

2020

Optimizing Arbovirus Surveillance using Risk Mapping and Coverage Modelling

Joni A. Downs

University of South Florida, downs@usf.edu

Mehrdad Vaziri

University of South Florida, mehrdadv@mail.usf.edu

George Deskins

Citrus County Mosquito Control

William Kellner

Citrus County Mosquito Control

Kristi M. Miley

University of South Florida, kmiley@mail.usf.edu

See next page for additional authors

Follow this and additional works at: https://digitalcommons.usf.edu/geo_facpub



Part of the [Earth Sciences Commons](#)

Scholar Commons Citation

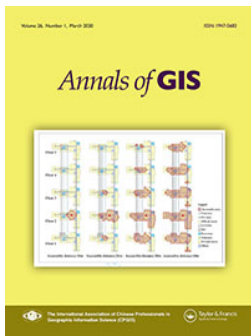
Downs, Joni A.; Vaziri, Mehrdad; Deskins, George; Kellner, William; Miley, Kristi M.; and Unnasch, Thomas R., "Optimizing Arbovirus Surveillance using Risk Mapping and Coverage Modelling" (2020). *School of Geosciences Faculty and Staff Publications*. 2240.

https://digitalcommons.usf.edu/geo_facpub/2240

This Article is brought to you for free and open access by the School of Geosciences at Digital Commons @ University of South Florida. It has been accepted for inclusion in School of Geosciences Faculty and Staff Publications by an authorized administrator of Digital Commons @ University of South Florida. For more information, please contact digitalcommons@usf.edu.

Authors

Joni A. Downs, Mehrdad Vaziri, George Deskins, William Kellner, Kristi M. Miley, and Thomas R. Unnasch



Optimizing arbovirus surveillance using risk mapping and coverage modelling

Joni Downs, Mehrdad Vaziri, George Deskins, William Kellner, Kristi Miley & Thomas R. Unnasch

To cite this article: Joni Downs, Mehrdad Vaziri, George Deskins, William Kellner, Kristi Miley & Thomas R. Unnasch (2020) Optimizing arbovirus surveillance using risk mapping and coverage modelling, *Annals of GIS*, 26:1, 13-23, DOI: [10.1080/19475683.2019.1688391](https://doi.org/10.1080/19475683.2019.1688391)

To link to this article: <https://doi.org/10.1080/19475683.2019.1688391>



© 2019 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group, on behalf of Nanjing Normal University.



Published online: 18 Nov 2019.



Submit your article to this journal [↗](#)



Article views: 540



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 2 View citing articles [↗](#)

Optimizing arbovirus surveillance using risk mapping and coverage modelling

Joni Downs^a, Mehrdad Vaziri^a, George Deskins^b, William Kellner^b, Kristi Miley^c and Thomas R. Unnasch^c

^aSchool of Geosciences, University of South Florida, Tampa, FL, USA; ^bCitrus County Mosquito Control, Lecanto, FL, USA; ^cGlobal Health Infectious Disease Research, University of South Florida, Tampa, FL, USA

ABSTRACT

Diseases carried by mosquitoes and other arthropods endanger human health globally. Though costly, surveillance efforts are vital for disease control and prevention. This paper describes an approach for strategically configuring targeted disease surveillance sites across a study area. The methodology combines risk index mapping and spatial optimization modelling. The risk index is used to identify demand for surveillance, and the maximum covering location problem is used to select a specified number of candidate surveillance sites that covers the maximum amount of risk. The approach is demonstrated using a case study where optimal locations for sentinel surveillance sites are selected for the purposes of detecting eastern equine encephalitis virus in a county in the state of Florida. Optimal sentinel sites were selected under a number of scenarios that modelled different target populations (horses or humans), coverage distances (0.5, 1.0, and 1.5 km), and numbers of sites to select (1–12). Sentinel site selections for the horse and human models displayed different spatial patterns, with horse sites located largely in the west-central region and human ones in the north-central. Minor amounts of spatial overlap between the horse and human sites were observed, especially as coverage distances and numbers of sites were increased. Additionally, a near linear increase in risk coverage was observed as sites were incrementally added to the scenarios. This finding suggests that the number of sentinel sites within the ranges explored should be based on the maximum that can be funded, since they provide similar levels of benefit.

ARTICLE HISTORY

Received 28 August 2018
Accepted 28 October 2019

KEYWORDS

Eastern equine encephalitis virus; index model; maximum covering location model; spatial optimization

Introduction

Arboviruses – viruses transmitted by mosquitoes, flies, and arthropods – endanger human health globally (Gubler 2002). Arboviral diseases, such as malaria, dengue, river blindness, and yellow fever, are responsible for illness and death in millions of people. Even the less common of these viruses, such as West Nile virus, Zika virus, and Eastern Equine Encephalitis virus, can pose serious health threats locally. Accordingly, risk control of arboviral diseases is an important global health priority (Dowdle 1998).

Measures aimed to control the transmission of arboviruses to humans and other species, including livestock and wildlife, typically involve a number of strategies (Eldridge 1987). Vector management efforts are implemented to reduce viral transmission. These vector control efforts commonly include pesticide application, biological control, and engineering solutions aimed at reducing vector abundance in order to reduce virus transmission risk (Kean et al. 2015). Prevention efforts are also implemented in order to reduce disease incidence or impacts. For example, vaccination is widely used to prevent illnesses caused by arboviruses in animals (Hotez 2009).

For these vector control and disease prevention efforts to be effective, adequate surveillance is critical to providing early warnings of potential outbreaks (Olliaro et al. 2018). Surveillance is used to estimate vector abundance, detect viral activity in vectors, and document cases of infections in humans or other hosts. Popular methods of surveillance include the use of traps in combination with viral culture or molecular assays to detect the presence of the viruses of interest, the use of serological assays to detect exposure in sentinel animals, and standard epidemiological reporting techniques to collect data on the incidence of disease caused by viral infections. Although they can be effective, these approaches are often labour intensive in both the field and laboratory, as well as financially expensive to implement. Accordingly, selecting which locations to survey is critical when resources limit the number of sites that can be sampled.

Sampling design is an important consideration for disease surveillance. Gu et al. (2008) describes two basic spatio-temporal strategies for surveillance: extensive and targeted. Extensive sampling involves widespread surveillance of vectors or diseases over space and time. Stratified random sampling designs are often

recommended with this approach, which is best suited for estimating vector abundance and surveying for easily detectable diseases (Sedda et al. 2019). Such approaches are commonly used to monitor mosquito abundance and common viruses across broad spatial areas over time (Burkett-Cadena et al. 2016). Targeted surveillance, which involves intensive surveys when and where there is a high likelihood of detection, is recommended when infection rates are low and possibly spatially isolated (Chevalier, Lecollinet, and Durand 2011). Use of sentinel animals are commonly used for targeted surveillance of rare, arboviral diseases. Generally, the number of sentinel sites that can be operated is limited to a small number, mostly due to labour and maintenance costs (Ritchie et al. 2007; Hall et al. 2012). This makes sentinel placement an important strategic decision for disease surveillance. However, in practice they are often selected based on historical case reports, suspected transmission areas based on expert opinion, convenience of the location, or some combination of these factors which can limit their effectiveness (Ramírez et al. 2018). However, the effectiveness and efficiency of sentinel programmes might be improved with a more strategic site selection method.

