Honey bees vary collective decision making in different landscapes

# Abstract

Honeybee foraging is an extraordinary collective behaviour which emerges from one of the most sophisticated communication systems in the animal kingdom: the waggle dance. Through a series of feedback loops, this elaborate behaviour -unparalleled in even a single other social insect- allows colonies to decide where to forage without requiring any individual to directly compare resources. This behaviour is a textbook example of an emergent collective process, yet a growing body of evidence now shows that using the dance does not always improve foraging success and sometimes not having access to the dance even increases it. Although the value of dancing varies across environments, it is not known if and to what extent colonies can adjust dance use in response. Here, develop a method to quantify dance use and show that colonies vary the use of the waggle dance in different landscapes, with some colonies relying on dance use almost exclusively and others not at all. Further probing of the drivers of this variation in dance use found a colonies use of the dance correlates with different landscape characteristics, suggesting colonies adjust collective decision-making in response to environmental conditions. Our results provide new evidence of flexible dance use in honey bee colonies and our methodology outlines a means to identify the ecological conditions in which honeybee colonies rely on dance communication, opening the door to large-scale experimental exploration of the selection pressures that may have driven the evolution of this remarkable collective behaviour.

# Introduction

In group living animals, many decisions are taken collectively by integrating information from multiple individuals to produce a social behaviour which extends beyond that of the individual units (*1*). Such systems are self-organised, relying on the constituents using socially provided information and following behavioural ‘rules’ to produce an emergent collective outcome (*2*, *3*). The behavioural architectures which produce these emergent collective behaviours have in some cases, such as in social insects, grown to dazzling levels of complexity. However, this complexity is not always beneficial. The benefits of working collaboratively versus independently vary across environmental contexts and individuals often adjust their own use of information accordingly (*4*, *5*). Although the benefits of collective behaviours are known to vary across ecological conditions, in many cases we have a more limited understanding of if and how a group can flexibly use collective information to make use of the most profitable strategy under different conditions.

The extraordinary waggle dance communication system of the honey bee (*Apis* spp.) is a key example of a behavioural mechanism which produces emergent collective decisions (*6*). Within colonies, a series of simple rules which determine when and how much bees perform dances (Fig. 1) mean that choices between feeding sites are made by the group rather than individuals (*7*–*9*). For example, because the number of dance circuits performed by a bee on returning from a food source reflects the net energetic benefits of the trip, more of the colony’s workforce will be recruited to the richer of two equidistant sources (*7*), or the closer of two equally rich sources (*9*). This mechanism allows a colony to pool individual assessments of the landscape and processes them collectively to identify the most profitable resources to focus on, all without the need for any individual to compare resources.



Figure 1. **The honeybee waggle dance carries information about the location of a resource.** The duration of the waggle run indicates the distance to the resource and the angle of the dance relative to the vertical indicates the direction of the resource (circle in right panel), relative to the direction of the sun (*10*). Through the observation and decoding of the waggle dance, a colony’s dance floor provides a unique opportunity to eavesdrop on the communication and decision making leading to collective foraging decisions. Overall resource “quality” -the net energetic gain of a foraging trip- is provided through the number of waggle runs performed (*10*–*14*). Although bees that follow dances do not specifically interpret this information on an individual level (*7*), the resulting over-representation of high-quality sites on the dance-floor means that they are more likely to encounter dances that advertise better forage (*6*), and provides the colony with a mechanism to select the most profitable resources in their environment (*8*).

Despite it’s complexity, research has shown that the ***value*** the waggle dance provides to foraging success varies according to environmental context and colonies without access to dance information achieve equal or even greater foraging success than their wild-type counterparts (*15*–*18*). Interestingly, studies have shown that individuals often ignore the dance when looking for forage sites (*19*). Whilst it likely that individual variation in dance use translates into variations in collective decision making at the colony level (*18*, *20*), this has yet to be empirically established. Identifying if and how colonies vary waggle dance ***use*** is not only important for deciphering how honey bees collectively respond to changes in environmental conditions (*21*), relating this to the environmental context of the behaviour is key to illuminating its ultimate evolutionary drivers (*22*).

Quantifying dance use within colonies requires measuring dance use across the foraging workforce (*6*). This has only previously been quantifiable by documenting foragers’ individual search history, which requires tracking individual bees and is time -and labour- intensive (*6*). Individual based models, which explore the collective effects of individuals following different behaviours, provide an insight into how individual information use scales to the level of the group (*5*, *23*), however, these are difficult to relate to empirical observations and so are limited to showing what a colony under certain conditions might do, rather than quantifying what they are doing. Empirically determining how dance information is used at a colony level requires new methods which can identify the behaviours of individual foragers and integrate them with the mechanisms by which the dance functions to make collective decisions.

Here, we derive a mathematical model to quantify waggle dance use in colony foraging by identifying the proportion of foraging trips that are made by bees recruited through waggle dances. To validate our model, we fit it to waggle dance observations collected from 20 sites across and agri-rural and urban environment (10 in each). We then use our model to explore how variations in landscape structure influences dance use in colony foraging.

# Results

The patterns of waggle dance-encoded foraging distances might differ under scenarios where colonies rely on individual search versus recruitment. To test this, we simulated honeybees foraging in a landscape where resource patches were randomly placed in the environment. Foragers could locate these under two different strategies: either acting as a scout and locating resources themselves, or following a recruit strategy and locating resources by following a random dance from the dance floor (*24*) (Fig. 2a,b, details in Materials and Methods). As it is known in the simulation which individuals in the hive forage under what strategy, we can compare the distributions of foraging distances reported on the dance floor by each type of forager. Fig 2c shows that the shapes of the resource distance distributions for bees engaging in the two types of foraging trips are different. The distance distribution for the scout trips is akin close to that of an exponential distribution (Fig. 2c), which is the nearest neighbour distance distribution for foragers operating in a one-dimensional environment (see Materials and Methods). The distribution of the distances reported for recruit trips (Fig. 2c) is a Rayleigh distribution which is the nearest-neighbour distribution in a two-dimensional environment (*25*) (see Materials and Methods).

