y	Y ⁸
M. California Risk Assessment	N	N	N	N	N	N	Y ⁹

¹Peak week could also be calculated for human cases but typically is not done in practice; therefore, this output was omitted from the table.

²The model has been upgraded since the initial publication to support probabilistic outputs.

³Counties with cases.

⁴In principle, the model could produce probabilistic output.

⁵The model uses vector index as a predictor but does not predict values for vector index.

⁶Can theoretically inverse cases and MIR, but model not tested for that.

⁷The model can be parameterized with either MLE infection rates or vector index, but empirically, the results from the vector index parameterization were not as strong, and, therefore, the final model is based on MLE.

⁸+/-25% of peak week, human cases, total infections over the season; +/-25% or 1 human case.

⁹Virus transmission risk to humans.

BRT, Boosted Regression Trees; EAKF, Ensemble-adjustment Kalman Filter; MIR, minimum infection rate; MLE, maximum likelihood estimate of infection rate.

<https://doi.org/10.1371/journal.pntd.0009653.t003>

Table 4. Model applications. Only published model applications were included. Each line corresponds to a separate model test; therefore, some models appear more than once. References are listed for further details.

Model	Study	Prediction Target	Sample Size	Spatial Domain	Time Domain	Testing Method ¹	Metric Score ²
B. Spatial Risk Random Forest	[29]	Mean annual incidence per 100,000 population	43,512 county-years	Conterminous US (3,108 counties)	2005–2018, averaged	Bootstrapping	$R^2_{pred} = 0.59$ [0.44–0.70], $RMSE = 3.7$
D. Spatial Risk High Resolution BRT	[30]	Ranked relative risk (0–1)	1,378 human cases	South Dakota	2004–2017	Out of sample data	$AUC = 0.727$
E. RF1	[28]	Annual human cases	882 county-years	New York and Connecticut	2000–2015	LOYOCV	$R^2_{pred} = 0.72$, $RMSE = 1.6$
E. RF1	[28]	Seasonal mosquito MLE	218 county-years	New York and Connecticut	2000–2015	LOYOCV	$R^2_{pred} = 0.45$, $RMSE = 2.3$
E. RF1	[28]	Seasonal mosquito MLE	2,596 trap-years	New York and Connecticut by trap	2000–2015	LOYOCV	$R^2_{pred} = 0.53$, $RMSE = 1.0$
F. NE_WNV County-years	[34]	2018 human cases ³	1,472 county-years	Nebraska	2002–2017	Out of sample data	$CRPS = 1.90$
F. NE_WNV County-years	[34]	2018 WNV positive counties ³	1,472 county-years	Nebraska	2002–2017	Out of sample data	$Accuracy = 0.717$
G. GLMER Ensemble	[20]	MLE mosquito infection rate	225 grid-years	Suffolk County, New York	2001–2015	LOYOCV	$RMSE = 4.27$
H. Harris County	[21]	MLE mosquito infection rate (1-month lead)	130,567 trap-nights	Harris County, Texas	2002–2016	Out of sample data	$R^2_{pred} = 0.8$
H. Harris County	[21]	Mosquito abundance (1-month lead)	10,533,033 mosquitoes	Harris County, Texas	2002–2016	Out of sample data	$R^2_{pred} = 0.2$
I. ArboMAP	[35]	Positive county-weeks	Approximately 9,504 county-weeks (training) Approximately 792 county-weeks (testing)	South Dakota	2004–2015 (training) 2016 (testing)	Out of sample data	$AUC = 0.836–0.856^4$
I. ArboMAP	[36]	Positive county-weeks	Approximately 11,088 county-weeks	South Dakota	2004–2017 ⁵	Fit to training data only	$AUC = 0.876$, $R_s = 0.84$
J. Chicago Ultra-Fine Scale	[22–24]	Human case probability (by hexagon)	1,346,940 hexagon-weeks	Variable, up to 5,345 1-km hexagons	2005–2016 ⁶	Fit to training data only	$R^2 > 0.85$; $RMSE < 0.02$; $AUC > 0.90$
K. Model-EAKF System	[25]	Annual human cases; peak mosquito infection rates; peak timing of infectious mosquitoes; annual infectious mosquitoes	21 county-years	2 counties (Suffolk, New York and Cook, Illinois)	Weekly, Varied by location	Retrospective data assimilation	<i>Threshold-based accuracy</i> ⁷
K. Model-EAKF System	[26]	Multiple ⁸	110 outbreak-years	12 counties	Weekly, Varied by location	Retrospective data assimilation	<i>Threshold-based accuracy</i> ⁷
K. Model-EAKF System	[39]	Multiple ⁸	4 county-years	4 counties	Weekly, 2017	Real-time data assimilation	<i>Threshold-based accuracy</i> ⁷
L. Temperature-forced Model-EAKF System	[26]	Multiple ⁸	110 outbreak-years	12 counties	Weekly, Varied by location	Retrospective data assimilation	<i>Threshold-based accuracy</i> ⁷
M. California Risk Assessment	[31]	Historical outbreaks of western equine encephalomyelitis and St. Louis encephalitis as proxy for WNV	14 agency-years	California	Half-months	Temporal correspondence	Early detection of arbovirus risk prior to outbreaks
M. California Risk Assessment	[32]	Onset and peak of human cases by geographic region	12 half-months in 3 regions	California	Half-months	Retrospective data assimilation	Early detection of WNV risk prior to onset and peak of human cases

(Continued)

Table 4. (Continued)

Model	Study	Prediction Target	Sample Size	Spatial Domain	Time Domain	Testing Method ¹	Metric Score ²
M. California Risk Assessment	[33]	Emergency planning threshold (risk \geq 2.6)	11,476 trap-nights	Los Angeles Country, California	2004–2010	Retrospective data assimilation	AUC = 0.982

¹LOYOCV: leave-one-year-out cross-validation; Out of sample data: accuracy based on data not used to develop the model; Fit to training data only: accuracy based on the same data used to develop the model; Retrospective data assimilation: finalized data until the time of forecast; Real-time data assimilation: data processed and available at the time of forecast.

² R^2_{pred} : predictive R^2 , i.e., an R^2 calculated on data outside the sample, R_s : Spearman correlation coefficient, AUC: area under the curve, *Threshold-based accuracy*: \pm 25% of peak week, human cases, total infections over the season; \pm 25% or 1 human case, RMSE: Root Mean Squared Error, CRPS: Continuous Ranked Probability Score.

³Results for 2018 reported here, validation was also performed separately for 2012–2017, see [34] for details.

⁴Three analyses presented: short-term: AUC = 0.856, annual made on July 5: AUC = 0.836, annual made on July 39: AUC = 0.855.

⁵Restricted to July–September for each year.

⁶Restricted to 21 epi weeks per year.

⁷Varied by analysis and lead time.

⁸Prediction targets: human cases in next 3 weeks; annual human cases; week with highest percentage of infectious mosquitoes; peak mosquito infection rate; annual infectious mosquitoes.

AUC, area under the curve; CRPS, Continuous Ranked Probability Score; LOYOCV: leave-one-year-out cross-validation; RMSE, Root Mean Squared Error; WNV, West Nile virus.

<https://doi.org/10.1371/journal.pntd.0009653.t004>

Model purposes. The models included in this review were developed for a variety of purposes, including generating present-day patterns of spatial risk, predicting risk under future climate change, and providing medium- to short-term planning guidance (see Table 6). The Spatial Risk High Resolution BRT model [30], Spatial Risk Random Forest model [29], and Temperature-trait-based Relative R_0 models [27] were all developed for spatial risk, with the latter two with an aim to also provide information about climate change risk. The RF1 model can also be used to make climate change risk predictions [87].

Medium-term risk guidance models include the RF1 model and the NE_WNV County-years model [34] for human cases at the county-annual scale (and mosquito-based risk for RF1), and the GLMER Ensemble [28] and the Harris County models aimed at mosquito-based infection rates (and abundance for the Harris County model) at the monthly scale. ArboMAP (Arbovirus Modeling and Prediction) [35,36], the Model-EAKF Systems [25,26], the California Risk Assessment [31–33], and the Chicago Ultra-Fine Scale (UFS) models [23,24] all provide short-term risk information. Model-EAKF System and the Chicago UFS models model human cases directly. ArboMAP focuses on the probability of a county having at least 1 human case in a given week, while the California Risk Assessment model provides an index of relative risk without a quantitative prediction of numbers of human cases. ArboMAP was specifically designed to facilitate WNV forecasting by epidemiologists working in state public health offices, while the California Risk Assessment model is currently used by the state of California to guide vector control operations [41]. The Model-EAKF Systems provide a data assimilation approach, which uses data from the current season to update the model predictions as the season progresses to make weekly predictions in areas with high levels of mosquito surveillance.