This paper presents a methodology for strategically selecting targeted arbovirus surveillance sites under constraints of limited resources. The approach is demonstrated using a case study where optimal locations for sentinel surveillance sites are selected for the purposes of detecting eastern equine encephalitis (EEE) in a county in the state of Florida. EEE is caused by infection of the mosquito-borne eastern equine encephalitis virus (EEEV). EEEV is an arbovirus native to the eastern coast of the United States, with the largest number of cases reported in Florida. EEEV is maintained in a complex transmission cycle involving birds, reptiles, and mammals, including humans (Figure 1). EEEV is thought to overwinter in small mammals, wetland birds, and reptiles, where it is maintained in an enzootic cycle through transmission to songbirds by the mosquito *Culiseta melanura*. The epizootic cycle is largely carried out by bridge vectors, mosquitoes of the genera *Aedes* and *Coquillettidia*, which also feed on songbirds but also transmit the virus to dead end hosts such as humans,

horses, and poultry (CDC 2017). Florida has historically exhibited relatively high levels of EEEV activity in horses, humans, and sentinel chickens (USDA 2017). It is also the only state in the USA where EEEV transmission occurs year round. Furthermore, recent studies suggest that Florida serves as a reservoir from which EEEV is periodically introduced to the rest of the eastern USA (Armstrong et al. 2008), so it is an important region to monitor EEEV activity.

This approach for selecting optimal sentinel surveillance sites initially relies upon risk index models to map EEEV risk based on the composition and configuration of habitats associated with viral transmission. The premise is that the goal of targeted surveillance is to place the sentinel sites in the locations most likely to support transmission in order to maximize disease detection. So, the goal is to survey as much high risk area as possible, as measured by a validated risk index model, given a limited number of sentinel sites that can be operated. First, risk is mapped for both horses and humans, two relevant dead-end hosts within the study area. Second, spatial modelling is used to identify optimal sentinel sites within the risk maps. Specifically, the maximum covering location problem is applied to select a set number of sentinel sites in order to cover the maximum amount of risk within a specified coverage radius. The model is solved for a number of scenarios for each host species based on different numbers of selected sites and coverage distances. A discussion of the relative merits of each scenario is also provided.

Materials and methods

Study area

The focus of this study is Citrus County, one of 67 administrative counties in the state of Florida, USA. Citrus County has a population of about 140,000 people (2010 US Census) and 1,150 horses and ponies (2012 USDA Census of Agriculture). Although the county is not known as a hotspot for EEEV activity, a number of equine, emu, and chicken cases have been reported during recent years (CDC 2017). Additionally, as the

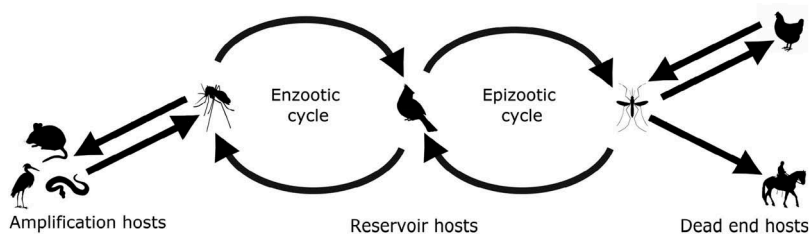


Figure 1. EEEV transmission cycle.

county borders multiple high incidence counties, it is an area of concern for future EEEV outbreaks.

Risk index mapping

Index models are a popular approach in GIScience for quantitatively assessing suitability of locations in a map for some purpose, such as wildlife habitat, economic development, or agricultural production (Brooks 1997; Downs, Gates, and Murray 2008). Typically, suitability is measured on a continuous scale from 0 to 1, ranging from completely unsuitable to optimally suitable, although sometimes alternate scales are used. These indices are derived from a number of individual variables that measure important suitability factors; these suitability variables are then combined using a mathematical equation to obtain a final suitability index value. Index values are computed for individual locations in a map, such as regularly spaced raster cells, using a geographic information system (GIS). In the context of disease, suitability is often interpreted as risk when using this approach. For example, risk index models have been developed for Lyme disease (Ogden et al. 2008; Nicholson and Mather 2014), *Dirofilaria immitis* (Vezzani and Carbajo 2006), and malaria (Hagenlocher and Castro 2015). This paper uses a risk index modelling approach to map EEEV risk in Citrus County. First, a previously developed and validated risk index model is used to map transmission risk to horses. Second, that model is modified to map transmission risk to people in residential areas.

Horse risk

Vander Kelen et al. (2014) published a risk index model for predicting EEEV transmission to horses in Florida. The model was developed based on previous studies that analysed habitats associated with documented equine cases of EEEV in the state, as well as other relative literature documenting habitat associations of EEEV. The model was later validated with an independent dataset of documented horse cases over a 5-year period, confirming its ability to predict high risk areas using land use/land cover data (Downs et al. 2018). The index model assesses EEEV risk to horses based on five individual risk variables (RVs) which are combined into a final risk index (RI) that measures transmission risk on a scale of 0.0 (no risk) to 1.0 (maximal risk). Risk is calculated for each individual raster cell in a map based on the habitat type of the cell along with the composition and configuration of habitats in neighbouring cells. High risk areas are identified as habitats likely to support horses that are near large amounts of habitats associated with the EEEV transmission cycle, as described below.

The index is computed as described in Table 1. The model incorporates five risk variables into a final risk index that measures risk of EEEV transmission to horses. RV1 (local habitat) evaluates EEEV transmission risk to horses based on the land use/land cover type of the focal raster cell (i.e. the individual cell under evaluation). RV1 values are assigned based on the relative likelihood of observing a horse case in that cover type (Vander Kelen et al. 2012b). RV2 (wetland proximity) measures EEEV transmission risk to horses based on the distance between the focal raster cell and the nearest cell classified as a wetland, as observed cases are frequently found within 1.5 km of wetlands which support *Culiseta melanura* (Vander Kelen et al. 2012b). RV3 (wetland composition) measures EEEV transmission risk to horses according to the abundance of wetlands surrounding the focal cell, as wetlands of various types support vectors of EEEV (Moncayo, Edman, and Finn 2000). RV4 (tree plantation–coniferous forest proximity) measures EEEV transmission risk to horses based on the distance between the focal raster cell and the nearest cell classified as either tree plantation or coniferous forests, which are habitats associated with relevant bridge vectors (Vander Kelen et al. 2012a). RV5 (tree plantation–coniferous forest composition) measures EEEV transmission risk to horses according to the abundance of coniferous forests surrounding the focal cell. Final RI (risk index) values are computed by mathematically combining RV1–RV5 according to the equation provided in Table 1. The RI values identify locations with the combinations of habitat features that are most likely to support transmission, i.e. those with cover types likely to support horses that are near an abundance of both wetlands and tree plantations/coniferous forests, which support both the enzootic and epizootic portions of the transmission cycle. Figure 2 illustrates how the index is computed from a classified land use/land cover map; the broader spatial pattern of values follows that of cover types supporting horses – especially agricultural lands and low-density residential areas – with the final RI values within those areas reflecting the spatial distribution and abundance of wetlands and coniferous forests. For more details about the development of the RI model, please consult Vander Kelen et al. (2014). RI values were mapped for Citrus County using 30-m resolution land use/land cover data obtained from the Southwest Florida Water Management District, as previously described (Downs et al. 2018).