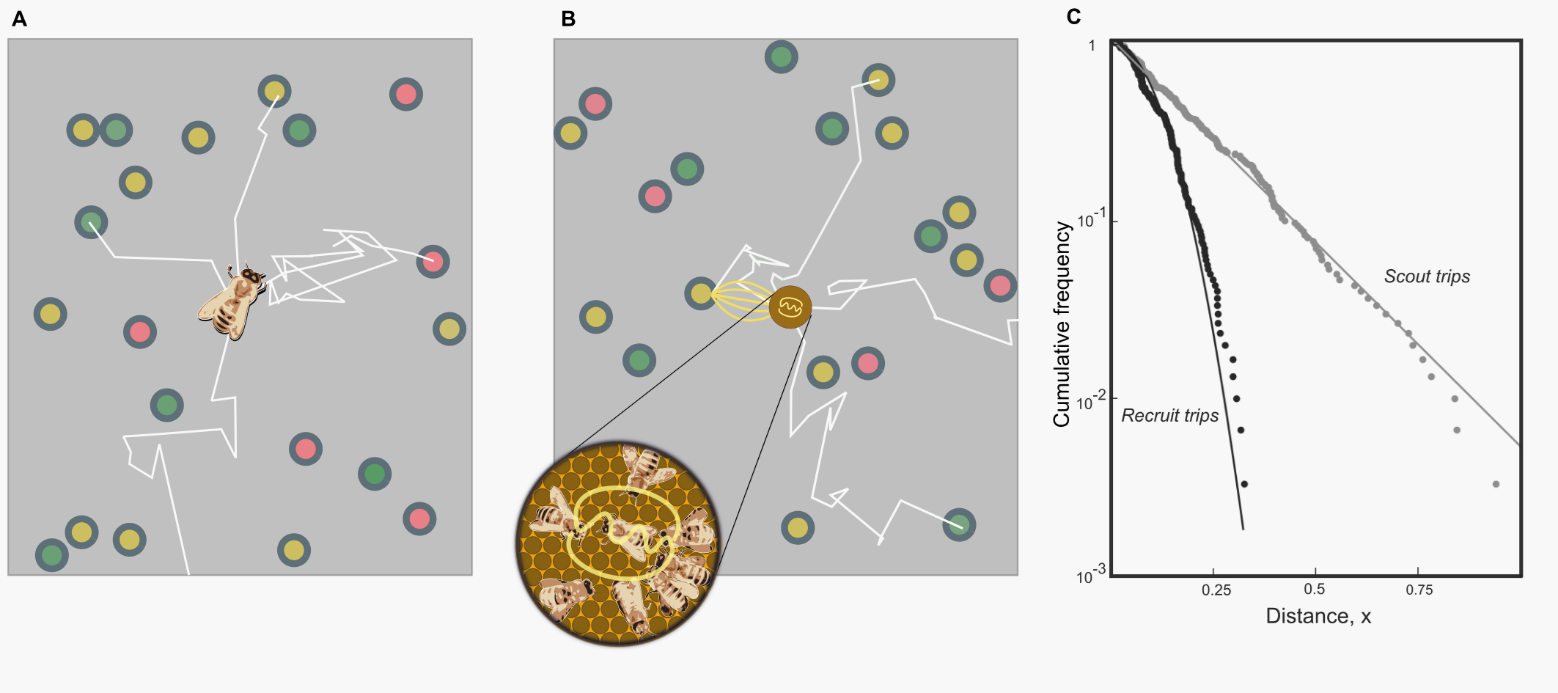


Figure 2. **Simulating honey bee foraging**. In our simulation model with scouting only (A), foragers leave the hive on a search path (white lines) and continue until they encounter a resource (circles, colours indicate different resource quality). When foraging with recruitment (B) foragers continue to identify resources in scouting trips (white lines) and convey this information on the dance floor (brown disc) where foragers can sample dances reporting on scouting and recruiting trips and follow these directions (yellow lines). (C) Complementary cumulative frequencies of foraging distances reported from scouting and recruit trips. Note the difference in the shape of the distributions. The scout distribution is best fit by an exponential (grey fit line), the recruit distribution by a Rayleigh distribution (black line)

On the dance-floor, the number of waggle runs performed for a resource depends on its profitability (*13*). Honeybees achieve this by measuring the energetic efficiency of a foraging trip through the ratio of energetic gain to energetic cost (*8*, *13*). By combining this profitability bias with the distributions identified in our simulations, we can accurately describe the distribution of waggle runs reported on the dance-floors of real honeybee colonies as a superposition of scout and recruit distributions (Fig 3, see Materials and Methods for details)

This description intrinsically captures the collective decision making underlying honeybee foraging, whereby the foraging sites represented on the dance floor derive from a mixture of individual search and waggle dance information, modified by the profitability rule that biases recruitment towards closer or richer patches. The extent to which collective decision making is used is expressed in the proportion of scout dances, , and the proportion of recruit dances, . Scouting and recruiting are not fixed behavioural categories, because individual bees can engage in both over the course of their foraging lifetime, and foragers can dance on return from any successful trip irrespective of whether they were recruited to the forage site or found it individually (*26*).

We constructed two models of foraging: an individual model describing a hypothetical colony that relies only on individual search to find foraging sites, and a collective model which describes a colony using waggle dance communication to exploit resources. In the individual model, all trips are “scout” trips, but as recruitment becomes more important, the proportion of “recruit” trips will increase. By setting , we obtain a model based on the sole use of individual search (scouts only) and, alternatively, by allowing the proportion of scout trips, , to take on any value between 0 and 1, we can model the extent to which foraging occurs collectively. (Fig 2). In fitting these two models to distribution of waggle run durations decoded from real honeybee colonies we are thus able, using model selection (*27*), to infer if and, by estimating the parameter , to what extent honeybee colonies use waggle dance information when foraging.