Statistical basis. Broadly, the models fall into 3 general approaches: machine learning techniques, traditional statistical approaches, and mathematical models. For the machine learning approaches, Hess and colleagues used Boosted Regression Trees [42], while the Spatial Risk Random Forest Model and the RF1 Models used Random Forest methods [43,44]. The

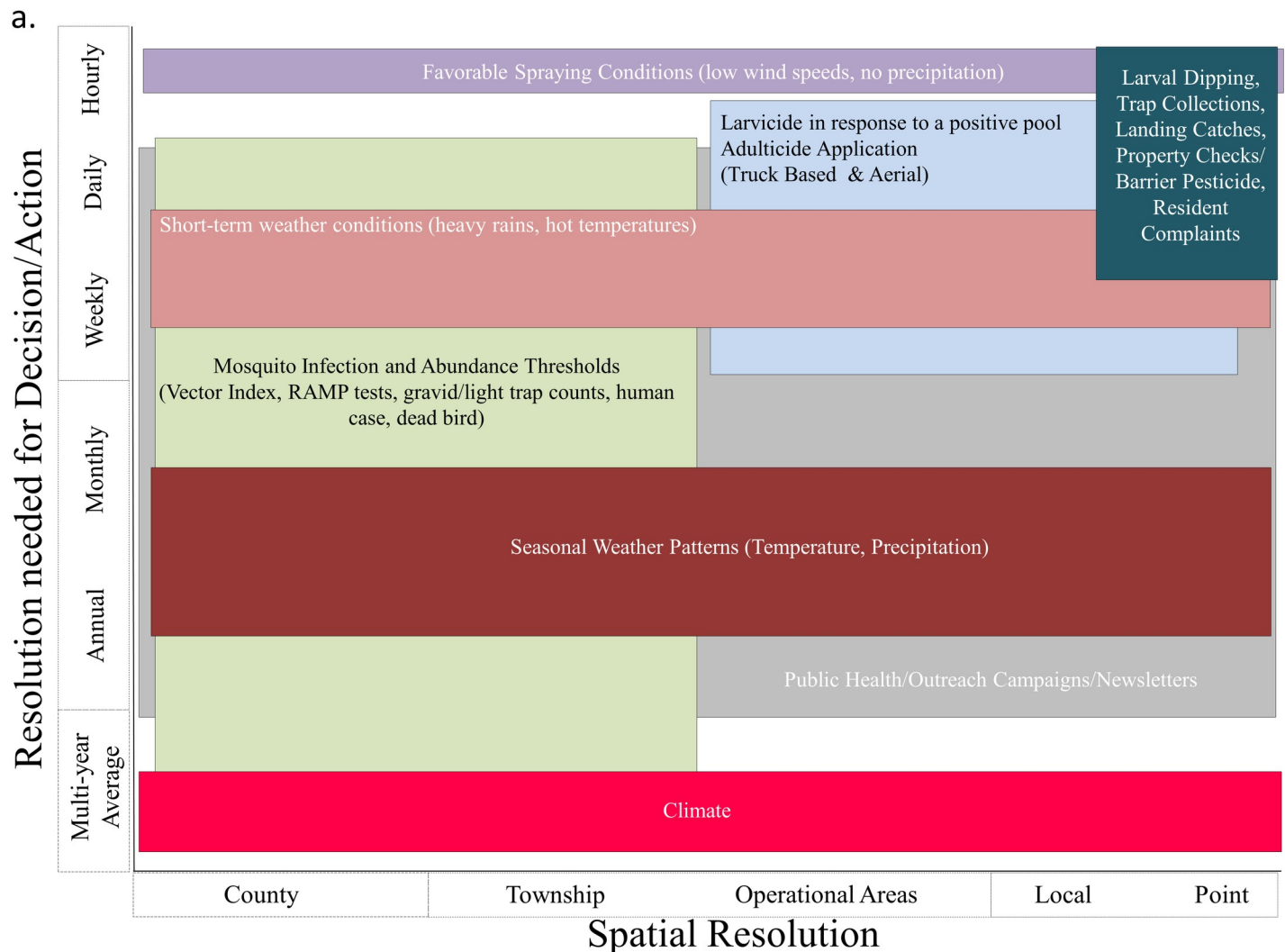


Fig 3. Generalized overview of major factors, tools, and decisions utilized by mosquito control agencies. This figure is based on 4 representative mosquito abatement districts: 2 in Chicago (IL), Slidell (LA), and Houston (TX). Management practices may differ from program to program, but similar challenges and decisions are made from across varying spatial (local to district-wide) and temporal (days to multiple months) scales.

<https://doi.org/10.1371/journal.pntd.0009653.g003>

RF1 Model was modified to produce probabilistic output using quantile random forests [45,46].

Traditional statistical approaches include the GLMER Ensemble, using negative binomial mixed-effects models (GLMER Ensemble). The Harris County, TX Model is a seasonally autoregressive forced model [21], i.e., a linear model that capture nonsymmetric features in the seasonality of the underlying data. The NE_WNV County-years model [34] used a general additive model with thin-plate splines (the R package mgcv) [47] for nonparametric modeling of distributed lags (lag lengths of 12, 18, 24, 30, and 36 months) of drought and temperature data, using restricted maximum likelihood estimation with a log link and negative binomial distribution. The ArboMAP model also used a distributed lags approach. ArboMAP used logistic regression models with environmental indices (temperature, precipitation, humidity, etc.) included as distributed lags, with shapes governed by splines [35,36]. The Chicago UFS model is also based on logistic regression, with 1 km-wide hexagonal (spatial) and 1-week (temporal) resolutions using environmental, land-use/land-cover, historical weather, light