Residential risk

The general framework of the RI model developed to predict EEEV transmission risk to horses in Florida has been shown to be flexible for mapping EEEV risk in other

Table 1. Calculation of an Eastern equine encephalitis virus risk index model for horses **using five risk variables (RV)** .

Variable	Input	Equation
RV1 Local Habitat	Cover type of focal cell	Low density residential = 1.00 Crop-pastureland = 1.00 Tree plantations = 0.48 Upland hardwood forest = 0.41 Medium density residential = 0.27 Commercial = 0.20 High density residential = 0.14 Upland coniferous forest = 0.13 Wetland mixed forest = 0.11 Wetland coniferous forest = 0.08 Shrubland and brushland = 0.08 Vegetated nonforested wetland = 0.07 Wetland hardwood forest = 0.07 Urban = 0.03 Mining = 0.02 All others = 0.01
RV2 Wetland Proximity	Metres to nearest wetland, d_w	$= \begin{cases} 1 & \text{if } d_w \leq 400 \\ 1 - \left(\frac{d_w - 400}{1500 - 400} \right) & \text{if } 400 < d_w < 1500 \\ 0 & \text{if } d_w \geq 1500 \end{cases}$
RV3 Wetland Composition	Proportion of wetlands within 1500 metres, p_w	$= \begin{cases} 1 & \text{if } p_w \geq 0.18 \\ \frac{p_w}{0.18} & \text{if } p_w < 0.18 \end{cases}$
RV4 Tree Plantation-Coniferous Forest Proximity	Metres to nearest tree plantation or coniferous forest, d_c	$= \begin{cases} 1 & \text{if } d_c \leq 250 \\ 1 - \left(\frac{d_c - 250}{1500 - 250} \right) & \text{if } 250 < d_c < 1500 \\ 0 & \text{if } d_c \geq 1500 \end{cases}$
RV5 Tree Plantation-Coniferous Forest Composition	Proportion of tree plantations and coniferous forest within 1500 metres, p_c	$= \begin{cases} 1 & \text{if } p_c \geq 0.20 \\ \frac{p_c}{0.20} & \text{if } p_c < 0.20 \end{cases}$
RI Risk Index		$= RV1 \times \left[\frac{RV2 + RV3 + RV4 + RV5}{4} \right]$

contexts. For example, Downs et al. (2015) adapted the RI model to map EEEV transmission risk to white-tailed deer in Michigan, USA, by: (1) modifying RV1 to reflect local habitat preferences of the species (Hiller, Campa, and Winterstein 2009) and (2) replacing tree plantations-coniferous forests in RV4 and RV5 with lowland forests, a habitat type associated with EEEV and relevant vector activity in that region (Rey et al. 2012; Snow 1955). Here, the horse model is adapted to map EEEV risk to humans in residential areas in a similar way. This is accomplished by reformulating RV1 to reflect residential land use/land cover types, where locations classified as residential are assigned values of 1.0 and everything else 0.0. RV2-RV5 remain the same as in the horse model, as humans and horses play identical roles in the EEEV transmission cycle as dead-end hosts. The final RI is then calculated using the same equation to generate final RI values. However, rather than assessing risk for 30-m grid cells in the map, residential risk was mapped for individual land parcels in the study area. Parcel data were obtained from the Citrus County Property Appraiser (www.citruspa.org). RI values were calculated by assigning each parcel an RV1 value of 0.0 or 1.0 based on whether it was classified as residential or not and then multiplying that value by the average of the previously calculated RV2-RV5 values that overlaid each parcel.

Optimal sentinel site selection

Spatial optimization is a sub-speciality within the discipline of geography that focuses on solving problems with an explicit spatial component, such as how to best allocate, route, or arrange some kind of resource across a geographic area. Tong and Murray (2012) provide a comprehensive review of spatial optimization models and their applications. Spatial optimization involves formulating a geographic problem of interest into a set of mathematical equations that represent objectives and constraints, and then using some algorithm or other approach to solve for the unknown variables in the equations (Church 1974). Objectives represent the measures that are to be optimized. Single objective problems either maximize or minimize some quantity, while multi-objective problems seek some trade-off between competing measures (Farhan and Murray 2008). In either case, constraints impose restrictions on the solution to the objective(s). Objectives and constraints are most commonly written as sets of linear equations, although non-linear models are sometimes utilized. Example applications of spatial optimization models include optimally siting hospitals (ReVelle, Toregas, and Falkson 1976), designing nature reserves (Williams, ReVelle, and Levin 2004), locating warehouses (Horner and Downs 2007; Poulos et al.

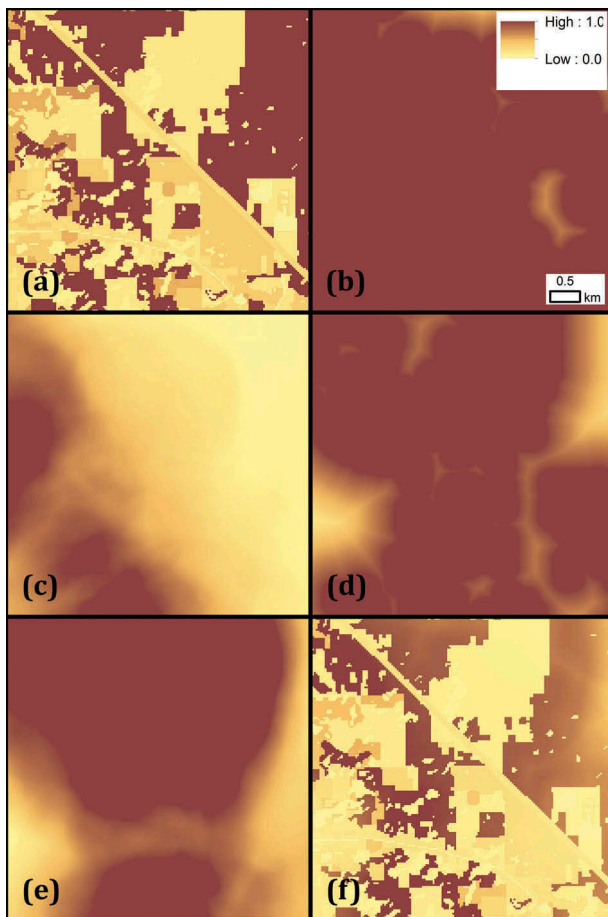


Figure 2. Computation of a risk index model for EEEV transmission to horses in Florida based on five risk variables (RVs): (a) RV1 local habitat, (b) RV2 wetland proximity, (c) RV3 wetland composition, (d) RV4 tree plantation-coniferous forest proximity, (e) RV5 tree plantation-coniferous forest composition, and (f) RI risk index.

