Route selection on inclines: Dynamics of foraging trails in leaf-cutter ants

Nihav Dhawale

Reproduced here are excerpts from the scientific paper that describe the analysis that is performed in the accompanying code.

1 Path following analysis

To characterize the development of a foraging trail, we draw on the temporal nature of our dataset in order to determine how ant trails early in the trial affect the final established foraging trails. First, we develop an algorithm to define the foraging trail(s) at a given point in time in the trial using cumulative trajectory data. The board surface is discretized into cells of 5 mm each and trajectory data up to time τ in the trial is used to determine the number of ant trajectories that pass through each cell. This 2D histogram is then normalized so that maximum cell count equals 1. Thus cells have a value of 0 if no ants in the trial (thus far) have passed through it, and a value of 1 if every ant in the trial (thus far) has passed through the cell. Since neighbouring cells can have widely different counts depending on the arbitrary choice of cell size chosen, we apply an exponential convolution filter to smooth this histogram, obtaining a 2D occupancy density function that is designed to mimic the pheromone density map which ants used to direct other foragers towards potential food sources. The foraging trails are then simply a result of joining local peaks of this 2D function. The smoothing filter applied renders the location of the peaks relatively insensitive to cell size parameters. section 1.1 discusses the parameters used for the smoothing convolution and for peak detection. A trial may have a single established foraging trail or multiple trails simultaneously. Trails may also disappear, and new ones appear over the course of the trial. The time-resolved nature of our data allows us to follow these dynamics by altering the trial time window τ.

Next, we develop a metric that measures the 'distance' between an ant trajectory late in the trial and an established foraging trail earlier in the trial. As described in Fig. **??**b, the mean of the distances between every point on the ant trajectory at time $t > \tau$ and the corresponding nearest point on the foraging trail established at time τ defines a distance metric,

$$D_{traj}(\tau, t) = \frac{\Sigma_{y=0}^{y=y_{max}} d(y)}{\Sigma_{y=0}^{y=y_{max}} 1}. \tag{1}$$

This a measure of how closely the ant trajectory at time t follows the established trail at time τ. In order to aggregate information from multiple trajectories, we define distance metrics $S_{med}(\tau) = \text{median}(D_{traj}(\tau, t), \forall t > \tau)$ and $S_{iqr}(\tau) = \text{IQR}(D_{traj}(\tau, t), \forall t > \tau)$ to measure the median distance between all future trajectories and an established foraging trail at time τ as well as the interquartile range (IQR) of distances between all future trajectories and an established foraging trail at time τ.

1.1 Parameters for path following analysis

Occupancy histograms are smoothened using a moving 2D Gaussian filter with a window size of $w = 5$ cells,

$$G_{i,j} = Ae^{-((i-c)^2+(j-c)^2)/(2\sigma)}, \tag{2}$$

where i, j refer to cells along the x, y directions within the window of the filter (hence i, j take integer values from $1 \to w$). The window size w was heuristically arrived at although varying it up to $w = 11$ cells