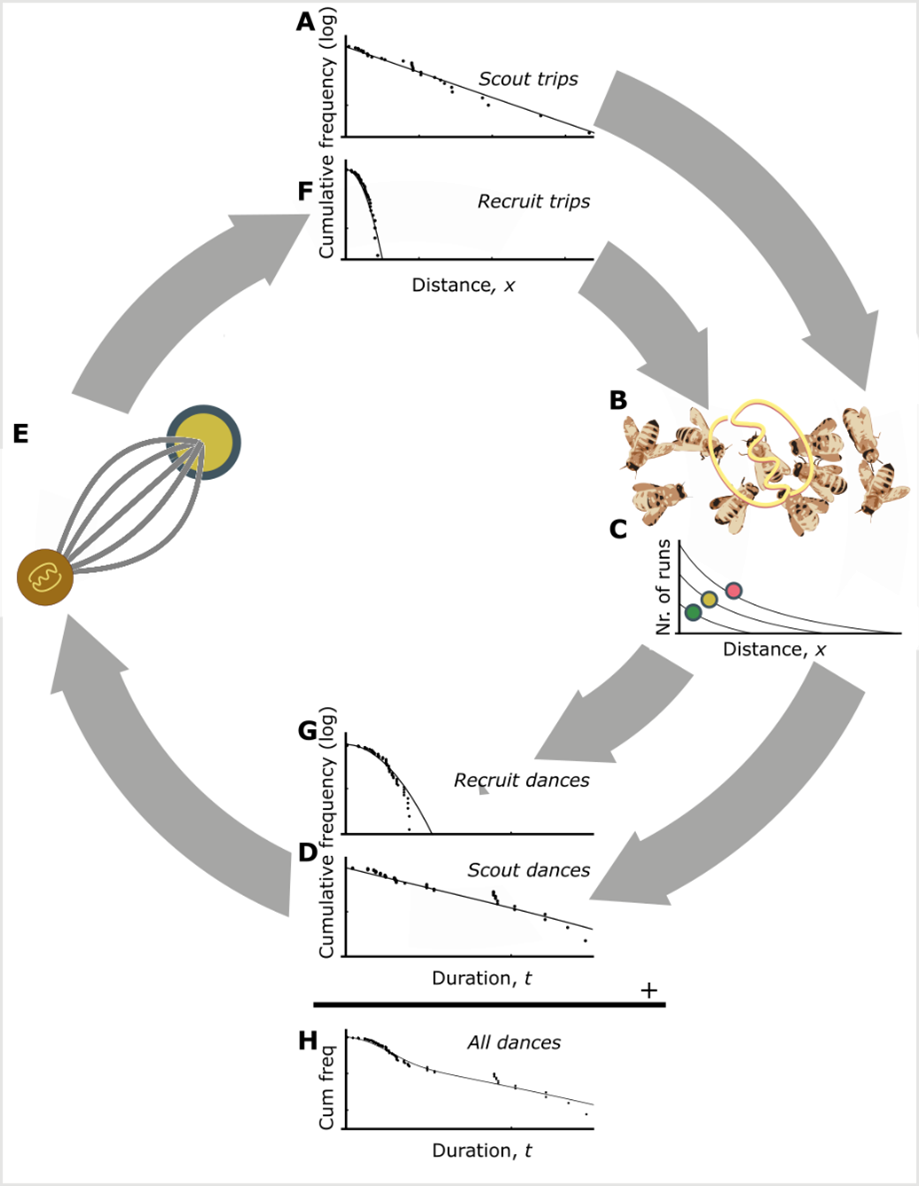


Figure 3. **The rationale of the foraging model**. The distances of resources encountered by scouts are distributed exponentially (A). These dances are advertised on the dance floor (B). Dances for resources that are closer or higher in quality are repeated more often (C). As a consequence, dances for more profitable resource are over-represented and sampling foragers are biased to the more profitable resources (D). After successfully visiting advertised resources, recruits also dance for them leading to further amplification of this bias towards the most profitable resource in the vicinity of the hive (E). The distances of recruiting trips are than distributed through a Rayleigh distribution (F). Recruits also report the locations on the dance floor (B) and repeat their runs more often depending on the profitability of the location (C), leading to a distribution of durations of recruit dances (G). By taking together the dance distributions for scouts (D) and recruits (G) the distributions of all dances on the dance-floor can be found. The distances reported on the dance floor this are a mixture of the scout and recruiting trips and can be calculated from the distance distributions of the scouting and recruiting trips, taking the reporting bias into account (see Materials and Methods for detail).

To evaluate the use of waggle dance information and individual search in honeybee colonies foraging in ‘natural’ landscapes, we analysed a pre-existing dataset of 2827 waggle dance observations from 20 observation hives, recorded between April-September 2017, (previously described in (*28*)). Hives that contributed to this dataset had been situated at different locations in South East England (see Materials and Methods, figure 4A) and visited every two weeks for a period of 24 weeks. On each visit, two hours of continuous waggle dance data was recorded by training a camcorder onto the dance floor. The footage of the dances was decoded manually (*28*–*30*) to extract waggle run durations. Using this data, for each site we fitted both the collective and individual models and used model selection to determine which provided the better explanation of the data, and (if the collective model provided a better fit) to quantify the relative use of social information through estimating the parameter . In each case, we calculated the goodness-of-fit using a Kolmogorov-Smirnov (KS) test to ascertain if the model provided a plausible explanation of the data (*31*, *32*).

For 16 out of 20 study hives, the collective model provided a better description of the data than the individual model (Fig. 4A). In the other 4 sites, despite the collective model having the higher maximum likelihood, the individual model had a higher AIC value and so is more parsimonious (Supplementary Table 1). In all but one site, the collective model had a good fit (using a Kolmogorov-Smirnov statistic of , see Materials and Methods) to the empirical waggle run durations (Fig. 4B), whereas the individual model was significantly different to the observed data in 8 sites (Kolmogorov-Smirnov statistic , Fig. 4b). The sites shown in Figs 4C-D are representative examples showing the model fits where the individual (Fig. 4C) and collective (Fig. 4D) models fit best. Note the closeness of the fit to the data, illustrating the overall quality of the model description.