2001), and designing public transit systems (Wu and Murray 2005), among many others.

Of all spatial optimization models, the maximal covering location model, or MCLP, is perhaps the most widely applied (Murray 2016). The MCLP was originally formulated by Church and ReVelle (1974) and is a single objective, linear spatial optimization model that seeks to maximize the coverage of demand within a desired service standard by locating a user-specified number of facilities. This paper uses the MCLP to optimally site sentinel locations for arboviral disease surveillance. In this context, demand is measured in terms of risk of EEEV transmission. Potential disease surveillance locations are considered as the facilities. The service standard, often termed a coverage distance, represents the radius around the facilities that covers demand. In the case of disease surveillance, coverage distance refers to the distance at which surveillance is considered to be effective. For EEEV, this can be interpreted the distance

at which we can detect the virus if it is actively circulating in the ecosystem. Mathematically, this problem can be formulated as:

The objective function (1) maximizes the total amount of risk r across all locations i that are covered by selected

INPUTS:	r_i = risk at location i	
	p = number of sentinel sites to locate	
	a_{ij} = 1 if candidate sentinel site j can cover risk at location i ; 0 otherwise	
MAXIMIZE:	$\sum_i r_i Z_i$	(1)
SUBJECT TO:	$\sum_j a_{ij} X_j - Z_i \geq 0$	$\forall i$ (2)
	$\sum_j X_j = p$	(3)
	$X_j = 1$ if sentinel site j is selected; 0 otherwise	$\forall j$ (4)
	$Z_i = 1$ if risk at location i is covered; 0 otherwise	$\forall i$ (5).

sentinel sites j . Essentially, the objective is a single linear equation that multiplies the risk r at each individual location i in the map by decision variable Z_i which determines whether that risk is covered as part of the solution. Constraints (2) ensure that for every location i , the risk r is only covered if a sentinel site j capable of covering that risk is selected by the model. In practice, values for a_{ij} are determined based on the proximity of the risk locations to the candidate sentinel sites; if location i is within the coverage radius of candidate sentinel site j , then a_{ij} is specified as 1 for the relevant equation, and 0 is specified if it is outside that radius. A unique equation is generated for each individual location i as a part of this set of constraints. Constraint (3) specifies the p number of sentinel sites the user wishes to locate. It is a single equation that sums the values of decision variables X_j which determine if a candidate site is selected or not. Finally, binary integer bounds are specified for decision variables X_j (4) and Z_i (5). Solving the MCLP yields binary values for all X_j and all Z_i under the objective of maximizing equation (1). In other words, the solution to the MCLP determines both which sentinel sites should be selected in order to maximally cover risk, as well as which locations are covered by each selected sentinel site. A complete set of MCLP equations for a sample problem for reference can be found in Downs et al. (2014). This formulation of the MCLP was applied to optimally site EEEV sentinel locations in Citrus County under several different scenarios. These scenarios included all combinations of three different factors: type of risk (horse or human), coverage distance (0.5, 1.0, and 1.5 km), and number of desired sentinel sites (1–12). Coverage distances up to 1.5 km were chosen since that threshold corresponds to the maximum distance that the most relevant EEEV vectors typically fly (Estep et al. 2010; Morris, Larson, and Lounibos 1991). Twelve was chosen for the maximum number of

sentinel sites, as that is approximately double what the county can currently support.

Risk was specified using the risk index values for horses and residential areas in the previous section. Because each map contained too many grid cells with risk values to apply the MCLP, since it can be computationally intensive to formulate and solve, risk values were aggregated to a slightly coarser scale in order to facilitate analysis. Risk index values were aggregated to a regular grid of hexagons, as tessellations of equal-area hexagons more accurately represent straight-line distances between cell centres than square-based grids (Birch, Oom, and Beecham 2007), which is an important consideration when mapping coverage radii for spatial optimization models. The hexagonal cells measured 150 m between their centroids and edge midpoints (Figure 3). The final grid, containing 21,891 cells, was overlaid each risk map. For the horse model, raster cells whose centres fell within a hexagon cell were summed to calculate an aggregated risk value, which essentially weights the risk by area. For the residential model, parcels whose centroids fell within hexagons were summed to calculate aggregated risk values. Use of parcel centroids captures a population density effect, as larger numbers of small parcels within a hexagon would have a higher aggregated risk than those with fewer low density parcels, *ceteris paribus*.

MCLP equations for the different scenarios were generated using the PySpatialOpt library for Python (<https://github.com/apulverizer/pyspatialopt>) integrated within a GIS (ArcGIS v 10.5 by ESRI, Inc). The PySpatialOpt tool operated from inside the GIS software outputs a text file

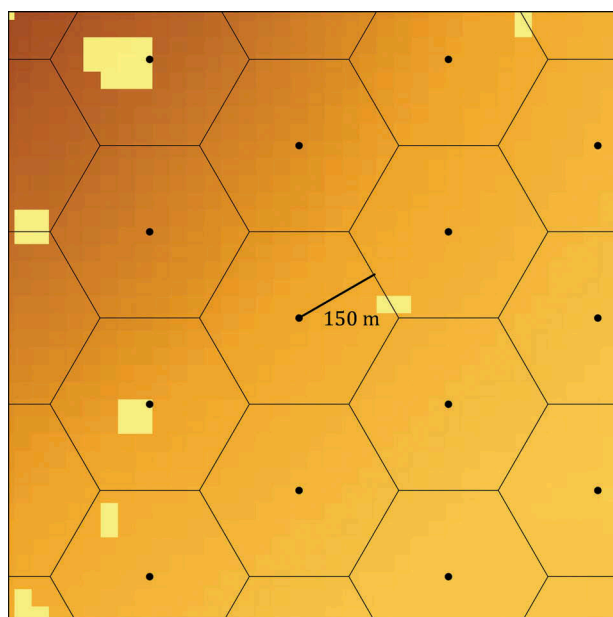


Figure 3. 150-m hexagons used to aggregate risk from individual raster cells. Darker coloured cells indicate higher risk.

of MCLP equations in linear programming (lp) format based on the input GIS layers, which include layers representing demand (risk at hexagon centroids), candidate sites (hexagon centroids), and coverage distances (buffer around hexagon centroids), as well as a specification of the number of candidate sites to select. The resulting sets of equations were then solved using a commercial optimization engine, IBM C-PLEX (IBM, Corp.). The solver output the values of the objective function and decision variables, which were then imported back into the GIS for visualization purposes.

Results

Risk index mapping

Risk index values measuring EEEV transmission to horses in Citrus County ranged from 0.0 to 1.0 (Figure 4a). RI values displayed an uneven spatial distribution. Optimal values were largely concentrated in the north- and west-central areas and the southeast of the county. Low values were

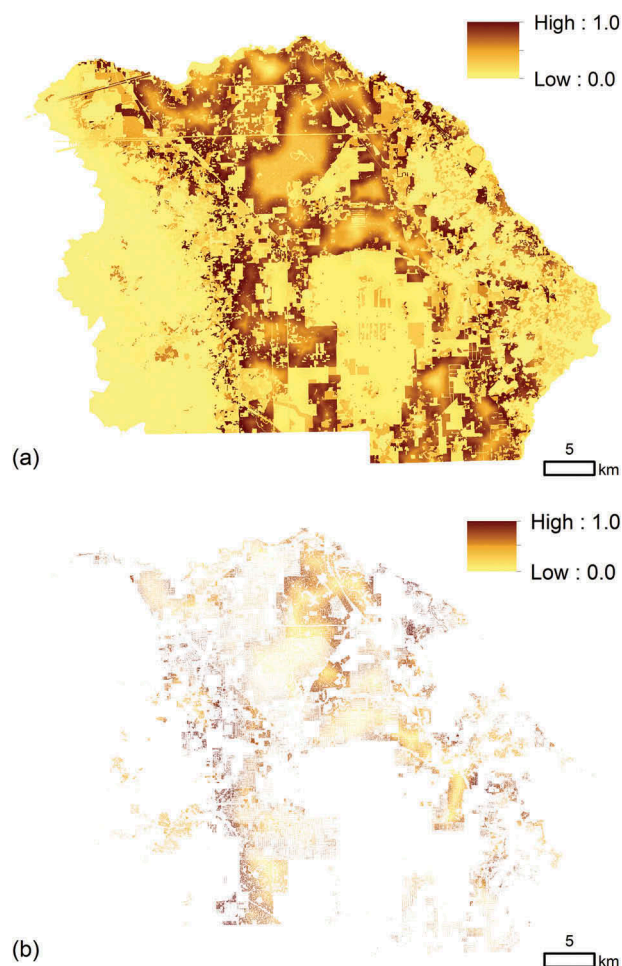


Figure 4. Final risk index (RI) maps for Citrus County: (a) horse model and (b) residential model.

widespread in coastal areas and the east-central parts of the county. The mean RI value in the map was 0.25 (s.d. = 0.26), with 15% of the raster cells receiving values between 0.50 and 0.75 and 6% receiving values greater than 0.75. This suggests that overall risk of EEEV transmission in the county is low, but there are considerable areas of high risk that are in need of monitoring.

Risk index values measuring EEEV transmission to humans in residential areas in Citrus County ranged from 0.0 to 1.0 (Figure 4b). RI values displayed a somewhat similar spatial distribution to the horse risk map, as low-density residential areas receive the highest values in both models. However, since the residential model assigns high risk to all residential areas for RV1 and assigns all other land use/cover types as no risk, the final map shows a much more restricted spatial distribution of RI values. Additionally, maximum values are also found in medium- and high-density residential areas in the latter. The average RI values assigned to residential parcels was 0.60 (s.d. = 0.18). This suggests that many residential areas in the county are located in close proximity to an abundance of wetlands and coniferous forests and should be monitored for potential EEEV transmission.

Optimal sentinel site selection

Risk index values for individual raster cells and parcels aggregated to 150-m hexagons are illustrated for both the horse and residential RI models in Figure 5. These maps represent generalized versions of the maps depicted in Figure 6. The aggregated values ranged

from about zero to 95 for the horse model and zero to 100 for the residential model. For the horse model, the highest values were generated when hexagons encompassed an area of nearly continuous optimal RI values. For the residential model, highest values occurred where high densities of high risk parcels were located. For both cases, the MCLP selected these high-value hexagons as optimal sentinel sites, particularly when they were surrounded by hexagons of similar values. Optimal sentinel sites are mapped in Figure 5 for four scenarios: (a) horse model, $p = 6$ sites, coverage radius = 1500 m; (b) horse model, $p = 12$ sites, coverage radius = 1500 m; (c) residential model, $p = 6$ sites, coverage radius = 1500 m; and (d) residential model, $p = 12$ sites, coverage radius = 1500 m. The six selected sites for the horse model are concentrated in a north-south corridor of high values that runs across the west-central portion of the county. When this number is raised to 12, four sites are added along this corridor and two in the north-central part of the county. The six selected residential sites are concentrated in the north-central part of the county; these sites are located in areas of concentrated residential parcels that are at high risk. When six additional sites are added, two occur in the same area, while the other four are located in the south-west and southeast portions of the county.

A summary of MCLP results across all scenarios is shown in Figure 6, which illustrates the objective values (total aggregated risk covered by the selected **sentinel** sites) for each coverage radius for both the horse model (a) and residential model (b). The near linear increase in the objective function as the number of sites increases –

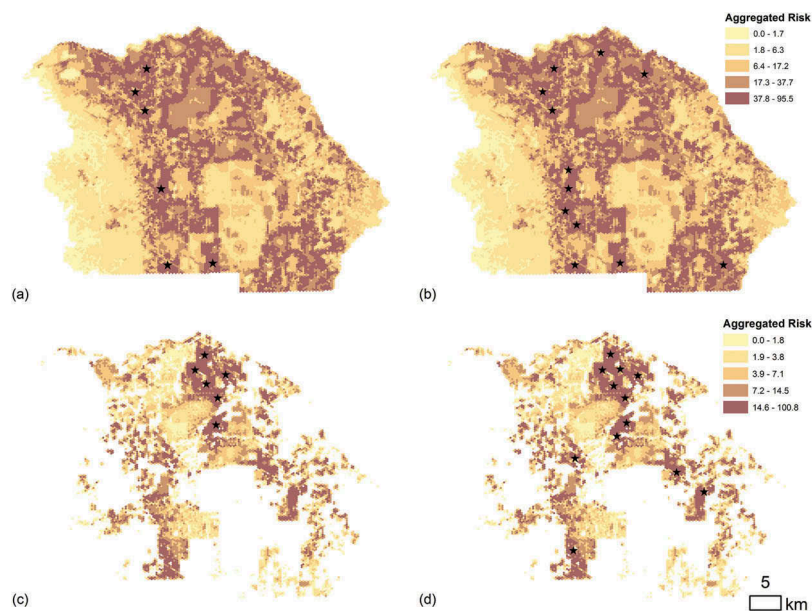


Figure 5. Optimal sites for EEEV surveillance in Citrus County for four scenarios: (a) horse model, $p = 6$ sites, coverage radius = 1500 m; (b) horse model, $p = 12$ sites, coverage radius = 1500 m; (c) residential model, $p = 6$ sites, coverage radius = 1500 m; and (d) residential model, $p = 12$ sites, coverage radius = 1500 m.

