

APPENDIX. Overview of the expectation maximization (EM) algorithm used to update conditional probabilities.

The EM algorithm (Dempster et al. 1977) updates conditional probability values from case data by integrating over missing variables and their values. For example, consider the probability P of habitat variable x_1 and population variable x_2 . The joint probability of habitat and population occurrence is a product of the conditional probability of the population given the habitat conditions, and the unconditional probability of the habitat conditions, or $P(x_2, x_1) = P(x_2 | x_1)P(x_1)$. Further, define the probability of absence of habitat as $P(x_1 = 0) = \alpha$, the probability of a zero population given absence of habitat as $P(x_2 = 0 | x_1 = 0) = \beta_1$, and the probability of a zero population given presence of habitat as $P(x_2 = 0 | x_1 = 1) = \beta_2$. Then, the overall probability of occurrence of habitat and

population is the joint function $L = \prod_{i=1}^m P(x_2 = x_{2i}, x_1 = x_{1i})$, and the maximum likelihood function is $\max(L) \forall \alpha, \beta_1, \beta_2$.

The EM algorithm is useful when a specific condition x_L is not observed so that the joint probability distribution $P(x_L, x_1)$ cannot be explicitly calculated. Instead, it can be estimated by integrating out x_2 over the domain (0,1) using $\int P(x_2, x_1) dx_2 = P(x_1)$. This yields the marginal (unconditional) probability distribution over x_1 , in other words, the marginal likelihood calculation. This calculation is used as the basis for updating the joint (and thus conditional) probabilities in the Bayesian network.