Testing models of bee foraging behavior through the analysis of pollen loads and floral density data

Philippe Marchand [a,*], Alexandra N. Harmon-Threatt [b], Ignacio Chapela [a]

[a] Department of Environmental Science, Policy and Management, University of California, 130 Mulford Hall, Berkeley, CA 94720-3114, USA
[b] Department of Entomology, University of Illinois at Urbana-Champaign, 505 South Goodwin Ave., Urbana, IL 61801, USA

ABSTRACT

The composition of social bees' corbicular pollen loads contains information about both the bees' foraging behavior and the surrounding floral landscape. There have been, however, few attempts to integrate pollen composition and floral landscape to test hypotheses about foraging behavior. Here, we present an individual-based model that generates the species composition of pollen loads given a foraging model and a spatial distribution of floral resources. We apply this model to an existing dataset of inflorescence counts and bumble bee pollen loads sampled at different field sites in California. For two out of three sites, a foraging model consisting in correlated random walks with constant preferences for each plant species provides a plausible fit for the observed distribution of pollen load content. Pollen load compositions at the third site could be explained by an extension of the model, where different preferences apply to the choice of an initial foraging patch and subsequent foraging steps. Since this model describes the expected level of pollen load differentiation due solely to the spatial clustering of conspecific plants, it provides a null hypothesis against which more complex descriptions of behavior (e.g. flower constancy) can be tested.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Along with nectar, pollen constitutes a major food source for social bee colonies. For example, a typical colony of honey bees (*Apis mellifera*) consumes 40 kg of pollen a year, requiring millions of foraging trips (Stanley and Linskens, 1974). When foraging, both honey bees and bumble bees (*Bombus* spp.) accumulate pollen in the corbicula (or pollen baskets) located on their hind legs. By determining the composition of these pollen loads, which can be sampled non-destructively by capturing bees in the field or installing traps at the entrance of the hive or nest, pollination ecologists can obtain a record of the foragers' visits to different floral types.

Bee pollen analysis has been used to estimate the foraging range of a colony (Beil et al., 2008), compare vegetation profiles between widely separated locations (Diaz-Losada et al., 1998), study changes in resource use by one or more colonies over time (Aronne et al., 2012; Boff et al., 2011) and predict the amount of pollen flow between transgenic and conventional crop varieties (Ramsay et al., 2003). Most published studies report aggregated results – i.e. the average proportion of pollen sources at the hive level – and until recently, there were few attempts to model the variability in composition between pollen loads (de Valpine and Harmon-Threatt, 2013).

Although it may simplify the determination of colony-level components of foraging behavior, the aggregation of pollen loads obscures other aspects of foraging – such as the area covered in a single bout or the level of flower constancy shown by pollinators – that could be investigated by analyzing pollen samples from individual bees. Extracting this information from a given case study requires, in addition to the compositional pollen data, an estimate of the spatial distribution of visited plants as well as a modeling framework that links plant distribution, foraging behavior and pollen load composition.

In this study, we establish such a link – between floral landscape, foraging behavior and pollen load composition – through an individual-based modeling approach. Using estimates of the spatial distribution of each pollen source in the field, our model simulates foraging paths by following stochastic rules describing the bees' movement and floral preferences. Its output is a distribution of pollen load compositions corresponding to the specific field configuration and foraging parameters.

We present a simple parametrization of the model, where foraging paths consist of correlated random walks and floral preferences

Table 1
Main variables and parameters of the bee foraging model.

Model inputs	
$\mathbf{a}_{\text{init}}$	Vector representing the relative attractiveness of each plant species when selecting the initial foraging position
$\mathbf{a}_{\text{succ}}$	Vector representing the relative attractiveness of each plants species for successive foraging steps
$\mathbf{d}(x,y)$	Vector of densities (inflorescences m^{-2}) for each of the plant species at position (x,y) in the field
n_s	Number of steps (inflorescences visited) in a foraging bout
l_s	Root mean square (RMS) length for a step (dimensionless[a])
f_j	Frequency of jumps (occasional large steps)
l_j	RMS length for a jump (dimensionless[a])
ρ	Mean cosine of the wrapped Cauchy distribution; represents the directional correlation between successive steps
Output and summary statistics	
$\mathbf{P}$	Pollen load composition matrix: P_{ij} is the proportion of species j in pollen load i
$\mathbf{p}$	Vector representing the average proportion of each species (over all pollen loads)
F_{ST}	Differentiation in composition between pollen loads (see definition in text)