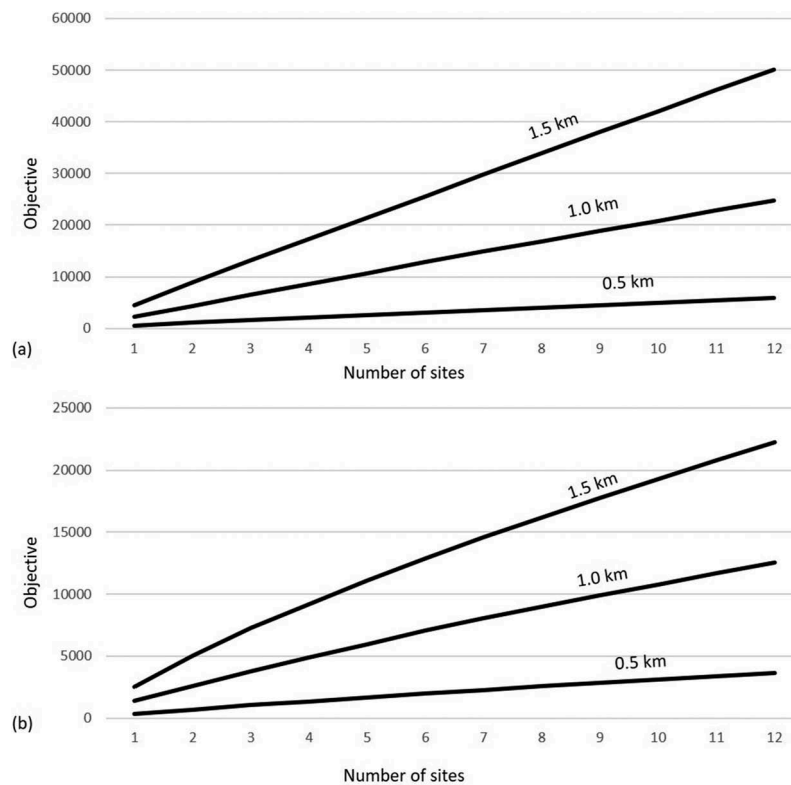


Figure 6. Objective values for the MCLP for three coverage distances (0.5, 1.0, and 1.5 km) for 1–12 selected sites.

especially at the lowest coverage distance – indicates each additional site adds a similar gain in coverage. This suggests that the number of sentinel sites within this range should be chosen based on the maximum the county can support, since each additional one provides a similar level of benefit. However, as additional sites beyond 12 are added, eventually increases in coverage will diminish. For example, for the residential scenario with a 1.5-km coverage distance, adding the 24th site (not illustrated) only adds an additional risk coverage equal to 37.7% of the first site. In spatial optimization, rates of diminishing returns are often used to determine an effective number of sites to select in practice. If much greater numbers of sentinel sites were feasible in the county, the MCLP could be solved to find the most efficient numbers of sites to sample.

Discussion and conclusions

The combined approach of risk mapping and spatial optimization offers a strategy for selecting optimal arbovirus surveillance sites under constraints of limited resources. The MCLP, in particular, offers flexibility in determining how to locate targeted sampling sites across space. In this example for EEEV, scenarios varied based on the population the model was intended to protect, the coverage distance, and the number of sites

to select. Separate solutions were found both horses and humans in residential areas, since the goals are somewhat different. The aims of the horse scenarios are to protect the largest areas at high risk rather than explicitly the largest number of horses. The rationale is two-fold: horse density data is unavailable, and incorporating density could bias the solutions towards locating the surveillance sites next individual farms with large numbers of horses at the expensive of covering larger numbers of farms. However, the human model does aim to cover the largest number of at risk people as measured by numbers of residential parcels. Despite the slightly different objectives, different strategies could be employed to survey both target populations simultaneously. One approach is to select a set number of sites from each scenario, potentially focusing on sites where risk is high in both models. An alternative approach could combine the demand from each population into a single problem. For example, this could be accomplished using multiple objectives and balancing the demand according to weights to ensure coverage of both populations (Kim, Murray, and Xiao 2008).

The main assumptions underlying this approach are that the index models accurately measure risk to the target populations and that the concept of a coverage distance is a valid method for estimating the effectiveness of sentinel surveillance methods. First, in the

context of target species, the risk model has been validated with independent data for horses but not for humans. Validation with human cases is challenging, because of the relative rarity of cases, privacy concerns, and difficulties with determining the location of transmission. Despite this, horses and humans are both bitten by the relevant mosquito species and are similarly susceptible to EEEV, so the risk model is expected to perform similar for both populations. Future work might test that assumption, however.

Second, there is uncertainty about the most appropriate coverage distance to specify when applying the MCLP to select sentinel locations. The closer to infected hosts and vectors that the sites are placed, the more likely the virus is expected to be detected. In this study, we chose coverage distances that represent average mosquito movements, but the amount of area a sentinel site can protect may be smaller or larger in practice. The coverage distance is important not only because it measures the effectiveness of sentinel sites, but also because it determines the spacing between them in the final solution. At lower distances, the chosen sites for our case study tended to cluster in high risk areas at the expense of locating sites as close as possible to risk, especially for the residential model. If sentinel sites do operate with a low coverage distance in reality but clustering is undesirable in order to more broadly survey viral activity in the county, then an additional set of constraints could be formulated in order to consider the spacing between selected sites. Future work should determine the effective distances and optimal spacing for EEEV surveillance methods in order to best protect the state.

Once the target population and appropriate coverage distances are determined, the number of sites to select can be determined based on economic or other limiting resources. Since the particular combination of sites can vary with the value of p , a long-term target might be used to designate sites if it is not practical to move sites once they are established. For example, if Citrus County has a target of 12 but can only currently service six, then the solution of 12 might be used to select the set of current and future sites. Six sites from the solution could be selected for initial sentinel sites, either the ones with the highest risk or a selection that offers broader spatial coverage across the county. Additionally, some selected sites might be inaccessible or infeasible to survey, in which case the model could be solved without those candidate sites. Future work might use field and laboratory studies to evaluate the effectiveness of these different strategies. The next step in our research is to determine if the surveillance sites selected by the optimization model do indeed allow us to detect

EEEV more efficiently than existing sites currently employed.

Although the MCLP is relatively simple to solve with commercial GIS and optimization software, other spatial optimization models might also provide useful approaches to select arbovirus surveillance sites. For example, the set covering location problem works similarly but, rather than maximizing coverage, it selects the minimum number of sites that provides complete coverage. Although such a strategy is not feasible in Citrus County, it could be useful in situations where the region of interest is small, resources are unlimited, or the virus is urgent to control. Additionally, in a similar way that a single model could be used to cover different populations, a combined approach could be used to develop a comprehensive strategy for surveying multiple types of arboviruses in an area. By considering risk for multiple diseases at once, an optimization model could be used to configure the most efficient arrangement of sites for surveying all of them. Multi-objective models could be used for this purpose if accurate index models are available for mapping risk for each virus. Future work might focus on developing risk index models for other types of arboviruses of interest, such as West Nile Virus, and developing a more comprehensive sampling strategy for Citrus County. In summary, risk modelling and spatial optimization can be used as tools to strategically design arbovirus surveillance networks.

Acknowledgements

This publication was supported by Cooperative Agreement Number U01CK000510, funded by the Centers for Disease Control and Prevention. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Centers for Disease Control and Prevention or the Department of Health and Human Services.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the Centers for Disease Control and Prevention [Cooperative Agreement Number U01CK000510].