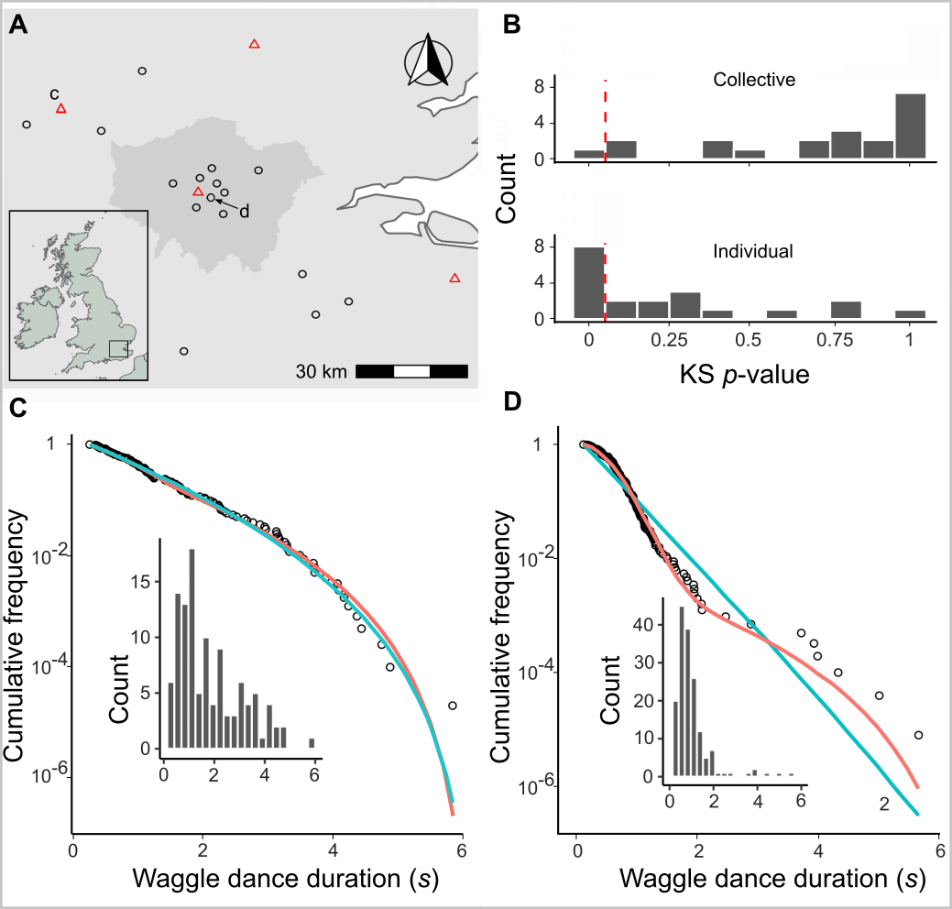


Figure 4. **The honey bee foraging model fitted to data from 20 hives**. (A) Location of study hives in Southern England, shaded area in the main plot indicates Greater London. For 16 hives for the collective foraging model provided best explanation (black circles) or for 4 hives the individual search model provided the best explanation (red triangles) as indicated by lowest AIC score. (B). Distribution of goodness of fit confidence values for each model fit to waggle run durations from each site. The p-value is derived from a bootstrapped two-sided KS test comparing the fitted model predictions to the empirical data, the red dashed line marks the significance threshold of 0.05. For values exceeding the threshold there is no statistically significant difference between the model and the data, indicating the model provides a good fit. For the hive in (C) the individual model (blue line) provided a better fit than the collective foraging model (red line). For the hive in (D) the collective foraging model (red line) provided a better fit than the individual model (blue line). The typical “hump” in the distribution in (D) which is indicative of contribution of recruitment dances (compare to Fig 3H). Panels show the compliment cumulative frequencies with binned frequency distributions as inset.

Our results show that, whilst colony-level foraging is mostly comprised of a mixture of scout and recruit foraging trips, in some circumstances, colony foraging can be better described by individual foraging alone. Thus, in some environments, the majority of foraging trips involve scouting to find new food sites rather than recruitment through dances. Note that this does not imply that these bees do not engage in dance following, because bees regularly follow dances but choose not to visit the advertised site (*33*), although it has also been shown that bees may cease dance following if it is proving unproductive (*18*, *33*). Evaluating how these individual decisions influence the collective, however, has historically been a challenge as it is effectively impossible to track an individual’s foraging behaviour over a landscape. As our results show that our model of individual foraging provides a good fit to colony foraging in 11 different sites, and provides a more parsimonious description of foraging than a model of collective foraging in 4 sites, these findings support the idea that individual decision making can dominate colony foraging and demonstrates further evidence for flexible waggle dance use by honey bee colonies.

quantifying the use of waggle-dance recruitment within all colonies, as a proportion of all foraging trips, can be achieved by extracting the estimated proportion of recruit trips, , for each site. Since our sites varied in land-use characteristics, and thus potentially forage distributions (although in this case, not by design for this study), we investigated whether these estimates might correlate with land-use. We first classified the different land-use types of the area surrounding each site (*34*) to obtain a standardised land-use profile for the urban and agri-rural environments separately as many land-use types present in urban areas do not occur in agri-rural environments and vice versa (see Materials and Methods). We then performed a Partial Least Squares (PLS) analysis (*35*) (see Materials and Methods) to determine the principal components that represent combinations of land-use types which explained the most variation in waggle dance use within agri-rural and urban environments. As for one of our sites neither model provided a plausible description, this site was removed from the PLS analysis.

Waggle dance use varied widely regardless of environment, ranging from 0 to 1 in both the urban and agri-rural environments whilst showing a significant, positive interaction with waggle dance use. In the agri-rural sites, a single principle component explained ~73% of the variation in waggle dance use (beta regression: , = 4.9, p < 0.05, Fig. 5A). This finding is also shown in the urban environment which also shows a single principle component explained ~73% of the variation in waggle dance use (beta regression: = 0.73, = 10.4, p < 0.05, Fig. 5C). Probing the individual loadings of these principal components allows us to evaluate the driving mechanisms in more detail. As our sample size is limited (10 urban and 9 agri-rural sites), we used jackknife resampling to evaluate the robustness of our results to influential points (see Materials and Methods, Supplementary Material) and evaluate how these land-use types contribute to the principle component which significantly correlates with waggle dance use.

