Our statistical model is similar to models that are frequently used for wildlife studies known as occupancy models. These occupancy models attempt to understand both the factors that determine which sites are occupied by the target species as well as which factors influence the detection of this species given that the site is occupied. In our case, we are interested in determining which factors influence the invasion probability of counties as well as which factors influence the detectability of leishmaniasis cases once a county has been successfully invaded.

Let denote if cases were observed in county i at time t (0 and 1 represent the absence and presence of cases, respectively). We assume the following Bernoulli distribution for this variable:

where is the detection probability and is the invasion status of county i at time t. Notice that if a county is not invaded (i.e., ), this model assumes that there will be no observed cases (i.e., ). We assume that the detection probability depends on several covariates through the following relationship:

where is the design vector containing covariates and is a vector of regression coefficients associated with the detection probability.

In relation to the latent invasion status for county i at time 1 , we assume that there is a fixed but unknown probability of each county already being invaded in the beginning of the study period, given by . More specifically, we assume that:

where is a real number between zero and one. For subsequent years, we use a logistic regression framework to model the invasion process assuming the focus county has not been invaded yet:

where is the design vector containing covariates and is a vector of regression coefficients associated with the invasion process.

We also assume that once a location has been invaded, it remains invaded. This can be written as:

Finally, we finish specifying this model by assuming that:

where and are diagonal matrices specified by the modeler.

#-------------------------------------------------------------------

We developed a Gibbs sampler to estimate the parameters and . The full conditional distributions required for this algorithm are described below:

We sample these parameters using an MH algorithm.

where A is the set of counties that have not been invaded in year 1 and is the year that county i was invaded (i.e., ). We sample these parameters using an MH algorithm.

Because of conjugacy, it can be shown that this is equivalent to:

where and are the locations estimated to be invaded and not invaded in year 1, respectively. In other words, and .

We jointly sample all the invasion status variables for each location with no leishmaniasis cases in year 1 (i.e., ). The full conditional distribution is given by:

To sample from this distribution, we proposed a new vector and we accept this proposal or reject it in favor of using the probability ratio described below.

Assume that there has been no cases for that county between years 1 and t (i.e., ). With equal probabilities we randomly proposed that this site was never invaded or that it was invaded in a given year chosen from 1,…,t. If we assume that , then we set for all . This proposal is accepted using the ratio: