Surveillance of Invading Mosquitoes Using Occupancy Estimation and Modeling

George W. Peck

Clackamas County Vector Control District, 1102 Abernethy Rd., Oregon City, OR 97045.
Phone: (661) 565-7583; Email: gwpeck5@gmail.com

ABSTRACT: Invasion of mosquitoes into previously unoccupied habitats is an ongoing concern in many regions of California. Recognizing and controlling precision and bias in the design and implementation of vector surveillance efforts could ensure the success of programs focused on detecting invasive mosquitoes. Surveillance programs may operate under the assumption that the mosquito detection probability (*p*) of their sampling method is perfect (no false negative errors). However, non-detection of a mosquito species at a survey site does not imply that the species is absent, unless the probability of detection is perfect ($p = 1$). This paper introduces new methods of planning and conducting mosquito surveillance using a simple model for estimating mosquito site occupancy probability. A simulation is used to examine how detection probability influences mosquito site occupancy probability estimates, within the context of a typical mosquito surveillance effort, with an emphasis on detection of rare or invasive species. Real surveillance data, analyzed for comparison with simulations, is also discussed.

INTRODUCTION

Surveillance of vectors of human disease usually focuses on two factors: presence and abundance. Surveillance actions are conducted to determine if vectors are present in a given habitat and if present, how abundant they are. There are many types of sampling methods for mosquitoes including ovi-cups, dip samples, and baited adult mosquito traps placed in appropriate locations. Trap placement strategies include random, systematic, stratified, favorite habitat and others (Reisen and Lothrop 1999). Trap type, number and placement are dictated by mosquito control agency logistical constraints, target species and mosquito biology. The question what trap to use and how often to sample for an invading mosquito in a given location can be complex and may be partially or wholly answered by occupancy model analysis. Wildlife biologists have known for over a decade that their analyses of animal presence and population size must be conditioned on an estimate of detection probability (McKenzie et al. 2002). Occupancy models explicitly incorporate detection probability in their estimation of occupancy estimates, controlling for parameter bias and precision. Interestingly, the models and methods of occupancy analysis have recently been adapted for use with selected vectors of human disease (Abad-Franch et al. 2010, Padilla-Torres et al. 2013).

California continues to experience invasions of mosquitoes into previously unoccupied areas (CDPH 2015). Mosquito surveillance data that is precise and unbiased are a critical component of programs aimed at detecting invasive mosquitoes. In order to translate the theory and methods of occupancy modelling into the context of mosquito surveillance, this paper uses the strict statistical definition of precision and bias as defined in MacKenzie, et al. (2006). Precision can be thought of as a measure of how closely repeated measurements of population parameters (e.g., mosquito presence [or absence] or population size) cluster together, whereas bias is a measure of how far away an estimated parameter is from a true (but unknown) population parameter. Mosquito surveillance programs may unconsciously operate under the assumption that the mosquito detection probability (*p*) of their sampling method is perfect (no false negative errors). However, non-detection of a mosquito species at a survey site does not imply that the species is absent unless the probability of detection is perfect ($p = 1$). Although theoretically possible, in practice this is rarely the case; *p* is always < 1. This paper provides a brief overview of the use of occupancy models to explore how sampling design effects bias and precision of occupancy estimates for mosquitoes in a range of theoretical habitat types that could correspond to a spatial invasion wave, and natural habitat within the Coachella Valley of California.

METHODS

Simulation. GENPRES software (USGS-PWRC 2015) was used to simulate mosquito occupancy data and thereby generate estimates of occupancy probability (OP) and related statistics. For simulations, input values of OP were chosen that corresponded to mosquito habitat quality under perfect detection ($p = 1.0$). Input OP values were assigned as follows: 0.1 (very poor habitat), 0.3 (poor habitat) and 0.5 (average habitat). These values of the input OP can also be interpreted as describing the dynamics of a mosquito invasive wave, with input OP = 0.1 corresponding to the outer edge of the wave, OP = 0.3 to the mid-point between the wave edge and population center and OP = 0.5 to the center of the invasive population. Simulation values of detection probability (*p*) were chosen to correspond to various sampling situations. A *p* value of 0.1 describes a very poor mosquito trap, or a very cryptic species; a *p* value of 0.3 would describe a poor mosquito trap or a relatively cryptic species; and a *p* value of 0.5 would describe a mosquito trap of average efficiency, or a relatively abundant species, or field technician whose larval dipping efficiency is about average, etc. Twenty was chosen as the number of sites being surveyed, as this might reflect a typical weekly surveillance effort from a mosquito control agency. A site can be envisioned as any likely mosquito habitat, including a permanent wetland, an urban neighborhood, an agricultural waste sump, etc. *S* is the site-specific number of trap deployments and varied from 2 to 5. Five hundred GENPRES occupancy data set simulations were generated for each combination of input OP, *p*, and *S* (36 total combinations), and summary statistics were calculated and reported for each OP-*S*-*p* combination.