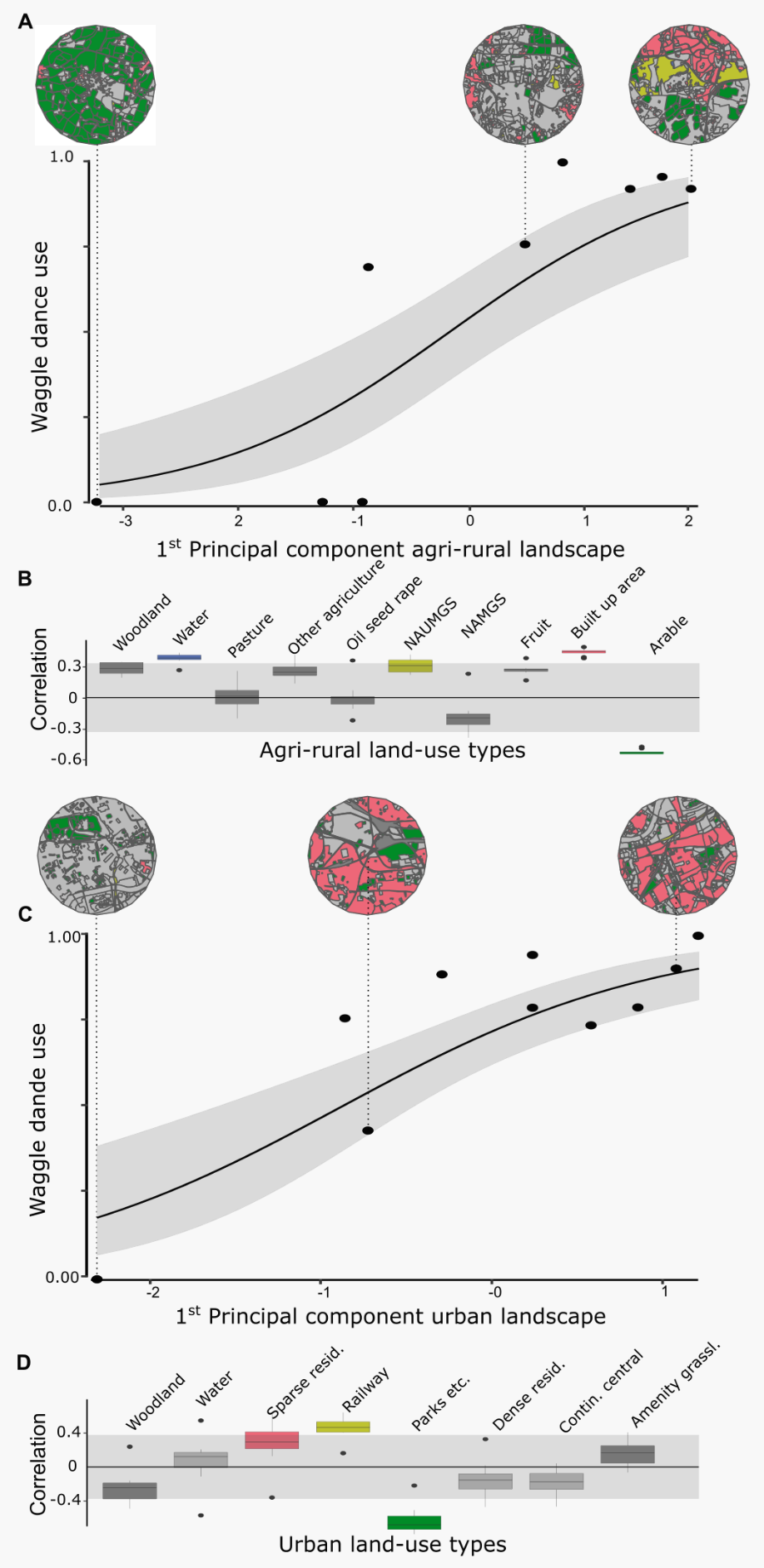


Figure 5. **Collective foraging correlates with land-use**. Estimated waggle dance use for each site against the first principal component derived from a Partial Least Squares analysis of land-use type. Beta regression shows the relationship (black line) between first principal component and waggle dance use, with 95% CI shown by the grey shaded area for agri-rural (A) and urban landscapes (C). The correlations between first principal component and each land-use type are shown for agri-rural (B) and urban landscapes (D). Correlations outside the shaded area significantly contribute to the first principal component. Colours correspond to the land use as shown in maps (circular insets) for selected sites. NAUMGS (resp. NAMGS) stands for non-agricultural unmanaged (resp. managed) green space

In the agri-rural environments the first principal component correlates positively with arable land (29% of land coverage; Table 1) and negatively with built-up areas (17% of land coverage); note that it also correlates negatively with non-agricultural unmanaged green space and water, but together these represent less than 3% of land-use (Table 1; Fig. 5B). These land-use types maintain a significant correlation with the first principal component over the jackknifed PLS (Fig. 5B), with the exception of non-agricultural unmanaged green space which sits on the boarder, indicating the results are robust. As arable land increases whilst built-up areas decrease, the proportion of trips that are driven by individual search increases. Arable land in the UK is typically considered nutritionally poor for bees (note that oilseed rape fields were not included within this category), while there is evidence to suggest that the residential areas that were captured within the “built-up” category are forage-rich hotspots, typically supporting relatively high bee diversity and abundance within gardens (*36*).

In the urban environment, our findings were more equivocal. The first principal component correlated negatively with the dominant land-use type by coverage, sparse residential (land coverage ~35%; Table 1), as would be expected since such land includes a high proportion of gardens. However, this relationship is very sensitive to site removal through the jackknife sampling (Supplementary Fig. 2.), and is not a significant contributor to the variance in the loadings (Fig. 5D.). The first principal component also correlates positively with railways but these constitute ~1% of land cover (Table 1). Furthermore, the first principal component correlates negatively with parks, allotments and cemeteries which would typically be considered forage-rich for honeybees (*36*) (although note that the amount of land cover for this component is limited and the jackknifed PLS identified substantial variation in the loadings identified for the first principal components; Supplementary Fig. 3.).