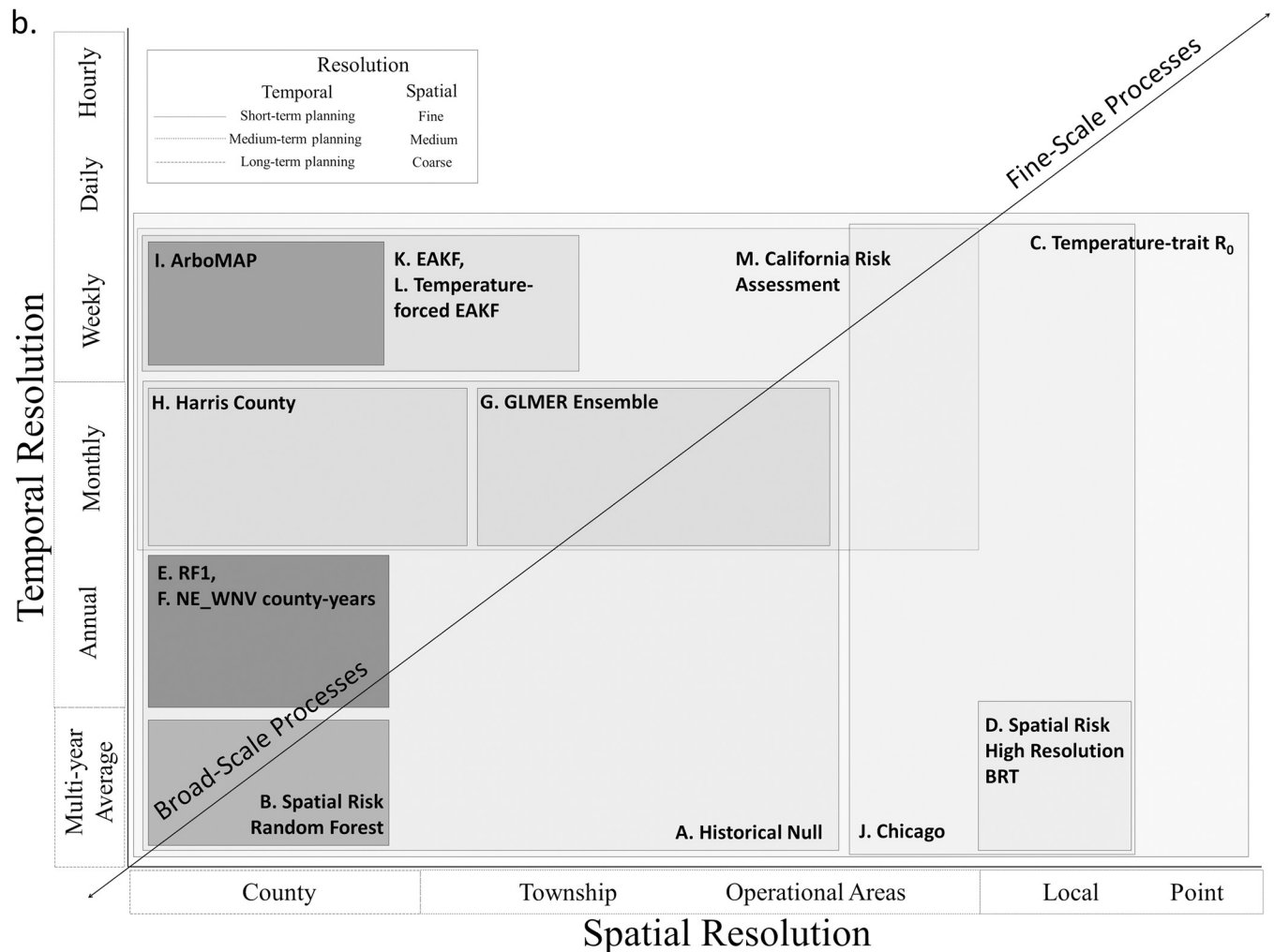


Fig 4. The 13 models reviewed in this paper arranged by spatial and temporal resolution. Rectangles with decreasing shades of gray indicate less coverage identifying potential knowledge gaps. These gaps may guide future model development or require additional data collection, as many models are at the county-annual scale due to data availability.

<https://doi.org/10.1371/journal.pntd.0009653.g004>

pollution, human socioeconomic and demographic, mosquito abundance and infection, mosquito landing rates on humans, and human activity/exposure risk as covariates.

For the mathematical models, the Temperature-trait-based Relative R_0 model used a modified Ross–McDonald equation that incorporates nonlinear thermal response curves fit to laboratory mosquito and virus trait data. The Model-EAKF System [25] and Temperature-forced Model-EAKF System [26] used a standard susceptible–infected–recovered epidemiological construct and were optimized using a data assimilation method (the ensemble adjustment Kalman filter (EAKF) [48] and 2 observed data streams: mosquito infection rates and reported human WNV cases. The models differ in that the Temperature-forced Model-EAKF System accounted for temperature modulation of the extrinsic incubation period for mosquitos [26]. The California Risk Assessment model estimates an overall level of WNV risk based on the average of all available risk elements (1) average daily temperature; (2) relative abundance of adult *Culex* mosquitoes versus the historical average; (3) WNV infection prevalence in *Culex* mosquitoes; (4) sentinel chicken seroconversions; (5) WNV infections in dead birds; and (6) human cases. Because human cases are affected by reporting lags and thus are unreliable

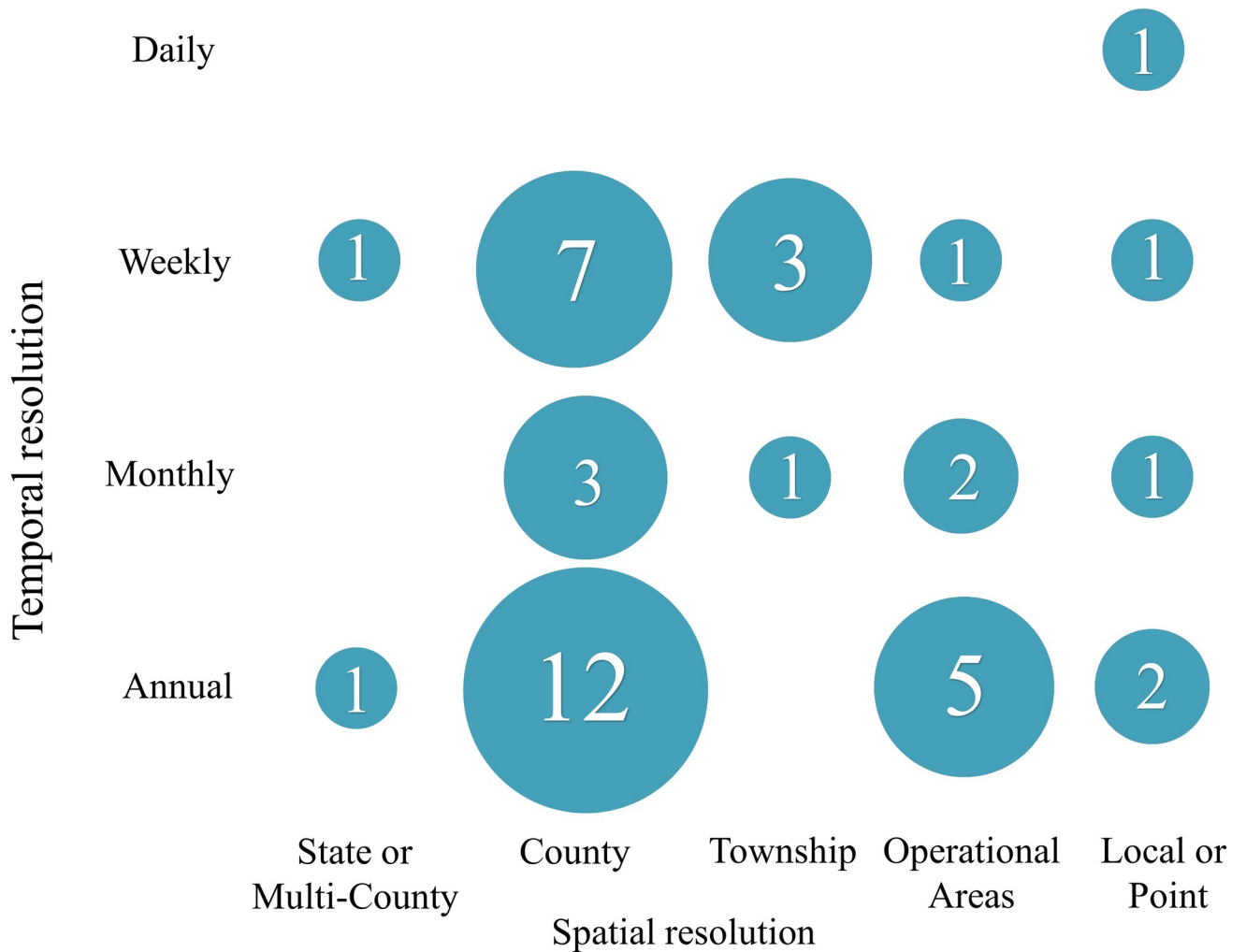


Fig 5. A summary of the spatial and temporal resolution for the 41 models reviewed in [17] that are not included in Fig 4. Numbers indicate the number of models at that spatial and temporal scale.

<https://doi.org/10.1371/journal.pntd.0009653.g005>

Table 5. List of common decisions made regarding a public health and vector control response to WNV. Letters correspond to models in Tables 1–4 and indicate models with an appropriate spatial or temporal resolution to inform the decision. Note that this pertains to the scale on which predictions are made and provides no information on the accuracy of the model predictions. As such, models with appropriate scale, but insufficient accuracy, would not be useful in an operational context.

Public health decisions	Potentially applicable models	
	When (timing)	Where (area)
Mosquito and WNV surveillance (trap sites)	C, M	C, D, M
Mosquito and WNV surveillance (county/district thresholds)	C, M	A–J, M
Public health and outreach	C, E–M	A–J, M
Larviciding	C, H–M	C, J, M
Truck-based adulticiding	C, I–M	C, J, M
Aerial adulticiding	C, I–M	C, J, M

WNV, West Nile virus.

<https://doi.org/10.1371/journal.pntd.0009653.t005>

Table 6. Classification of temporal and spatial resolutions relevant to vector control and public health decision-making.

Classification Term	Spatial or Temporal	Resolution
Long-term planning	Temporal	Years to decades
Medium-term planning	Temporal	Months to year
Short-term planning	Temporal	Days to weeks
Coarse grain	Spatial	Multiple/large management districts (e.g., county or above)
Medium grain	Spatial	Single management district or county subdivision
Fine grain	Spatial	Meters to km, within a management district

<https://doi.org/10.1371/journal.pntd.0009653.t006>

indicators of real-time risk, they are typically omitted from risk calculations that guide mosquito control operations during the season. Each surveillance element is assigned a value on an ordinal scale (1 to 5 for lowest to highest risk), and the mean value of all factors is calculated to estimate the WNV transmission risk and corresponding response level (i.e., normal season (1.0 to 2.5), emergency planning (2.6 to 4.0), and epidemic (4.1 to 5.0)).