Occupancy Analysis of Coachella Valley Trap Data. PRESENCE software (USGS-PWRC 2015) was used to analyze real mosquito occupancy data. For comparison with an estimated OP, PRESENCE calculates a naïve OP, equivalent to the input OP in GENPRES under perfect detection ($p = 1.0$). For the study presented here, a subset of count data from a two-year mosquito trapping study using 63 dry ice-baited CDC traps set across the northern end of the Salton Sea (Figure 1) was used (Reisen and Lothrop 1999).

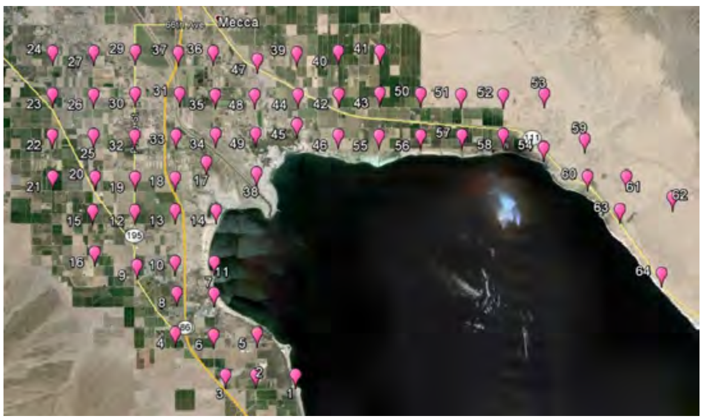

Figure 1. Sampling design study area, northern Salton Sea, CA (Reisen and Lothrop 1999).

The data subset was comprised of occupancy data for five trap nights during the spring of the 1995 season (calendar weeks 16, 18, 20, 22 and 24 of April, May and June; all traps were deployed overnight at two-week intervals). Collected mosquitos were transported live to the laboratory, anesthetized with triethylamine and enumerated by species (see Reisen and Lothrop 1999 for details). For occupancy analysis, counts of mosquitoes were converted to a binary representation, with 0 indicating absence and 1 indicating presence for a given trap site-survey date combination. The focus of the original Coachella Valley study was *Culex tarsalis* L. For the data subset examined here, 83.1% of the mosquitoes captured (N = 118,071) were *Cx. tarsalis*. However, relatively rare species are the focus of the present study, especially since rare species are used as a surrogate for the leading edge of a wave of an invading mosquito species. Therefore, this occupancy study focused on three relatively rare species: *Culex quinquefasciatus* Say (N = 5,631 or 4.0%), *Aedes vexans* (Meigan) (N = 289 or 0.2%) and *Anopheles franciscanis* McCraken (463 or 0.%).

RESULTS

Simulation. Increasing the number of site specific deployments (*S*) and increasing the detection probability (*p*) were found to decrease bias in simulated OP values (Figure 2A); increasing values of *p,* and *S* within a given *p,* allowed the simulated OP values to more closely approximate the input OP (legend in Figure 2A). Precision, measured as a relative decrease in the value of the simulated OP standard deviation (smaller values of OP standard deviation indicate greater estimate precision, $n = 500$ simulations), increased with increasing input OP, *p* within OP, and for *S* within *p* within OP (Figure 2B).