**Table** : Percentage area covered for each land-use type in the agri-rural and urban environments in the sites studied.

| **Environment** | **Land-use** | **% coverage** |
| --- | --- | --- |
| Agri-rural | Arable | 28.3 |
| Agri-rural | Pasture | 23.3 |
| Agri-rural | Woodland | 21.1 |
| Agri-rural | Built Up Area | 15.0 |
| Agri-rural | Fruit | 3.1 |
| Agri-rural | Oilseed Rape | 2.9 |
| Agri-rural | Non Agricultural Unmanaged Green Space | 2.8 |
| Agri-rural | Non Agricultural Managed Green Space | 1.8 |
| Agri-rural | Other Agricultural | 1.5 |
| Agri-rural | Water | 0.2 |
| Urban | Sparse Residential | 34.8 |
| Urban | Continuous Central | 24.3 |
| Urban | Dense Residential | 21.8 |
| Urban | Parks Allotments Cemeteries | 7.9 |
| Urban | Woodland | 4.2 |
| Urban | Water | 3.4 |
| Urban | Amenity Grassland | 2.6 |
| Urban | Railway | 1.0 |

# Discussion

Understanding how individual variation in information use scales to produce flexible collective processes at the group level is a key challenge in behavioural ecology. Here, we have presented a method to quantify information use at a collective level in a model social insect. After verifying that this method accurately captures empirical patterns of collective and individual foraging, we have shown how it can be used to evaluate how environmental factors shape collective foraging at the colony level. Whilst recruitment is well known to occur through other mechanisms besides the waggle dance, such as through olfactory cues (*37*), the exceptionally close fit of our model to waggle run durations underlines the importance of the waggle dance in honey bee foraging. Further examining our estimates of dance use show that a colonies use of the waggle dance varies across environments and does not appear to be randomly determined.

Within both the agri-rural and urban environments, dance use had a significant interaction with land-use type, ranging from total dance use to complete reliance on individual information. The notion that individual variation in information use influences the collective effect has only previously been inferred theoretically. Our findings demonstrate this empirically. Individual-based models (*5*, *18*, *23*) have outlined how variations at an individual level could drive group level dynamics, however, without verifying this upon empirical data the insights remain theoretical. Here, we have shown how an individual based model can be used to better understand a system in a way which allows us to build more informative models which can be empirically verified. Not only do our findings add support for our understanding of how the dance functions to make collective decisions (*6*), but they also provide a new dimension to our understanding of how variation at an individual level influences the collective level.

Whilst the benefits of dance communication are known to vary across environments (*18*, *20*, *33*, *38*), our results show that colonies vary their use in response, affording the colony a flexible dance use strategy which they can tailor to different conditions to maximise foraging success. Identifying the circumstances in which dance communication is important have been hampered by the requirement to monitor real-world colony weight (*15*, *16*, *23*, *39*, *40*) -a noisy proxy of foraging success that is influenced by many other abiotic and biotic factors, including colony health- over ecologically credible time periods (*38*). Doing this at a scale that allows inter-colony variation in foraging environments at sufficient replication is a major logistical hurdle (but see (*39*)), particularly given that multiple landscape variables may interact to determine the utility of dance communication. Our method circumvents this issue by providing a time and labour efficient methodology to quantify this key collective behaviour and map the environment along the major axis of honeybee information use.

Examining the land-use types which contribute most to the variation in waggle dance use in each of our study environments reveals insights into the drivers of waggle dance use. Our agri-rural results suggest a potential decrease in reliance on waggle dance recruitment as resources become harder to find and foraging trips become longer (*28*), broadly agreeing with the findings from other theoretical (*23*) (11) and empirical (*18*) studies which examined foraging performance with and without the dance. These results suggest that, even though collective foraging is not always beneficial, in environments where high-quality resources are present but relatively scarce, collective foraging is being carried out by colonies and under these conditions exchanging social information through the waggle dance confers benefits. In both agri-rural and urban landscapes we have shown that the waggle dance use changes with land-use (illustrated in the change in land-use shown in the maps from left to right in Fig 5A and 5C). These findings illustrate how our model can be used to estimate reliance on collective behaviour within different landscapes, for exploration within datasets where land-use is systematically chosen by design.

Although the number of colonies used in this analysis are comparable to similar studies evaluating collective foraging in honey bees (*15*, *16*, *18*, *39*, *40*), probing the landscape drivers of colony level waggle dance use in more detail requires more colony level data. As longer term effects, such as resource stability (*38*), may also influence waggle dance use, these data need to be gathered over sufficient time frames to ensure these effects are accounted for. Recently, technological advances have emerged which enable colony metrics to be collected faster, more accurately and over greater time spans than could be gathered by hand, allowing individuals to be tracked within colonies and theories of individual behaviour to be evaluated in more depth than could have been done previously (*41*). Here, we have presented an accurate mathematical description of colony foraging which extends our ability to quantify collective behaviour across environments. With the advances in the decoding of the waggle dance through automated methods (*42*), we face the prospect of waggle dance data becoming “big data”. Our methodology thus provides a means of analysing such large data sets and provide a framework to start exploring how plasticity in collective decision making can help to buffer the effects of landscape change, and in this case, how flexibility in the use of the waggle dance may help to exploit resources when experiencing novel landscape mosaics, such as in highly urbanized conurbations.

# Materials and Methods

## Data collection

Details of data collection , waggle dance decoding and classification of land-use types can be found in full in the Materials and Methods section of (*28*).

## Simulation

All simulation code was written in Python version 3.9 and uses the Pandas (*43*) and Scipy (*44*) packages.