Use of surveillance data. Models varied in their use of mosquito and human surveillance data. Models using mosquito data calculated infection rates but differed in the approaches used to do so. The RF1 model used a published R method [49] applied at the county level, pooled for 3 *Culex* species: *Culex pipiens*, *Culex restuans*, and *Culex salinarius*. The Harris County Model used the maximum likelihood method by Farrington [50]. In the GLMER Ensemble, the *Culex* spp. infection rate was calculated for each NLDAS grid cell and year using maximum likelihood approaches. Similarly, the 2 Model-EAKF Systems estimate mosquito infection rates by week but require at least 1 positive mosquito and at least 300 mosquito samples per week. In ArboMAP, mosquito data are modeled in their own mixed-effects models, in which exponential growth curves are imposed on mosquito infection rates in the early season. The estimated growth rate is then used as a covariate in the human models. The Chicago UFS model used MIRs in conjunction with abundance to estimate vector index. For the California Risk Assessment model, mosquito abundance is compared to the 5-year average for the same area and time period. Viral infection rates are expressed as either MIRs [51] or MLEs [52] per 1,000 female mosquitoes tested. Due to differences in the attractiveness of traps to different subsets of the population, abundance and infection prevalence data are not pooled across trap types, but the most sensitive trap type's value is used in the risk assessment. Also, due to differences in the sensitivity of traps between species and spatial heterogeneity in the distribution of *Culex tarsalis* and *Cx. pipiens* complex mosquitoes relative to humans, separate risk calculations for each species are suggested. The Temperature-trait-based Relative R_0 model does not incorporate surveillance data as currently structured.

The 2 Model-EAKF Systems include human case data from the current year, unlike virtually all the other models (the California Risk Assessment allows it to be incorporated, but this is generally not done in practice). Historical human case data are used by several of the models, including the RF1, ArboMAP, the Chicago UFS, and the NE_WNV County-years models. In addition to training on past estimates of human cases, the NE_WNV County-years model included the rate of cumulative incidence of human cases as the total number of previous cases, for each county and each year, per 100,000 population on the basis that previous exposure to WNV reduces human infection rates [53]. The Chicago UFS model included human cases following a zero-inflated Poisson distribution.

Climate data inputs. While many models used climate data, models were constructed with different climate data sources. The Spatial Risk Random Forest model was based on 4 km gridded data from the Precipitation elevation Regressions on Independent Slopes Model

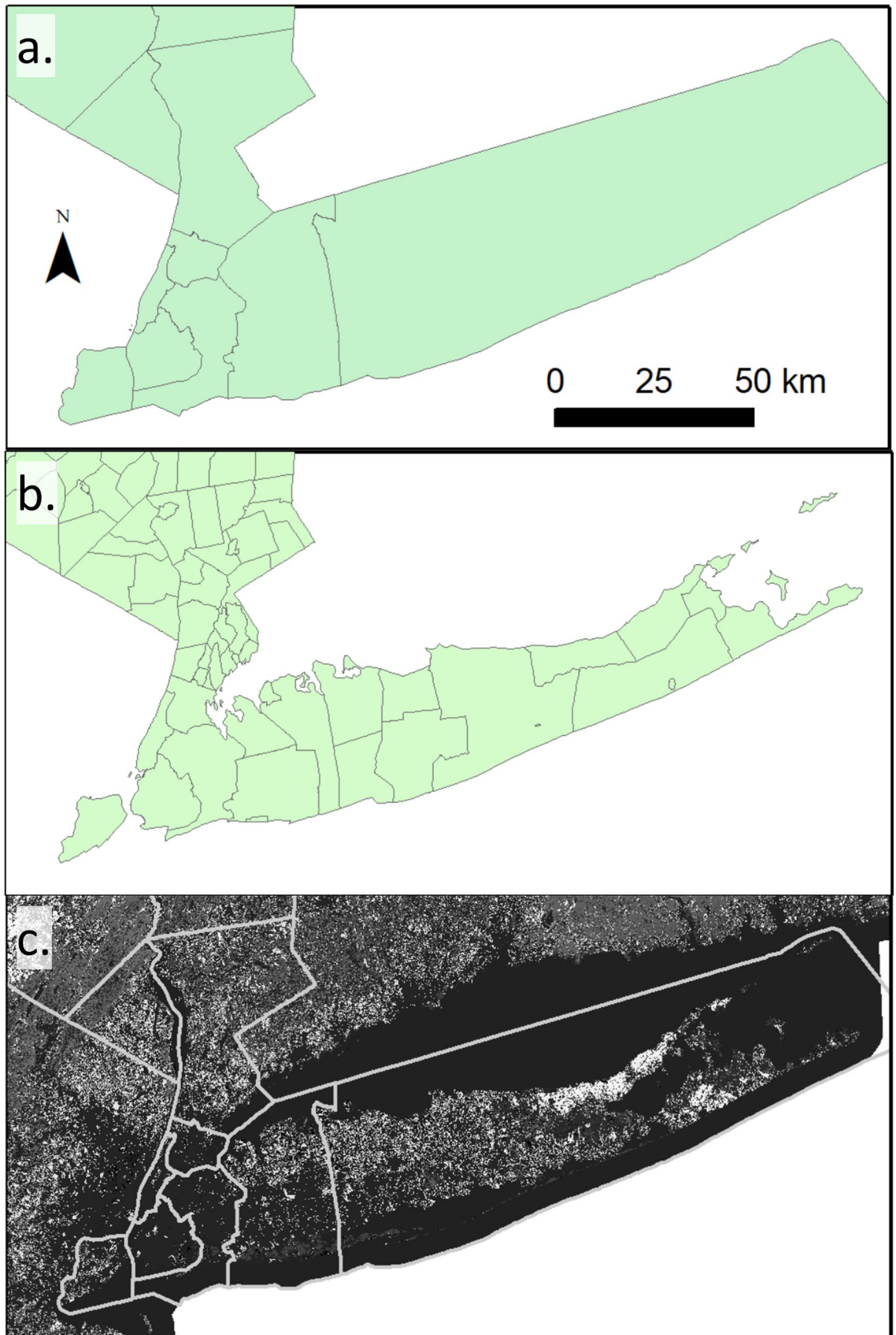


Fig 6. Examples of the 3 spatial scales described in Table 6 for Long Island, NY. (a) Coarse-grain: county, (b) medium-grain: county subdivision, and (c) fine-grain: 30 × 30 m resolution for vegetation types [40], with the NY county outlines in gray for context. County outlines and county subdivisions from the 2017 US Census <https://www.census.gov/geo/maps-data/data/tiger-line.html>.

<https://doi.org/10.1371/journal.pntd.0009653.g006>

(PRISM) [54,55]. Time series of weather data used in ArboMAP are typically obtained from gridMET [56] through Google Earth Engine (GEE) [57] using a custom downloader script [58]. gridMET combines data from both PRISM [55] and NLDAS-2 [59] into a single high-resolution gridded data set. The GLMER Ensemble model used approximately 13 km² gridded monthly averages in temperature, precipitation, specific humidity, and soil moisture [59] from the North American Land Data Assimilation System (NLDAS) Mosaic submodel data set. The RF1 model used soil moisture data from the NLDAS Noah submodel data set, and initially with a published ensemble of temperature and precipitation data [60], and later with data from gridMET as above. The NE_WNV County-years model used lags of drought (1-month Standardized Precipitation Index (SPEI); 1-month Standardized Precipitation and Evapotranspiration Index) [61] and temperature variables (standardized temperature deviations from the mean, standardized precipitation deviations from the mean, from NOAA's Climate Divisional Database) [62]. The Harris County model used the mean, standard deviation, and kurtosis for temperature and rainfall from local weather stations. While the Temperature-trait-based Relative R₀ Model requires a temperature input, the model is flexible with respect to the choice of temperature input. The Temperature-forced Model-EAKF System used a mean climatology based on NLDAS-2 data from 1981 to 2000 and each outbreak year for each region of interest.

Model selection approaches. The GLMER Ensemble approach, the Harris County Model, and the NE_WNV County-years model all used the Akaike information criterion (AIC) for model selection. In the Harris County and NE_WNV County-years models, the model that minimized the AIC score was selected. In the GLMER Ensemble, all combinations of predictor variables were considered. Those models for which all explanatory variables were significant with 95% confidence were ranked by AIC [63,64]. The Akaike weight was calculated, and the set of models whose Akaike weights sum to 0.95 were used for the inference. The RF1 model used a two-stage fitting process for the Random Forest, removing all variables below a calculated mean importance score, and then removing additional variables that did not increase the model's explanatory power using a variance partitioning approach [28]. The Spatial Risk Random Forest model did not use variable selection and included all predictor variables in the inference. The Model-EAKF System optimizes a 300-member ensemble of model simulations and, in so doing, provide an improved, posterior estimate of the true state as well as estimates of unobserved state variables and parameters. Model-EAKF System forecasts were repeated 10 times with different randomly selected initial conditions and evaluated for accuracy according to prescribed forecast metrics.