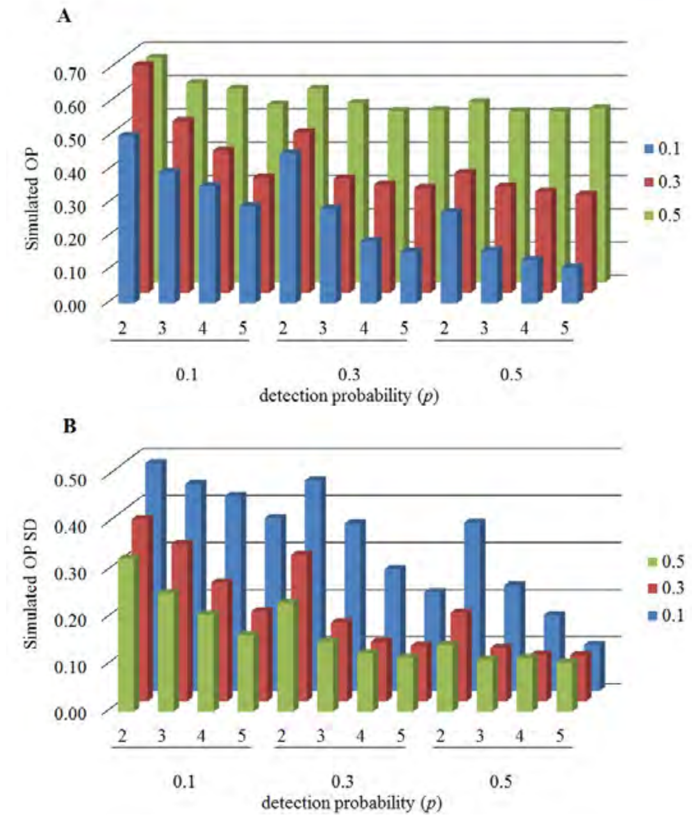

Figure 2. Occupancy Probability (OP) bias and precision analysis using GENPRES simulation results. **A**. Bias analysis. Increasing the number of site specific deployments (numbers below columns, *S*) and increasing the detection probability (*p*) were found to decrease bias [e.g., simulated OP values more closely approximated the input OP (legend values) with increasing *p,* and increasing *S* within *p*]. **B.** Precision analysis. Increasing input OP (legend values), *S* and *p* increased precision of OP estimates in all simulated data sets [e.g., the relative size of simulated OP standard deviation (denoted as SD) decreased with increasing input OP, *S*, and *p*]. OP values in legends are the occupancy probabilities under the assumption of perfect detecton ($p = 1.0$) and were used as the input value for simulations: OP = 0.1 simulates the edge of an invasion wave or very poor habitat; OP = 0.3 simulates a region near the edge of an invasion wave or poor habitat; and OP = 0.5 simulates the middle of an invasion wave or average habitat quality. Detection probability values (*p*) can be interpreted as follows: 0.3 = poor trap design, 0.5 average trap design and 0.7 = good trap design. Each combination of input OP, *p* and *S* (36 total combinations) was simulated 500 times with GENPRES.

Coachella Valley Trap Data. Increasing *S* strongly increased estimate precision (decreasing standard error [SE]) for *p* and OP for all three species (Table 1). Increasing *S* decreased OP estimate bias, measured as a percent difference between the estimated OP and the naïve OP. Rarity of species also influenced percent difference between estimated OP and the naïve OP, with the largest value (41.2%) for the rarest species (*Ae. vexans*).

Species	Survey # (*S*)	Est *p* (SE)	Est OP (SE)	Naïve OP	% difference OP
Cx. quinques	2	0.748 (0.062)	0.674 (0.071)	0.619	8.5
	3	0.697 (0.046)	0.692 (0.062)	0.667	3.6
	4	0.661 (0.037)	0.760 (0.056)	0.746	1.8
	5	0.669 (0.031)	0.784 (0.053)	0.778	0.7
Ae. vexans	2	0.372 (0.134)	0.459 (0.160)	0.270	41.2
	3	0.323 (0.091)	0.398 (0.110)	0.270	32.2
	4	0.262 (0.057)	0.597 (0.119)	0.413	30.8
	5	0.285 (0.047)	0.559 (0.090)	0.444	20.6
An. franciscanis	2	0.600 (0.106)	0.424 (0.085)	0.349	17.7
	3	0.550 (0.070)	0.441 (0.072)	0.397	10.0
	4	0.494 (0.050	0.564 (0.070)	0.524	7.1
	5	0.488 (0.042)	0.581 (0.066)	0.556	4.3