A circular environment is first created with radius . The number of resources in the environment is generated as a random Poisson variable with rate equal to 5000 multiplied by the area of the environment. These are placed on polar coordinates with a uniformly selected angle, , between 0 and 2 and a radial value, , between 0 and , determined from the square root of the uniform position values multiplied by . These polar coordinates are converted to Cartesian coordinates. Each location is then assigned to an instance of a resource object along with a random quality of between 0 and 10. This quality is combined with the distance of the resource to the centrally located hive to form a measure of how profitable the resource is (see model, equations).

One-hundred honeybee objects are created, 20 of which are on scouting trips and the rest recruited to follow scout dances. Scouts leave the environment following a random path through the environment generated by sampling a uniform random step length and angle. The number of paths the scout draws when searching is also determined as a uniform random number. Each straight line in the random path is converted to a rectangle with length equal to the path section length and a constant width of ~0.01 to represent an area the scout searches along that path. Of all the resources contained in the boxes drawn from the scout’s path, the one closest to the colony is selected as the resource patch that the scout will report and will communicate its location if the quality of the resource exceeds a minimum threshold. Communication is simulated by pooling together all the resource patches found. If no resources are contained in the scout’s path, they will not add any resources to the scout pool and draw a new path in the next foraging iteration.

Recruits represent honeybee objects which do not perform the searches the scouts do. Instead, they sample from the pool of resource objects reported by scouts. This sampling is done by selecting resources with a probability which is skewed towards the profitability of the resource, meaning more profitable resources have a greater chance of being selected by recruits. Recruits will then visit these resources and in the next iteration will add their resource to the pool of scout dances. Consequently, the pool of dances represents resources discovered by scouts and resources exploited by recruits. When a resource is depleted, it is removed from the environment and so any foragers that were foraging on it would select a different resource from the dance floor.

The simulation was run 100 times and every 5 time steps all distances reported by scouts and recruits were recorded and combined. This was done to reflect the way foraging data is collected in real honeybee studies. We fit an exponential and minimum of an exponential distribution to both the distribution of foraging distances reported by the scout and recruit objects. Fitting was done by deriving the maximum likelihood estimate for each model fit on each data source through their analytical solutions: , minimum of the exponential with a minimum foraging distance: . As the exponential assumes distributions start from 0 the data was transformed to start from 0 by subtracting the minimum foraging distance from all foraging distances () before fitting.

## Model

To describe the distribution of dance durations on the dance floor we formulated a generic model for the duration of waggle dances. In the model resource patches are assumed to be randomly placed in the environment. Foragers scout for these patches. The rationale of the model is illustrated in Fig. 3: upon visiting a resource patch, foragers translate the profitability of a resource into the number of repeats of the dance. The number of repeats of the dance is a function of quality and distance. Recruits sample random dances and report the location of successful visits to resource patches on the dance floor. Through the feedback and over-representation of profitable resources on the dance floor recruits will converge to visiting the most profitable resource in vicinity of the hive. The distribution of dance durations is the superposition of scouting and recruiting trips.

As the resources patches are randomly placed in the environment, the distance after which the first resource is discovered approximately follows an exponential distribution (given by ). Through the feedback mechanism that the dance floor provides, the colony can, collectively, locate the most profitable resource in its environment. For randomly placed resources in a two dimensional environment the distance to the nearest point is distributed according to a Rayleigh distribution (given by )) (*25*). Our simulation model shows that this describes the distances at which recruits visit resources well. Knowing the distance distributions of scout and recruit trips we then assume that the a proportion of all trips are scout trips. With this information we can specify the distributions of distances on the dance floor (see Supplementary Material for details).

We implemented this in a full model which describes the distance distribution of the environment has different resource types (See Supplementary Material). However, in the full model the number of parameters increases with . Even if the number of resources is low, it turned out to be cumbersome to estimate the parameters and the model tends to overfit. To facilitate estimation of the parameter we therefore used a simplified model to estimate the fraction of scout trips, where represents the lowest duration considered, and here a minimum waggle run duration in the data set.

For the simplified the model we assumed that the number of dances depends weakly on distance and there is a sizable quality differences between resources of a non-negligible size and that there is a sizable intensity of the high-quality resource (See Supplementary Material for detailed derivation). Foragers on scouting trips are more likely to report larger distances than foragers on recruiting trips. By linearising the function that translates the profitability into the number of waggle dance run for the largest dance duration and normalising, we arrive at simplified distribution for dance durations for scouting trips:

where we used the shorthand The parameter is the maximum dance duration by scouts, is the intensity of resources found by scouts and is factor that normalises the distribution.

Recruit trips will be predominantly to high-quality resources. Only if the nearest high-quality patch is very far away will there be a more profitable patch of lesser quality available, and this happens only rarely if the intensity of the best quality resource is sizable. After linearising the function that translates the profitability into the number of waggle dance runs for short dance durations and normalising the distribution of dance durations reported from recruit trips in the simplified model is:

where

is the normalisation factor, the parameter is the rate with which dances repeats depends on distance for recruit trips and is the intensity of high-quality resources reported by recruited foragers.

The simplified distribution function is

which we used for parameter estimation and model fittings.

## Statistical analysis

All analysis code is written in R (*45*).

### Model fitting

All models are fit using Maximum likelihood estimation (*27*) by summation of the log of the simplified distribution function outlined in the methods section: model. The numerical optimisation routine is written in c++ and uses the Nelder-Mead simplex algorithm (*46*) implemented in the ‘NLopt’ library (*47*) and interfaced to R (*45*) using ‘Rcpp’ (*48*).

The most parsimonious model is assessed using Akaike information criterion (AIC) (*27*, *49*) and Akaike weights. The model with the lowest AIC score is deemed to be the most parsimonious.

Goodness of fit is assessed using the two-sample Kolmogorov-Smirnov (KS) test (*31*) and implemented in R using the ks.boot function of the package ‘Matching’ in R (*50*).

### Partial Least Squares analysis

Prior to conducing the PLS we removed any sites in which the models fit was significantly different to the waggle dance durations for that site. This resulted in one agri-rural site and no urban sites being removed from the analysis.