Discussion

We qualitatively compared the models in the context of 6 potential decisions related to public health and vector control response to WNV (Table 5). This qualitative comparison was made based on the scale of the model and the scale of the decision. We found that some models, such as those developed at the coarser spatial and temporal scales (i.e., county/annual), are not useful for many of the decisions needed for vector control operations. Indeed, only 3 out of 13 models (the Temperature-trait-based Relative R₀ model, the California Risk Assessment, and the Chicago UFS model), would be potentially capable to guide spatial and temporal adulticiding based on model resolution. However, the Temperature-trait-based Relative R₀ model had

not been developed or validated for this purpose. The California Risk Assessment model provides a threshold-based risk but does not quantitatively predict the number of human cases in the present year.

The models reviewed here included 3 different classes of models. Nearly all of the models have been implemented in the statistical software R [65]. Most of the models reviewed used human data as an input with 4 exceptions (the Temperature-trait-based Relative R_0 model, the GLMER Ensemble, the Harris County model, and the California Risk Assessment as applied for real-time decisions by most vector control agencies). Climate was also a common input for models, having been used in all models except the Model-EAKF System (but note there is a temperature-forced version) and the null models. This is unsurprising, given the importance of climatic conditions on the mosquito life cycle (e.g., [66]). Landcover, sociological inputs, and other inputs were less common, even though these inputs may also be important in understanding disease dynamics (e.g., [67,68]). Model outputs were more heterogeneous, making comparisons across models more challenging. Most models focused on either annual cases or seasonal mosquito infection rates, or both. Fewer models examined patterns within a season, notably the ArboMAP model, the Chicago UFS model, the Harris County model, the 2 Model-EAKF System models, and the California Risk Assessment. Model applications are difficult to compare qualitatively and will be best examined through quantitative comparisons on common data sets, as R^2 values and RMSE values can be difficult to interpret across scales and in the context of different numbers of cases or infection rates. Most of the models have been applied with nonoverlapping domains, making direct comparisons more difficult (Fig 1). Some of the models are specific examples of more general approaches that can be applied at finer scales.

Many of the models were designed to give an indication of whether it will be a “good year” or a “bad year” for WNV, without a direct, specific connection to decisions related to WNV control (e.g., Models E, F, G, I; Table 1). Temperature and precipitation may play a large role in such determinations [69], and such relationships may be the complex result of several interacting traits [27]. None of the models directly address the question of initial vector or viral surveillance, although such surveillance could be guided by spatial risk, in which case models A to H could be used to help guide general regions. The California Risk Assessment (M) provides recommendations for enhanced surveillance as risk levels increase. The Model EAKF Systems (I and J) specify specific quantitative surveillance requirements in order to be implemented but do not address the question of where such surveillance should take place.

Broadly, many of these models were developed within a local context. As a consequence, they are not necessarily the “best” model, but one sufficient to the task. In addition, being highly local means the models may be difficult to generalize to new locations. Regional variation is expected in the underlying processes. In some cases, the models are very closely tied to a specific region (e.g., the Chicago UFS model and the GLMER Ensemble) and influenced by variability in surveillance programs (spatiotemporal resolution). Regions vary in the dominant mosquito vector(s), the degree to which they are rural or urban, human risk-taking behavior (e.g., time spent outdoors, presence of window screens, and presence of mosquito breeding habitat), mosquito surveillance, and human socioeconomic status and ability to report biting mosquitoes.

Our analyses demonstrated that there is no “one size fits all” model—different models may be needed to guide the vector control and public health decisions considered here. Some decisions are made early in the season, while others are made later in the season. Decisions during the season may be constrained by planning made prior to the mosquito season. The decisions also vary on the spatial scale at which they take place, with public outreach taking place potentially across an entire state, while truck-based or aerial insecticide applications take place on

localized scales of up to several square kilometers. It is important to note that these decisions are informed by WNV risk but are also influenced by social factors (e.g., interest in vector-borne disease around outdoor activities [70] or specific holidays such as Memorial Day or Fourth of July), financial (e.g., budget constraints) [70], environmental (e.g., current weather), and regulatory/political factors (e.g., protected ecosystems and restrictions on adulticide applications and willingness to pay for control) [70–73]. Vector control operations in the US are highly localized, and substantial regional variation exists in the timing of decisions, the thresholds used for decisions, and the willingness to apply adulticides. For example, Harris County, TX is primarily managed by a single agency (Harris County Public Health Mosquito and Vector Control Division, although several small cities and municipalities will also control nuisance mosquitoes) covering a geographic area of approximately 4,600 km². Cook and DuPage Counties, Illinois have 4 mosquito abatement districts covering a geographic area of almost 3,500 km².

WNV case incidence rates also vary from region to region. A change in the associated causal factors would likely influence the particular model's performance. Regions also differ in their surveillance efforts, surveillance methods, and trap density. These factors affect the quality of the data going into the models and the quality of the data being used to evaluate the model. For example, a model could do a very good job of predicting the “true value,” but with poor data, the model may be scored worse than a model that predicted an observed prevalence that was consistent with biases created through data collection or sampling. Regions also vary in their turnaround times for data [39], and this may influence the degree to which different models can be implemented. Thus, more than one model may be necessary, and models may need to accommodate additional location-specific constraints.

In addition, the workshop discussions highlighted the importance of quantifying the value of information associated with model results. In many mosquito abatement districts, larvicide and public outreach are routine actions and are unlikely to be strongly affected by variations in the predicted risk of WNV. Other districts may dynamically increase public health outreach or larvicide application during “bad” years. In contrast, decisions regarding adulticide application are usually made based on perceived risk at a given point in time. However, these decisions typically take place on scales below the spatial resolution of most models. In practice, adulticide applications may be reactive to positive detection of WNV in mosquitoes, birds, or humans or a specified metric such as MIR or the vector index. While some of the models used vector index as an input, none of them included vector index as a predicted output (Table 3), despite common use of this metric by vector control. Increased temporal and spatial resolution in model results and a focus on these quantities could make models more applicable to these decisions. As it stands, most models are parameterized on the county scale or larger, and this prevents them from being utilized for decisions at local scales. In part, this is driven by the availability of data—human case data in particular are difficult to obtain at scales finer than the county because of privacy concerns.

Models may also be limited by accuracy thresholds needed for decisions. Often, while models may provide more information than a null model, these models may not provide enough confidence to be used for decision-making. On the other hand, relying on null models for early warning of upcoming high-risk events is not feasible either. Trade-offs between confidence in imperfect model predictions of extreme events and uninformative priors (e.g., null models) have to be made. In case of rare events, the uncertainty is usually higher than for more regular events. Some of this may be related to inadequate data—for rare events, especially large data sets may be needed for model training. Models may also be limited by heterogeneity in underlying processes. Models are typically aggregated over multiple trap sites, with the assumption that similar processes are operating at all trap sites. If traps differ strongly from

one another in the underlying mosquito or disease dynamics, this heterogeneity may be averaged over during the modeling process and lead the model to produce mean predictions that are incorrect for all locations. Improved models for identifying regions of homogenous risk could aid in this aggregation process.

Incorporating the effects of public health interventions such as vector control efforts into models may be difficult as well. Interventions, even when applied at discrete locations, typically have effects that extend beyond the place and time of treatment that are not easily quantified. Also, most interventions do not have suitable controls, as interventions are required to protect public health. Therefore, finding a control site with no intervention that is equivalent to the treatment area is difficult (as any sites with equal risk would likely be treated). Before–after controls are challenging, as mosquito populations can be dynamic. For example, even if populations do not decrease after an intervention, it is unclear whether the intervention did not work, or whether the mosquito populations would be much higher in the absence of the intervention.

Future directions

To improve the quality of modeling for decision-making, a clear mapping between model outputs and information needed for decisions would be beneficial. Quantifying the gain of information achieved by the model and the value of that information gain would provide clear guidance for when to apply a model to guide decision-making. Metrics, such as human cases averted, or resources saved due to an early intervention, could strengthen the justification for decisions made on the basis of a model. These metrics would need to be carefully described, however, as it may be difficult to know exact numbers of cases averted, and, therefore, language should reflect uncertainty in the results.