References

- Armstrong, P. M., T. G. Andreadis, J. F. Anderson, J. W. Stull, and C. N. Mores. 2008. "Tracking Eastern Equine Encephalitis Virus Perpetuation in the Northeastern United States by Phylogenetic Analysis." *American Journal of Tropical*

- Medicine and Hygiene* 79 (2): 291–296. doi:10.4269/ajtmh.2008.79.291.
- Birch, C. P. D., S. P. Oom, and J. A. Beecham. 2007. "Rectangular and Hexagonal Grids Used for Observation, Experiment and Simulation in Ecology." *Ecological Modelling* 206 (3): 347–359. doi:10.1016/j.ecolmodel.2007.03.041.
- Brooks, R. P. 1997. "Improving Habitat Suitability Index Models." *Wildlife Society Bulletin (1973–2006)* 25 (1): 163–167.
- Burkett-Cadena, N. D., J. Gibson, M. Lauth, T. Stenn, C. Acevedo, R. D. Xue, J. McNelly, et al. 2016. "Evaluation of the Honey-card Technique for Detection of Transmission of Arboviruses in Florida and Comparison with Sentinel Chicken Seroconversion." *Journal of Medical Entomology* 53 (6): 1449–1457. doi:10.1093/jme/tjw106.
- CDC. 2017. "Eastern Equine Encephalitis, Epidemiology and Geographic Distribution." <https://www.cdc.gov/easterequineencephalitis/tech/epi.html>
- Chevalier, V., S. Lecollinet, and B. Durand. 2011. "West Nile Virus in Europe: A Comparison of Surveillance System Designs in A Changing Epidemiological Context." *Vector Borne and Zoonotic Diseases (Larchmont, N.Y.)* 11 (8): 1085–1091. doi:10.1089/vbz.2010.0234.
- Church, R., and C. ReVelle. 1974. "The Maximal Covering Location Problem." Paper presented at the Papers of the Regional Science Association.
- Church, R., and C. ReVelle. 1974. "The Maximal Covering Location Problem." *Papers Of The Regional Science Association* 32 (1): 101–118. doi:10.1007/BF01942293.
- Dowdle, W. R. 1998. "The Principles of Disease Elimination and Eradication." *Bulletin of the World Health Organization* 76 (Suppl 2): 22.
- Downs, J., M. Horner, R. Loraamm, J. Anderson, H. Kim, and D. Onorato. 2014. "Strategically Locating Wildlife Crossing Structures for Florida Panthers Using Maximal Covering Approaches." *Transactions in GIS* 18 (1): 46–65. doi:10.1111/tgis.2014.18.issue-1.
- Downs, J., M. Vaziri, A. Jenkins, and T. Unnasch. 2018. "Validation of a Risk Index Model for Predicting Eastern Equine Encephalitis Virus Transmission to Horses in Florida." *Journal of Medical Entomology*. doi:10.1093/jme/tjy067.
- Downs, J. A., G. Hyzer, E. Marion, Z. J. Smith, P. Vander Kelen, and T. R. Unnasch. 2015. "Mapping Eastern Equine Encephalitis Virus Risk for White-tailed Deer in Michigan." *Applied Geography* 64 (Supplement C): 66–73. doi:10.1016/j.apgeog.2015.09.006.
- Downs, J. A., R. J. Gates, and A. T. Murray. 2008. "Estimating Carrying Capacity for Sandhill Cranes Using Habitat Suitability and Spatial Optimization Models." *Ecological Modelling* 214 (2–4): 284–292. doi:10.1016/j.ecolmodel.2008.02.006.
- Eldridge, B. F. 1987. "Strategies for Surveillance, Prevention, and Control of Arbovirus Diseases in Western North America." *American Journal of Tropical Medicine and Hygiene* 37 (3_Part_2): 775–865. doi:10.4269/ajtmh.1987.37.775.
- Estep, L. K., N. D. Burkett-Cadena, G. E. Hill, R. S. Unnasch, and T. R. Unnasch. 2010. "Estimation of Dispersal Distances of Culex Erraticus in a Focus of Eastern Equine Encephalitis Virus in the Southeastern United States." *Journal of Medical Entomology* 47 (6): 977–986. doi:10.1603/ME10056.
- Farhan, B., and A. T. Murray. 2008. "Siting Park-and-ride Facilities Using a Multi-objective Spatial Optimization Model." *Computers & Operations Research* 35 (2): 445–456. doi:10.1016/j.cor.2006.03.009.
- Gu, W., T. R. Unnasch, C. R. Katholi, R. Lampman, and R. J. Novak. 2008. "Fundamental Issues in Mosquito Surveillance for Arboviral Transmission." *Transactions of the Royal Society of Tropical Medicine and Hygiene* 102 (8): 817–822. doi:10.1016/j.trstmh.2008.03.019.
- Gubler, D. J. 2002. "The Global emergence/resurgence of Arboviral Diseases as Public Health Problems." *Archives of Medical Research* 33 (4): 330–342. doi:10.1016/S0188-4409(02)00378-8.
- Hagenlocher, M., and M. C. Castro. 2015. "Mapping Malaria Risk and Vulnerability in the United Republic of Tanzania: A Spatial Explicit Model." *Population Health Metrics* 13 (1): 2. doi:10.1186/s12963-015-0036-2.
- Hall, R. A., B. J. Blitvich, C. A. Johansen, and S. D. Blacksell. 2012. "Advances in Arbovirus Surveillance, Detection and Diagnosis." *BioMed Research International* 16: 2012.
- Hiller, T. L., H. Campa, and S. R. Winterstein. 2009. "Estimation and Implications of Space Use for White-Tailed Deer Management in Southern Michigan." *Journal of Wildlife Management* 73 (2): 201–209. doi:10.2193/2008-041.
- Horner, M., and J. Downs. 2007. "Testing a Flexible Geographic Information System-based Network Flow Model for Routing Hurricane Disaster Relief Goods." *Transportation Research Record: Journal of the Transportation Research Board* 2022: 47–54. doi:10.3141/2022-06.
- Hotez, P. J. 2009. "Mass Drug Administration and Integrated Control for the World's High-Prevalence Neglected Tropical Diseases." *Clinical Pharmacology & Therapeutics* 85 (6): 659–664. doi:10.1038/clpt.2009.16.
- Kean, J., S. M. Rainey, M. McFarlane, C. L. Donald, E. Schnettler, A. Kohl, and E. Pondeville. 2015. "Fighting Arbovirus Transmission: Natural and Engineered Control of Vector Competence in Aedes Mosquitoes." *Insects* 6 (1): 236–278. doi:10.3390/insects6010236.
- Kim, K., A. T. Murray, and N. Xiao. 2008. "A Multiobjective Evolutionary Algorithm for Surveillance Sensor Placement." *Environment and Planning B: Planning and Design* 35 (5): 935–948. doi:10.1068/b33139.
- Moncayo, A. C., J. D. Edman, and J. T. Finn. 2000. "Application of Geographic Information Technology in Determining Risk of Eastern Equine Encephalomyelitis Virus Transmission." *Journal of the American Mosquito Control Association* 16 (1): 28–35.
- Morris, C. D., V. L. Larson, and L. P. Lounibos. 1991. "Measuring Mosquito Dispersal for Control Programs." *Journal of the American Mosquito Control Association* 7 (4): 608–615.
- Murray, A. T. 2016. "Maximal Coverage Location Problem: Impacts, Significance, and Evolution." *International Regional Science Review* 39 (1): 5–27. doi:10.1177/0160017615600222.
- Nicholson, M. C., and T. N. Mather. 2014. "Methods for Evaluating Lyme Disease Risks Using Geographic Information Systems and Geospatial Analysis." *Journal of Medical Entomology* 33 (5): 711–720. doi:10.1093/jmedent/33.5.711.
- Ogden, N. H., L. St-Onge, I. K. Barker, S. Brazeau, M. Bigras-Poulin, D. F. Charron, C. M. Francis, A. Heagy, L. Lindsay, and A. Maarouf. 2008. "Risk Maps for Range Expansion of the Lyme Disease Vector, Ixodes Scapularis, in Canada Now