[a] When determining the RMS step or jump length, l_s and l_j are multiplied by the inverse square root of the local inflorescence density (see text).

are summarized in fixed vectors describing the relative attractiveness of each species. We apply the resulting model to a data set that includes pollen counts from bumble bees (*Bombus vosnesenskii*) foraging at different sites in Northern California as well as inflorescence counts sampled on a quadrat grid at each site. We demonstrate through a sensitivity analysis that the effective parameter space of the movement model can be reduced as many parameters have redundant effects on the model output. Finally, we use approximate Bayesian computation (ABC) to estimate the model's parameters based on a comparison of simulated and observed summary statistics—specifically, the average prevalence of each species' pollen and the compositional differentiation between individual pollen loads.

2. Materials and methods

2.1. Model specification

The model description in this section follows the ODD (Overview, Design concepts, Details) protocol (Grimm et al., 2006, 2010).

2.1.1. Purpose
Our model aims to predict the composition of bee pollen loads given the spatial distribution of pollen sources and parameters of bee foraging behavior (shape of foraging paths and preferences for certain floral species); conversely, the model would serve to infer parameters of foraging behavior from the observed composition of pollen loads sampled from a known floral landscape.

2.1.2. Entities, state variables and scales
The spatial distribution of floral resources is described as a density field $\mathbf{d}$, where $d_j(x,y)$ is the density (inflorescences m^{-2}) of species j at the field coordinates (x,y). Inflorescences (clusters of flowers arranged on a stem) are used as the basic floral unit since their abundance was found to be a better predictor of resource use for the different sites studied here (Harmon-Threatt, unpublished data). From empirical estimates of $\mathbf{d}$ on a rectangular grid of n_x by n_y grid points with spacings of Δ_x and Δ_y, the model interpolates the density field at any point within the grid (see interpolation submodel below).

Each run of the model simulates the activity of n_b foraging bees based on two sets of parameters, which are the same for all bees in a given run. The first set of parameters is used in modeling foraging paths as correlated random walks (CRW) and includes: the number of inflorescences visited on the path (n_s); the root mean square (RMS) value of the step length (l_s); the mean cosine of the turning angle distribution (ρ); the frequency (f_j) and RMS length (l_j) of occasional larger steps or "jumps" in the path. Details of the CRW model are presented in the submodels section below.

The second set of parameters describes the relative attractiveness of each species to foragers and includes two vectors, $\mathbf{a}_{\text{init}}$ and $\mathbf{a}_{\text{succ}}$: the former affects bee preferences for the starting point of the foraging bout while the latter affects bee preferences for successive flower visits. This distinction is motivated by the observation that bees may be attracted to a patch by a specific floral type, but will also visit less preferred flowers located in the same patch (Seifan et al., 2014). Therefore, we expected $\mathbf{a}_{\text{succ}}$ to show either the same or more even (less discriminative) preferences than $\mathbf{a}_{\text{init}}$.

During the simulation, the state of a bee is represented by its current position (x,y), its current direction θ (based on a line from the last to the current position), the plant species visited at the current position as well as the total number of steps taken to this point. The output of a model run is the composition matrix $\mathbf{P}$ where P_{ij} represents the proportion of the jth floral species in the ith bee's pollen load. A summary of the parameters and outputs of the model is presented in Table 1.

We note that while all positions and distances in the field are expressed in meters, the foraging parameters l_s and l_j are dimensionless. As we explain in the CRW submodel description, the actual step size is adjusted dynamically based on the local inflorescence density.

2.1.3. Process overview and scheduling
The model simulates each forager's activity as an independent realization of the following stochastic algorithm:

- The bee selects one of the grid points as the first position on its foraging path. Its initial direction (θ, in radians) is selected uniformly over $(-\pi, \pi)$. The probability of starting from a given point is proportional to a weighted sum of the densities of all species present at that point, with weights given by $\mathbf{a}_{\text{init}}$. That is, using j as an index over species and k as an index over grid points:

$$Pr(k) = \frac{\sum_j a_{\text{init}(j)} d_j(x_k, y_k)}{\sum_k \sum_j a_{\text{init}(j)} d_j(x_k, y_k)}. \tag{1}$$

Foraging paths are restricted to start on one of the grid points because the interpolation method does not produce a simple analytical form of the density field. Thus sampling the initial position in continuous space would be impractical.
- Starting from that point, the bee's path is simulated using the CRW submodel.
- The inflorescence density at each point in the path, which serves both to scale the CRW step size and to determine the probability of visiting each species, is calculated by the interpolation submodel.