Formal quantitative comparisons of existing models may be useful to ensure that all decision-makers are able to select the best models for their regions and the decisions they need to make. Multiple models exist at the county-annual scale (Figs 4 and 5), and comparisons could be performed for 4 of the 6 decisions (Table 5). Formal model comparisons have been performed for dengue [74] and leishmaniasis [75]. A comparison of linear and classification and regression tree analysis methods has been performed locally for WNV [76]. Models for spatial risk, models for “good” versus “bad” years, and models that guide local decisions such as application of adulticides could all be compared separately. Quantitative comparisons would provide a degree of rigor and could also contribute to assessing the gain of information associated with each model. A formal quantitative comparison should consider the lead times associated with each model in the context of the lead times needed for control efforts [77]. In addition, the creation of standard data sets for each key output would aid in model comparison. Standard data sets are used in machine learning (e.g., the Anderson’s Iris data) [78] and provide a basis for comparing different methods. This will be a challenging task given the complexity and regional variation of the disease system.

There is a critical need for more social science research, particularly the need to incorporate human behavior related to vector control and exposure risk in the models. Mosquito transmission takes place within social-ecological systems [79]. Integrated mosquito management (IMM) [80] in part aims to influence human behavior and the interaction of humans and mosquitoes. Predictive computational models of human behaviors [81] are potentially powerful tools to support IMM interventions (e.g., source reduction, public education, and community involvement). This is particularly true if the models can link the individual attributes and behaviors with the dynamics of the socioenvironmental systems within which individuals/residents operate [81]. Such considerations as well as making available actionable, customizable

(e.g., predictive analysis), and easy-to-use model outputs can also encourage mosquito control practitioners to use such model outputs for local level decision-making [71]. In addition, models should be culturally responsive to the needs of state, tribal, local, or territorial public health and mosquito control agencies.

Finally, models need to be based on sound scientific data. A recent study identified over 1,000 mosquito control agencies in the continental US. Of these, 152 agencies had publicly available open access mosquito data sets, while 148 agencies had live data that can be leveraged and used with good effect [82]. Indeed, improved integration of IMM interventions such as public health campaigns, larvicide applications [83,84], and adulticide applications [85,86] into the models will be critical to assessing the role of interventions in a modeling framework.

Supporting information

S1 Text. Model description template.

(DOCX)

S2 Text. Model description instructions.

(DOCX)

S3 Text. Regional decision-makers template.

(DOCX)

S4 Text. Detailed model descriptions.

(DOCX)

Acknowledgments

We thank K. Caillouet, D. Sass, C. Lippi, S. Ryan, E. Mordecai, K. Price, L. Kramer, M. Childs, N. Nova, P. Stiles, S. Mundis, and A. Ciota for participation in the WNV Model Comparison Project.

References

1. Kramer LD, Li J, Shi P-Y. West Nile virus. *Lancet Neurol*. 2007; 6:171–81. [https://doi.org/10.1016/S1474-4422\(07\)70030-3](https://doi.org/10.1016/S1474-4422(07)70030-3) PMID: 17239804
2. Farajollahi A, Fonseca DM, Kramer LD, Kilpatrick AM. “Bird biting” mosquitoes and human disease: a review of the role of *Culex pipiens* complex mosquitoes in epidemiology. *Infect Genet Evol*. 2011; 11:1577–85. <https://doi.org/10.1016/j.meegid.2011.08.013> PMID: 21875691
3. CDC. Final Cumulative Maps & Data for 1999–2019. *Ctr Dis Control Prev*; 2020. Available from: <https://www.cdc.gov/westnile/statsmaps/cumMapsData.html#one>.
4. Limaye VS, Max W, Constible J, Knowlton K. Estimating the Health-Related Costs of 10 Climate-Sensitive U.S. Events During 2012. *GeoHealth*. 2019; 3:245–265. <https://doi.org/10.1029/2019GH000202> PMID: 32159045
5. Mongoh MN, Hearne R, Dyer N, Khaita M. The economic impact of West Nile virus infection in horses in the North Dakota equine industry in 2002. *Trop Anim Health Prod*. 2008; 40:69–76.
6. CDC. Species of dead birds in which West Nile virus has been detected, United States, 1999–2016. 2017 [cited 2018 May 21]. Available from: <https://www.cdc.gov/westnile/resources/pdfs/BirdSpecies1999-2016.pdf>.
7. George TL, Harrigan RJ, LaManna JA, DeSante DF, Saracco JF, Smith TB. Persistent impacts of West Nile virus on North American bird populations. *Proc Natl Acad Sci U S A*. 2015; 112:14290. <https://doi.org/10.1073/pnas.1507747112> PMID: 26578774
8. LaDeau SL, Kilpatrick AM, Marra PP. West Nile virus emergence and large-scale declines of North American bird populations. *Nature*. 2007; 447:710. <https://doi.org/10.1038/nature05829> PMID: 17507930

9. Kilpatrick AM, Wheeler SS. Impact of West Nile virus on bird populations: limited lasting effects, evidence for recovery, and gaps in our understanding of impacts on ecosystems. *J Med Entomol* 2019; 56:1491–7. <https://doi.org/10.1093/jme/tjz149> PMID: 31549723
10. Stauffer GE, Miller DAW, Williams LM, Brown J. Ruffed grouse population declines after introduction of West Nile virus. *J Wildl Manag*. 2018; 82:165–72. <https://doi.org/10.1002/jwmg.21347>
11. Crosbie SP, Koenig WD, Reisen WK, Kramer VL, Marcus L, Carney R, et al. Early impact of West Nile virus on the Yellow-billed Magpie (*Pica nuttalli*). *Auk*. 2008; 125:542–50.
12. Mostashari F, Bunning ML, Kitsutani PT, Singer DA, Nash D, Cooper MJ, et al. Epidemic West Nile encephalitis, New York, 1999: results of a household-based seroepidemiological survey. *Lancet*. 2001; 358:261–4. [https://doi.org/10.1016/S0140-6736\(01\)05480-0](https://doi.org/10.1016/S0140-6736(01)05480-0) PMID: 11498211
13. Busch MP, Wright DJ, Custer B, Tobler LH, Stramer SL, Kleinman SH, et al. West Nile virus infections projected from blood donor screening data, United States, 2003. *Emerg Infect Dis*. 2006; 12:395. <https://doi.org/10.3201/eid1205.051287> PMID: 16704775
14. CDC. Nationally notifiable arboviral diseases reported to ArboNET: Data release guidelines. *Ctr Dis Control Prev*; 2019.
15. Chiang CL, Reeves WC. Statistical estimation of virus infection rates in mosquito vector populations. *Am J Hyg*. 1962; 75:377–91. <https://doi.org/10.1093/oxfordjournals.aje.a120259> PMID: 13878878
16. Walter SD, Hildreth SW, Beaty BJ. Estimation of infection rates in populations of organisms using pools of variable size. *Am J Epidemiol*. 1980; 112:124–8. <https://doi.org/10.1093/oxfordjournals.aje.a112961> PMID: 7395846
17. Barker CM. Models and Surveillance Systems to Detect and Predict West Nile Virus Outbreaks. *J Med Entomol*. 2019 [cited 2019 Sep 27]. <https://doi.org/10.1093/jme/tjz150> PMID: 31549727
18. Reiner RC Jr, Perkins TA, Barker CM, Niu T, Fernando Chaves L, Ellis AM, et al. A systematic review of mathematical models of mosquito-borne pathogen transmission: 1970–2010. *J R Soc Interface*. 2013; 10. <https://doi.org/10.1098/rsif.2012.0921> PMID: 23407571
19. Moise IK, Riegel C, Muturi EJ. Environmental and social-demographic predictors of the southern house mosquito *Culex quinquefasciatus* in New Orleans, Louisiana. *Parasit Vectors*. 2018; 11:249. <https://doi.org/10.1186/s13071-018-2833-5> PMID: 29665862
20. Little E, Campbell SR, Shaman J. Development and validation of a climate-based ensemble prediction model for West Nile Virus infection rates in *Culex* mosquitoes, Suffolk County, New York. *Parasit Vectors*. 2016; 9:443. <https://doi.org/10.1186/s13071-016-1720-1> PMID: 27507279
21. Poh KC, Chaves LF, Reyna-Nava M, Roberts CM, Fredregill C, Bueno R Jr, et al. The influence of weather and weather variability on mosquito abundance and infection with West Nile virus in Harris County, Texas USA. *Sci Total Environ*. 2019; 675:260–72. <https://doi.org/10.1016/j.scitotenv.2019.04.109> PMID: 31030133
22. Uelmen JA, Irwin P, Bartlett D, Brown W, Karki S, Ruiz MO, et al. Effects of Scale on Modeling West Nile Virus Disease Risk. *Am J Trop Med Hyg*. 2020. <https://doi.org/10.4269/ajtmh.20-0416> PMID: 33146116
23. Uelmen JA, Irwin P, Brown W, Karki S, Ruiz MO, Li B, et al. Dynamics of data availability in disease modeling: An example evaluating the trade-offs of ultra-fine-scale factors applied to human West Nile virus disease models in the Chicago area, USA. *PLoS ONE*. 2021. <https://doi.org/10.1371/journal.pone.0251517> PMID: 34010306
24. Karki S, Brown WM, Uelmen J, Ruiz MO, Smith RL. The drivers of West Nile virus human illness in the Chicago, Illinois, USA area: Fine scale dynamic effects of weather, mosquito infection, social, and biological conditions. *PLoS ONE*. 2020; 15:e0227160. <https://doi.org/10.1371/journal.pone.0227160> PMID: 32437363
25. DeFelice NB, Little E, Campbell SR, Shaman J. Ensemble forecast of human West Nile virus cases and mosquito infection rates. *Nat Commun*. 2017; 8:14592. <https://doi.org/10.1038/ncomms14592> PMID: 28233783
26. DeFelice NB, Schneider ZD, Little E, Barker C, Caillouet KA, Campbell SR, et al. Use of temperature to improve West Nile virus forecasts. *PLoS Comput Biol*. 2018; 14. <https://doi.org/10.1371/journal.pcbi.1006047> PMID: 29522514
27. Shocket MS, Verwillow AB, Numazu MG, Slamani H, Cohen JM, El Moustaid F, et al. Transmission of West Nile and five other temperate mosquito-borne viruses peaks at temperatures between 23°C and 26°C. Franco E, Malagón T, Gehman A, editors. *Elife*. 2020; 9:e58511. <https://doi.org/10.7554/eLife.58511> PMID: 32930091
28. Keyel AC, Elison Timm O, Backenson PB, Prussing C, Quinones S, McDonough KA, et al. Seasonal temperatures and hydrological conditions improve the prediction of West Nile virus infection rates in