- and with Climate Change." *International Journal of Health Geographics* 7 (1): 24. doi:[10.1186/1476-072X-7-24](https://doi.org/10.1186/1476-072X-7-24).
- Olliaro, P., F. Fouque, A. Kroeger, L. Bowman, R. Velayudhan, A. C. Santelli, D. Garcia, R. S. Ramm, L. H. Sulaiman, and G. S. Tejeda. 2018. "Improved Tools and Strategies for the Prevention and Control of Arboviral Diseases: A Research-to-policy Forum." *PLoS Neglected Tropical Diseases* 12 (2): e0005967. doi:[10.1371/journal.pntd.0005967](https://doi.org/10.1371/journal.pntd.0005967).
- Poulos, P. N., G. G. Rigatos, S. G. Tzafestas, and A. K. Koukos. 2001. "A Pareto-optimal Genetic Algorithm for Warehouse Multi-objective Optimization." *Engineering Applications of Artificial Intelligence* 14 (6): 737–749. doi:[10.1016/S0952-1976\(01\)00036-7](https://doi.org/10.1016/S0952-1976(01)00036-7).
- Ramírez, A. L., A. F. van den Hurk, D. B. Meyer, and S. A. Ritchie. 2018. "Searching for the Proverbial Needle in a Haystack: Advances in Mosquito-borne Arbovirus Surveillance." *Parasites & Vectors* 11 (1): 320. doi:[10.1186/s13071-018-2901-x](https://doi.org/10.1186/s13071-018-2901-x).
- ReVelle, C., C. Toregas, and L. Falkson. 1976. "Applications of the Location Set-covering Problem." *Geographical Analysis* 8 (1): 65–76. doi:[10.1111/j.1538-4632.1976.tb00529.x](https://doi.org/10.1111/j.1538-4632.1976.tb00529.x).
- Rey, J., W. Walton, R. Wolfe, C. Connelly, S. O'Connell, J. Berg, G. Sakolsky-Hoop, and A. Laderman. 2012. "North American Wetlands and Mosquito Control." *International Journal of Environmental Research and Public Health* 9 (12): 4537. doi:[10.3390/ijerph9124537](https://doi.org/10.3390/ijerph9124537).
- Ritchie, S. A., A. F. Van Den Hurk, P. Zborowski, T. J. Kerlin, D. Banks, J. A. Walker, J. M. Lee, et al. 2007. "Operational Trials of Remote Mosquito Trap Systems for Japanese Encephalitis Virus Surveillance in the Torres Strait, Australia." *Vector-Borne and Zoonotic Diseases* 7 (4): 497–506. doi:[10.1089/vbz.2006.0643](https://doi.org/10.1089/vbz.2006.0643).
- Sedda, L., E. R. Lucas, L. S. Djogbenou, A. V. Edi, A. Egyir-Yawson, B. I. Kabula, J. Midega, E. Ochomo, D. Weetman, and M. J. Donnelly. 2019. "Improved Spatial Ecological Sampling Using Open Data and Standardization: An Example from Malaria Mosquito Surveillance." *Journal of the Royal Society Interface* 16 (153): 20180941. doi:[10.1098/rsif.2018.0941](https://doi.org/10.1098/rsif.2018.0941).
- Snow, W. E. 1955. "Feeding Activities of Some Blood-Sucking Diptera with Reference to Vertical Distribution in Bottomland Forest." *Annals of the Entomological Society of America* 48 (6): 512–521. doi:[10.1093/aesa/48.6.512](https://doi.org/10.1093/aesa/48.6.512).
- Tong, D., and A. T. Murray. 2012. "Spatial Optimization in Geography." *Annals of the Association of American Geographers* 102 (6): 1290–1309. doi:[10.1080/00045608.2012.685044](https://doi.org/10.1080/00045608.2012.685044).
- USDA. 2017. "Eastern Equine Encephalitis - States with Equine Cases." *United States Department of Agriculture*. https://www.aphis.usda.gov/aphis/ourfocus/animalhealth/animal-disease-information/horse-disease-information/sa_encephalitis/ct_eee_distribution_maps
- Vander Kelen, P., J. A. Downs, T. Unnasch, and L. Stark. 2014. "A Risk Index Model for Predicting Eastern Equine Encephalitis Virus Transmission to Horses in Florida." *Applied Geography* 48: 79–86. doi:[10.1016/j.apgeog.2014.01.012](https://doi.org/10.1016/j.apgeog.2014.01.012).
- Vander Kelen, P. T., J. A. Downs, L. M. Stark, R. W. Loraamm, J. H. Anderson, and T. R. Unnasch. 2012b. "Spatial Epidemiology of Eastern Equine Encephalitis in Florida." *International Journal of Health Geographics* 11 (1): 47. doi:[10.1186/1476-072X-11-47](https://doi.org/10.1186/1476-072X-11-47).
- Vander Kelen, P. T., J. A. Downs, N. D. Burkett-Cadena, C. L. Ottendorfer, K. Hill, S. Sickerman, J. Hernandez, J. Jinright, B. Hunt, and J. Lusk. 2012a. "Habitat Associations of Eastern Equine Encephalitis Transmission in Walton County Florida." *Journal of Medical Entomology* 49 (3): 746–756. doi:[10.1603/ME11224](https://doi.org/10.1603/ME11224).
- Vezzani, D., and A. E. Carbajo. 2006. "Spatial and Temporal Transmission Risk of *Dirofilaria immitis* in Argentina." *International Journal for Parasitology* 36 (14): 1463–1472. doi:[10.1016/j.ijpara.2006.08.012](https://doi.org/10.1016/j.ijpara.2006.08.012).
- Williams, J. C., C. S. ReVelle, and S. A. Levin. 2004. "Using Mathematical Optimization Models to Design Nature Reserves." *Frontiers in Ecology and the Environment* 2 (2): 98–105. doi:[10.1890/1540-9295\(2004\)002\[0098:UMOMTD\]2.0.CO;2](https://doi.org/10.1890/1540-9295(2004)002[0098:UMOMTD]2.0.CO;2).
- Wu, C., and A. T. Murray. 2005. "Optimizing Public Transit Quality and System Access: The Multiple-Route, Maximal Covering/Shortest-Path Problem." *Environment and Planning B: Planning and Design* 32 (2): 163–178. doi:[10.1068/b31104](https://doi.org/10.1068/b31104).