Table 1. Results of PRESENCE analysis for selected population parameters of Coachella Valley mosquitoes. Estimated detection probability is given by Est *p*, displayed with one standard error (SE, $n = 63$). Estimated occupancy probability is given by Est OP, displayed with one standard error (SE, $n = 63$). The naïve OP value assumes that detection probability is perfect ($p = 1.0$). The percent difference between naïve and estimated OP (a measure of bias) was influenced by species and survey number, with the rarest species (*Ae. vexans*) having the greatest bias for $S = 2$.

DISCUSSION

This preliminary study explored how a simple occupancy model could be used to provide precise and unbiased mosquito population OP estimates, and how those OP estimates might be influenced by different habitat types (modeled as input OP), changing detection probabilities (*p*), and changing survey number (*S*). Input OP, *p* and *S* strongly influenced all simulated OP estimates and associated statistics. For very rare species or for the initial colonizers of an invasive wave of mosquitoes, understanding the interplay between the expected OP for a given habitat, detection probability of a given trap type and the choice of *S* in space or time may be crucial for surveillance program success. OP for a given habitat can be estimated from previous invasion surveillance data, while the detection probability of a given trap for a given species can be estimated from field data or laboratory experiments using the occupancy models described here. The effect of changing *S* on a given habitat-trap combination can be estimated using previously collected data as well. Mosquito district personnel may want to consider exploring simple and complex occupancy models (McKenzie et al. 2006) with careful attention to the interplay of OP, *p* and *S* when formulating and implementing plans for mosquito surveillance. When invading species are expected, extra surveys (increasing *S* above a baseline) in the most likely habitats will increase the precision of any estimate of occupancy, while decreasing the bias between the true (but unknown) and estimated occupancy probability. The simple binary occupancy model used here can be easily modified to incorporate Poisson counts, leading to estimates of population density (mosquitoes per unit area) and thus can be used as a decision tool for management personnel grappling with questions about whether a given site has reached a threshold that requires control interventions.

ACKNOWLEDGEMENTS

Thanks to Bill Reisen and the American Mosquito Control Association for permission to use the mosquito count data from Reisen and Lothrop (1999). Thanks to Hugh Lothrop for details on trap placement and for generating the initial file on Google Earth for Fig. 1 (Map data ©2015 Google).

REFERENCES CITED

Abad-Franch, F., G. Ferraz, C. Campos, F. S. Palomeque, M. J. Grijalva, H. M. Aguilar, and M. A. Miles. 2010. Modeling Disease Vector Occurrence when Detection Is Imperfect: Infestation of Amazonian Palm Trees by Triatomine Bugs at Three Spatial Scales. PLoS Negl Trop Dis 4: e620.

CDPH (California Department of Public Health). 2015. Source: http://www.cdph.ca.gov/HealthInfo/discond/Documents/AedesDistributionMap5-28-15.pdf

MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. Andrew Royle, and C. A. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology 83: 2248-2255.

MacKenzie, D., J. Nichols, J. Royle, K. Pollock, L. Bailey, and J. Hines. 2006. Occupancy Estimation and Modeling: Inferring Patterns and Dynamics of Species Occurrence, Elsevier Academic Press, San Diego, CA.

Padilla-Torres, S. D., G. Ferraz, S. L. B. Luz, E. Zamora-Perea, and F. Abad-Franch. 2013. Modeling Dengue Vector Dynamics under Imperfect Detection: Three Years of Site-Occupancy by *Aedes aegypti* and *Aedes albopictus* in Urban Amazonia. PLoS ONE 8: e58420.

Reisen, W. K., and H. D. Lothrop. 1999. Effects of sampling design on the estimation of adult mosquito abundance. J Amer Mosq Cont Assoc 15: 105-114.

USGS-PWRC. 2015. GENPRES and PRESENCE. 9.3 - Estimates patch occupancy rates and related parameters. http://www.mbr-pwrc.usgs.gov/software/presence.html