- Culex mosquitoes and human case counts in New York and Connecticut. *PLoS ONE*. 2019; 14: e0217854. <https://doi.org/10.1371/journal.pone.0217854> PMID: 31158250
29. Gorris M.E., 2019. Ch 4: Climate controls on the spatial pattern of West Nile virus incidence in the United States. In *Environmental infectious disease dynamics in relation to climate and climate change*. University of California, Irvine. Available from: <https://www.proquest.com/openview/4879510a3806015bd7fc27384b18dd18/1/advanced>
 30. Hess A, Davis JK, Wimberly MC. Identifying Environmental Risk Factors and Mapping the Distribution of West Nile Virus in an Endemic Region of North America. *GeoHealth*. 2018; 2:395–409. <https://doi.org/10.1029/2018GH000161> PMID: 32159009
 31. Barker C, Reisen W, Kramer V. California state mosquito-borne virus surveillance and response plan: A retrospective evaluation using conditional simulations. *Am J Trop Med Hyg*. 2003; 68:508–18. <https://doi.org/10.4269/ajtmh.2003.68.508> PMID: 12812335
 32. Barker CM, Park BK, Melton FS, Eldridge BF, Kramer VL, Reisen WK. 2007 Year-in-Review: integration of NASA's Meteorological Data into the California Response Plan. *Proceedings and Papers of the Seventy-Sixth Annual Conference of the Mosquito and Vector Control Association of California*. 2008; 7–12.
 33. Kwan JL, Park BK, Carpenter TE, Ngo V, Civen R, Reisen WK. Comparison of Enzootic Risk Measures for Predicting West Nile Disease, Los Angeles, California, USA, 2004–2010. *Emerg Infect Dis*. 2012; 18:1298–306. <https://doi.org/10.3201/eid1808.111558> PMID: 22840314
 34. Smith KH, Tyre AJ, Hamik J, Hayes MJ, Zhou Y, Dai L. Using Climate to Explain and Predict West Nile Virus Risk in Nebraska. *GeoHealth*. 2020; 4:e2020GH000244. <https://doi.org/10.1029/2020GH000244> PMID: 32885112
 35. Davis JK, Vincent G, Hildreth MB, Kightlinger L, Carlson C, Wimberly MC. Integrating Environmental monitoring and mosquito surveillance to predict vector-borne disease: prospective forecasts of a West Nile virus outbreak. *PLoS Currents*. 2017; 9. <https://doi.org/10.1371/currents.outbreaks.90e80717c4e67e1a830f17feeaaf85de> PMID: 28736681
 36. Davis JK, Vincent GP, Hildreth MB, Kightlinger L, Carlson C, Wimberly MC. Improving the prediction of arbovirus outbreaks: A comparison of climate-driven models for West Nile virus in an endemic region of the United States. *Acta Trop*. 2018; 185:242–50. <https://doi.org/10.1016/j.actatropica.2018.04.028> PMID: 29727611
 37. Keyel AC, Kilpatrick AM. Probabilistic Evaluation of Null Models for West Nile Virus in the United States. *BioRxiv*. 2021. <https://doi.org/10.1101/2021.07.26.453866>
 38. Shumway R, Stoffer D. *Time series analysis and its applications*. 3rd ed. New York: Springer; 2011.
 39. DeFelice NB, Birger R, DeFelice N, Gagner A, Campbell SR, Romano C, et al. Modeling and Surveillance of Reporting Delays of Mosquitoes and Humans Infected With West Nile Virus and Associations With Accuracy of West Nile Virus Forecasts. *JAMA Netw Open*. 2019; 2:e193175–5. <https://doi.org/10.1001/jamanetworkopen.2019.3175> PMID: 31026036
 40. LANDFIRE. LANDFIRE Remap 2014 Existing Vegetation Type (EVT) CONUS. LANDFIRE, Earth Resources Observation and Science Center (EROS), US Geological Survey; 2017. Available from: https://www.landfire.gov/lf_schedule.php.
 41. CDPH. California Department of Public Health, Mosquito and Vector Control Association of California, and University of California, California mosquito-borne virus surveillance and response plan; 2020. Available from: https://westnile.ca.gov/download.php?download_id=4502.
 42. Elith J, Leathwick JR, Hastie T. A working guide to boosted regression trees. *J Anim Ecol*. 2008; 77:802–13. <https://doi.org/10.1111/j.1365-2656.2008.01390.x> PMID: 18397250
 43. Breiman L. Random forests. *Mach Learn*. 2001; 45:5–32.
 44. Liaw A, Wiener M. Classification and Regression by randomForest. *R News*. 2002; 2:18–22.
 45. Meinshausen N. Quantile regression forests. *J Mach Learn Res*. 2006; 7:983–99.
 46. Meinshausen N. *quantregForest: Quantile Regression Forests*; 2017. Available from: <https://CRAN.R-project.org/package=quantregForest>.
 47. Wood SN. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *J R Stat Soc B*. 2011; 73:3–36.
 48. Anderson JL. An ensemble adjustment Kalman filter for data assimilation. *Mon Weather Rev*. 2001; 129:2884–903.
 49. Williams CJ, Moffitt CM. Estimation of pathogen prevalence in pooled samples using maximum likelihood methods and open-source software. *J Aquat Anim Health*. 2005; 17:386–91.
 50. Farrington C. Estimating prevalence by group testing using generalized linear models. *Stat Med*. 1992; 11:1591–7. <https://doi.org/10.1002/sim.4780111206> PMID: 1359621

