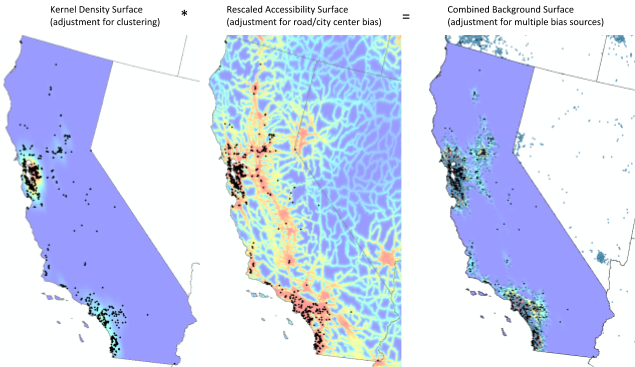
Simulation of Background Data for California.

The development of a survey bias surface for generating background points was a two-step process. First we accounted for clustering of survey effort by generating a kernel density estimator (KDE) using the smoothed cross-validation bandwidth selector from the R Kernel Smoothing (“ks”) library [110]. Roadside accessibility is a common form of survey bias that can affect model predictions [111]. We obtained a global map of accessibility which quantifies travel time to locations of interest such as a road or urban center [112]. Using the positive detection data in California, we intersected the detections with the accessibility surface to extract accessibility values. We fitted an exponential model to the extracted data to quantify the frequency of gypsy moth detection with accessibility (y = 5113.3e(-0.006\*Accessibility), R2=0.9149). We applied the fitted model to the accessibility surface. We standardized both surfaces to 0-1 and multiplied them together times 100 to generate a combined probability surface of survey bias. This final 0-100 probability surface was used in the Vistrails Software for Assisted Habitat Modeling modules to generate background points for California.



**Figure S1.** Sample bias surface statistically derived from clustering of survey effort and accessibility. The two bias surfaces on the left are standardized and then multiplied to create the bias surface on the right, from which background samples are generated and combined with known trap negatives (blue dots).