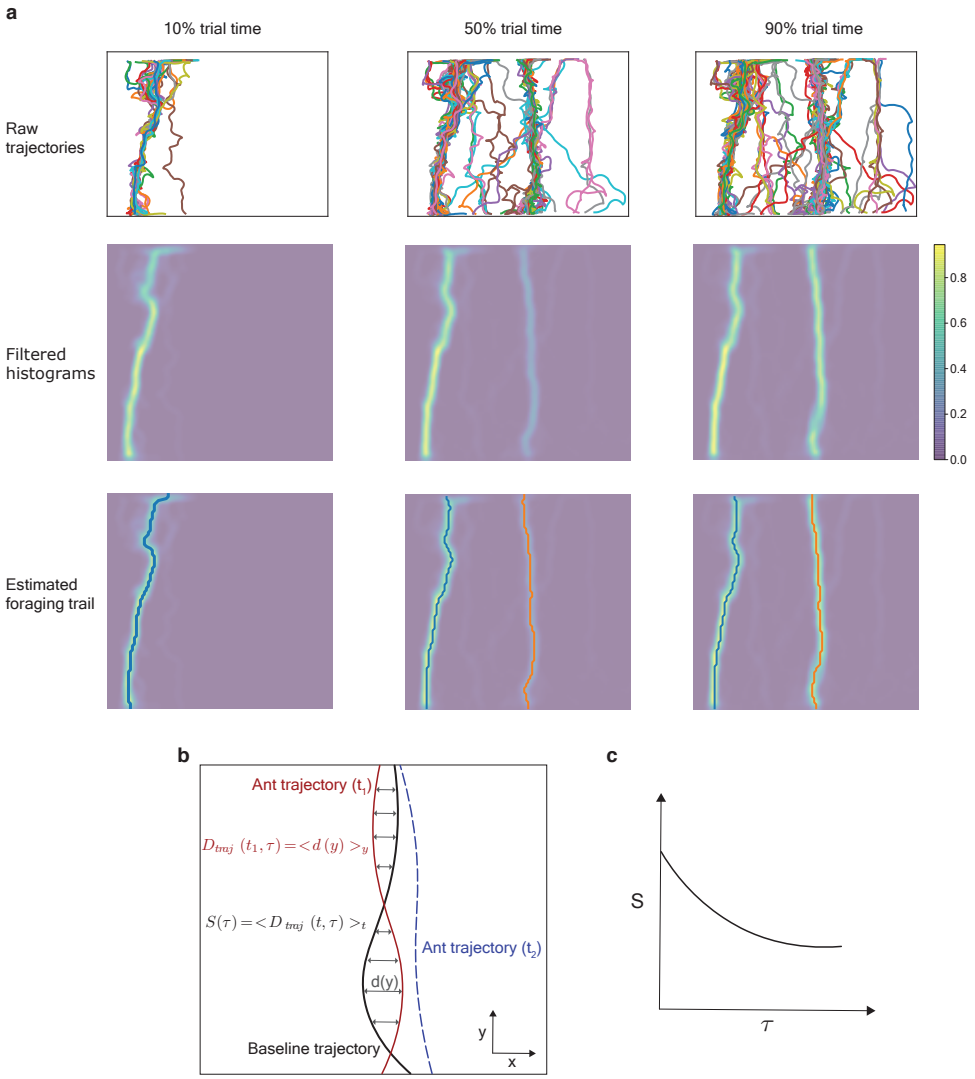

Figure 1: *Path following analysis.* To determine how ant trails early in the trial influence the established foraging route at the end of the trial, we develop methods to **a**, define an 'average' or baseline foraging trail using time-resolved ant trajectory data and **b,c**, develop a scalar metric to quantify how closely an ant trajectory follows the estimated baseline trail. **a**, Using raw trajectory data at various times within a trial (first row), we calculate a normalized spatial histogram of ant occupancy on the board. Convolving the spatial histogram with a 2D Gaussian filter we obtain a scaled-filtered histogram (2nd row) that is designed to mimic the pheromone density on the board at any given point in the trial. Joining the peaks of this scaled histogram (i.e. the estimated line of highest local pheromone density) we obtain the baseline foraging trail(s) (3rd row, solid line(s)). **b**, Using this, we calculate the mean point-wise distance $D(\tau, t)$ between an individual trail at time t and its closest estimated baseline foraging route from earlier time τ. The distance score metric $S_{med}(\tau)$ is then defined as the median $D_{traj}(\tau, t)$ of all trajectories that start after time τ. We also calculate the interquartile range (IQR) of $D_{traj}(\tau, t)$ from all trajectories that start after time τ as a metric of distance variability, $S_{iqr}(\tau)$. **c**, A hypothetical shape of distance score metric $S(\tau)$ (median and IQR) as a function of trial time τ shows that baseline foraging trails established very early in the trial have relatively less influence on ant trajectories later in the trial (high $S(\tau)$ for lower values of τ). When the distance score metric levels out, it implies that the established trail(s) are relatively unchanged over the remainder of the trial time and that ant trajectories are clustered around these established trails.

(odd numbers to preserve the symmetry of the Gaussian window) did not significantly change the results. Parameters A, c, σ are always tied to the window size with $A = 1/w$, $c = (w-1)/2$, $\sigma = w/2$.

From the scaled and smoothened occupancy histograms, peak detection is done based on parameters such as absolute peak height (p_{ht}); peak prominence (p_{prm}), i.e. peak height relative to height of the peak's lowest contour line; and minimum distance (in of units of cell number) between peaks along the width of the board (x-direction) d_{pk}. The value of these parameters are varied heuristically till the estimated foraging trails are visually matched to the density fluctuations of the scaled and smoothened occupancy histograms. Other parameters in the peak finding algorithm relate to the maximum allowed overlap between two distinct foraging trails m_{ovl}, used in determining when a foraging trail branches. This parameter, expressed as a fraction of trail length, balances the detection of a truly branched trail versus misidentification of loops that initially branch and rejoin the same trail, and may falsely register as trail branch. Finally, a 'saccade' distance d_{sac} is used to account for localized loss of data resulting in a gap between peaks. The parameter d_{sac} sets the maximum distance between peaks that the algorithm can determine is a continuation of the same foraging trail and linearly interpolate between. The value of this parameter balances identifying continuations of trails in the face of true localised loss in data versus misidentification of disparate trails or density hotspots as a single trail.

The values of all parameters are tuned on a trial-by-trial basis but kept constant within a trial. Table 1 shows ranges of the parameters used for peak detection across all trials.

p_{ht} (a.u.)	p_{prm} (a.u.)	d_{pk} (units of cell num.)	m_{ovl}	d_{sac} (units of cell num.)
0.09-0.12	0.06-0.09	10-20	0.3-0.7	15-25

Table 1: Ranges of parameter values used for peak detection.