51. Reeves WC, Hammon WM. The role of arthropod vectors. *Epidemiology of the arthropod-borne viral encephalitides in Kern County, California, 1943–1952*. Berkeley, CA: University of California Press; 1962. p. 75–190.
52. Hepworth G, Biggerstaff BJ. Bias correction in estimating proportions by pooled testing. *J Agric Biol Environ Stat*. 2017; 22:602–14. <https://doi.org/10.1007/s13253-017-0297-2> PMID: 30636859
53. Paull SH, Horton DE, Ashfaq M, Rastogi D, Kramer LD, Diffenbaugh NS, et al. Drought and immunity determine the intensity of West Nile virus epidemics and climate change impacts. *Proc R Soc B*. 2017; 284:20162078. <https://doi.org/10.1098/rspb.2016.2078> PMID: 28179512
54. Daly C, Halbleib M, Smith JI, Gibson WP, Doggett MK, Taylor GH, et al. Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *Int J Climatol*. 2008; 28:2031–64.
55. PRISM Climate Group. Parameter-elevation Regression on Independent Slopes Model. Oregon State University. 2019 [cited 2019 Aug 28]. Available from: <http://prism.oregonstate.edu>.
56. Abatzoglou JT. Development of gridded surface meteorological data for ecological applications and modelling. *Int J Climatol*. 2013; 33:121–31.
57. Gorelick N, Hancher M, Dixon M, Ilyushchenko S, Thau D, Moore R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens Environ*. 2017. <https://doi.org/10.1016/j.rse.2017.06.031>
58. Wimberly MC, Davis JK. GRIDMET_downloader.js. University of Oklahoma; 2019. Available from: <https://github.com/ecograph/arbomap>.
59. Mitchell KE, Lohmann D, Houser PR, Wood EF, Schaake JC, Robock A, et al. The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple GCIP products and partners in a continental distributed hydrological modeling system. *J Geophys Res Atmos*. 2004; 109:1–32.
60. Newman AJ, Clark MP, Craig J, Nijssen B, Wood A, Gutmann E, et al. Gridded ensemble precipitation and temperature estimates for the contiguous United States. *J Hydrometeorol*. 2015; 16:2481–500.
61. Abatzoglou JT, McEvoy DJ, Redmond KT. The West Wide Drought Tracker: drought monitoring at fine spatial scales. *Bull Am Meteorol Soc*. 2017; 98:1815–20.
62. Vose RS, Applequist S, Squires M, Durre I, Menne MJ, Williams CN, et al. NOAA's Climate Divisional Database (nCLIMDIV). National Climatic Data Center; 2014. Available from: <https://doi.org/10.7289/V5M32STR>.
63. Akaike H. A new look at the statistical model identification. *IEEE Trans Autom Control*. 1974; 19:716–23. <https://doi.org/10.1109/TAC.1974.1100705>
64. Burnham KP, Anderson DR. *Model selection and Multimodel inference*. New York: Springer-Verlag; 2002.
65. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing; 2017. Available from: <https://www.R-project.org/>.
66. Ciota AT, Matakchiero AC, Kilpatrick AM, Kramer LD. The effect of temperature on life history traits of *Culex* mosquitoes. *J Med Entomol*. 2014; 51:55–62. <https://doi.org/10.1603/me13003> PMID: 24605453
67. Ruiz MO, Tedesco C, McTighe TJ, Austin C, Kitron U. Environmental and social determinants of human risk during a West Nile virus outbreak in the greater Chicago area, 2002. *Int J Health Geogr*. 2004; 3:8. <https://doi.org/10.1186/1476-072X-3-8> PMID: 15099399
68. Rochlin I, Turbow D, Gomez F, Ninivaggi DV, Campbell SR. Predictive Mapping of Human Risk for West Nile Virus (WNV) Based on Environmental and Socioeconomic Factors. *PLoS ONE*. 2011; 6:e23280. <https://doi.org/10.1371/journal.pone.0023280> PMID: 21853103
69. Ciota AT, Keyel AC. The Role of Temperature in Transmission of Zoonotic Arboviruses. *Viruses*. 2019; 11:1013. <https://doi.org/10.3390/v11111013> PMID: 31683823
70. Moise IK, Zulu LC, Fuller DO, Beier JC. Persistent Barriers to Implementing Efficacious Mosquito Control Activities in the Continental United States: Insights from Vector Control Experts. *Current Topics in Neglected Tropical Diseases*. IntechOpen; 2018.
71. Tedesco C, Ruiz M, McLafferty S. Mosquito politics: local vector control policies and the spread of West Nile Virus in the Chicago region. *Health Place*. 2010; 16:1188–95. <https://doi.org/10.1016/j.healthplace.2010.08.003> PMID: 20709590
72. Dickinson K, Paskewitz S. Willingness to pay for mosquito control: how important is West Nile virus risk compared to the nuisance of mosquitoes? *Vector Borne Zoonotic Dis*. 2012; 12:886–92. <https://doi.org/10.1089/vbz.2011.0810> PMID: 22651384
73. Dickinson KL, Hayden MH, Haenchen S, Monaghan AJ, Walker KR, Ernst KC. Willingness to pay for mosquito control in key west, Florida and Tucson, Arizona. *Am J Trop Med Hyg*. 2016; 94:775–9. <https://doi.org/10.4269/ajtmh.15-0666> PMID: 26903603

74. Johansson MA, Apfeldorf KM, Dobson S, Devita J, Buczak AL, Baugher B, et al. An open challenge to advance probabilistic forecasting for dengue epidemics. *Proc Natl Acad Sci*. 2019; 116:24268–74. <https://doi.org/10.1073/pnas.1909865116> PMID: 31712420
75. Chaves LF, Pascual M. Comparing Models for Early Warning Systems of Neglected Tropical Diseases. *PLoS Negl Trop Dis*. 2007; 1:1–6. <https://doi.org/10.1371/journal.pntd.0000033> PMID: 17989780
76. Ruiz MO, Chaves LF, Hamer GL, Sun T, Brown WM, Walker ED, et al. Local impact of temperature and precipitation on West Nile virus infection in *Culex* species mosquitoes in northeast Illinois, USA. *Parasit Vectors*. 2010; 3. <https://doi.org/10.1186/1756-3305-3-19> PMID: 20302617
77. Hii YL, Rocklöv J, Wall S, Ng LC, Tang CS, Ng N. Optimal Lead Time for Dengue Forecast. *PLoS Negl Trop Dis*. 2012; 6:e1848. <https://doi.org/10.1371/journal.pntd.0001848> PMID: 23110242
78. Dua D, Graff C. UCI Machine Learning Repository. University of California, Irvine, School of Information and Computer Sciences; 2017. Available from: <http://archive.ics.uci.edu/ml>
79. LaDeau SL, Leisnham PT, Biehler D, Bodner D. Higher mosquito production in low-income neighborhoods of Baltimore and Washington, DC: understanding ecological drivers and mosquito-borne disease risk in temperate cities. *Int J Environ Res Public Health*. 2013; 10:1505–26. <https://doi.org/10.3390/ijerph10041505> PMID: 23583963
80. CDC. Integrated Mosquito Management. *Ctr Dis Control Prev*. 2020. Available from: <https://www.cdc.gov/mosquitoes/mosquito-control/professionals/integrated-mosquito-management.html>. <https://doi.org/10.1093/jme/tjaa078> PMID: 32474595
81. Ajelli M, Moise IK, Hutchings TCS, Brown SC, Kumar N, Johnson NF, et al. Host outdoor exposure variability affects the transmission and spread of Zika virus: Insights for epidemic control. *PLoS Negl Trop Dis*. 2017; 11:e0005851. <https://doi.org/10.1371/journal.pntd.0005851> PMID: 28910292
82. Rund SSC, Moise IK, Beier JC, Martinez ME. Rescuing Troves of Hidden Ecological Data to Tackle Emerging Mosquito-Borne Diseases. *J Am Mosq Control Assoc*. 2019; 35:75–83. <https://doi.org/10.2987/18-6781.1> PMID: 31442186
83. Anderson JF, Ferrandino FJ, Dingman DW, Main AJ, Andreadis TG, Becnel JJ. Control of mosquitoes in catch basins in Connecticut with *Bacillus thuringiensis israelensis*, *Bacillus sphaericus*, and spinosad. *J Am Mosq Control Assoc*. 2011; 27:45–55. <https://doi.org/10.2987/10-6079.1> PMID: 21476447
84. Harbison JE, Nasci R, Runde A, Henry M, Binnall J, Hulsebosch B, et al. Standardized operational evaluations of catch basin larvicides from seven mosquito control programs in the Midwestern United States During 2017. *J Am Mosq Control Assoc*. 2018; 34:107–16. <https://doi.org/10.2987/18-6732.1> PMID: 31442163
85. Reddy MR, Spielman A, Lepore TJ, Henley D, Kiszewski AE, Reiter P. Efficacy of resmethrin aerosols applied from the road for suppressing *Culex* vectors of West Nile virus. *Vector Borne Zoonotic Dis*. 2006; 6:117–27. <https://doi.org/10.1089/vbz.2006.6.117> PMID: 16796509
86. Lothrop H, Lothrop B, Palmer M, Wheeler S, Gutierrez A, Miller P, et al. Evaluation of pyrethrin aerial ultra-low volume applications for adult *Culex tarsalis* control in the desert environments of the Coachella Valley, Riverside County, California. *J Am Mosq Control Assoc*. 2007; 23:405–19. <https://doi.org/10.2987/5623.1> PMID: 18240517
87. Keyel, A.C., Raghavendra, A., Ciota, A.T. and Alison Timm, O. (2021), West Nile virus is predicted to be more geographically widespread in New York State and Connecticut under future climate change. *Global Change Biology Accepted Author Manuscript*. <https://doi.org/10.1111/gcb.15842>