As our estimates of waggle dance use is continuous on the interval we used the R package plsRbeta (*51*) to conduct the partial least squares analysis and performed a beta regression on the results using the R package betareg (*52*). As the betareg package only works on the open interval the data, , was transformed using the following equation: , as outlined in the betareg package documentation. After analysis the data was back transformed to the original values for the plots in Fig 5.

For the jackknifed resampling we iterated through the each site and removed it from the pool of data and then ran the PLS as described above, recoding the loadings for each iteration (see Supplementary Material for loadings with each site removed). The PLS loadings for each land-use type are plotted as a box plot in Fig 5. to show the spread of these variable types. A loading was determined to be significantly correlating with the first principal component if contributed more than its expected variance.

# Acknowledgments

**Funding**

This work was supported by the Biotechnology and Biological Sciences Research Council (BBSRC) through grant BB/M011178/1.

**Author Contributions**

Conceptualization: JP, RG, EL, VJ. Methodology: JP, EL, VJ. Software: JP. Formal analysis: JP. Investigation: AS, EL. Writing - Original Draft: JP, VJ. Writing - review & editing: JP, RG, EL, VJ. Supervision: RG, EL, VJ.

**Competing interests**

Authors declare they have no competing interests.

**Data and materials availability**

All of the data and code required to reproduce this article is contained as within the ‘wagglefit’ which can be found at (link to finalised Zenodo archive finalised)

<https://github.com/joseph-palmer/wagglefit>

# References

1. D. J. Sumpter, The principles of collective animal behaviour. *Philosophical Transactions of the Royal Society B: Biological Sciences*. **361**, 5–22 (2006).

2. I. D. Couzin, J. Krause, Self-organization and collective behavior in vertebrates. *Advances in the Study of Behavior*. **32**, 1–75 (2003).

3. E. Bonabeau, G. Theraulaz, J. L. Deneubourg, S. Aron, S. Camazine, Self-organization in social insects. *Trends in Ecology and Evolution*. **12**, 188–193 (1997).

4. F.-X. Dechaume-Moncharmont, A. Dornhaus, A. I. Houston, J. M. McNamara, E. J. Collins, N. R. Franks, The hidden cost of information in collective foraging. *Proceedings of the Royal Society B: Biological Sciences*. **272**, 1689–1695 (2005).

5. M. Beekman, J. B. Lew, Foraging in honeybees - when does it pay to dance? *Behavioral Ecology*. **19**, 255–262 (2008).

6. T. Seeley, *The wisdom of the hive* (1995; <https://pdfs.semanticscholar.org/3230/a6f96e88dc64ca33e1ac7915dda906e1888e.pdf>).

7. T. D. Seeley, S. Camazine, J. Sneyd, Collective decision-making in honey bees: How colonies choose among nectar sources. *Behavioral Ecology and Sociobiology*. **28**, 277–290 (1991).

8. T. D. Seeley, Honey bee foragers as sensory units of their colonies. *Behavioral Ecology and Sociobiology*. **34**, 51–62 (1994).

9. M. J. Hasenjager, W. Hoppitt, E. Leadbeater, Do honey bees modulate dance following according to foraging distance? *Animal Behaviour*. **184**, 89–97 (2022).

10. K. V. Frisch, *The dance language and orientation of bees* (Harvard University Press, 1967).

11. R. Boch, Die tänze der bienen bei nahen und fernen trachtquellen. *Zeitschrift für Vergleichende Physiologie*. **38**, 136–167 (1956).

12. H. Esch, Über die schallerzeugung beim werbetanz der honigbiene. *Zeitschrift für Vergleichende Physiologie*. **45**, 1–11 (1961).

13. T. D. Seeley, C. A. Tovey, Why search time to find a food-storer bee accurately indicates the relative rates of nectar collecting and nectar processing in honey bee colonies. *Animal Behaviour*. **47**, 311–316 (1994).

14. T. D. Seeley, A. S. Mikheyev, G. J. Pagano, Dancing bees tune both duration and rate of waggle-run production in relation to nectar-source profitability. *Journal of Comparative Physiology - A Sensory, Neural, and Behavioral Physiology*. **186**, 813–819 (2000).

15. G. Sherman, P. K. Visscher, Honeybee colonies achieve fitness through dancing. *Nature*. **419**, 920–922 (2002).

16. A. Dornhaus, L. Chittka, Why do honey bees dance? *Behavioral Ecology and Sociobiology*. **55**, 395–401 (2004).

17. C. Grüter, F. H. Segers, F. L. Ratnieks, Social learning strategies in honeybee foragers: Do the costs of using private information affect the use of social information? *Animal Behaviour*. **85**, 1443–1449 (2013).

18. R. I. Price, N. Dulex, N. Vial, C. Vincent, C. Grüter, Honeybees forage more successfully without the “dance language” in challenging environments. *Science Advances*. **5** (2019), doi:[10.1126/sciadv.aat0450](https://doi.org/10.1126/sciadv.aat0450).

19. C. Grüter, W. M. Farina, The honeybee waggle dance: Can we follow the steps? *Trends in ecology & evolution*. **24**, 242–247 (2009).

20. M. C. Donaldson-Matasci, A. Dornhaus, How habitat affects the benefits of communication in collectively foraging honey bees. *Behavioral Ecology and Sociobiology*. **66**, 583–592 (2012).

21. C. Grüter, E. Leadbeater, Insights from insects about adaptive social information use. *Trends in Ecology & Evolution*. **29**, 177–184 (2014).

22. R. I. Price, C. Grüter, Why, when and where did honey bee dance communication evolve? *Frontiers in Ecology and Evolution*. **3**, 125 (2015).

23. A. Dornhaus, F. Klügl, C. Oechslein, F. Puppe, L. Chittka, Benefits of recruitment in honey bees: Effects of ecology and colony size in an individual-based model. *Behavioral Ecology*. **17**, 336–344 (2006).

24. T. D. Seeley, W. F. Towne, Tactics of dance choice in honey bees: Do foragers compare dances? *Behavioral Ecology and Sociobiology*. **30**, 59–69 (1992).

25. G. H. Pyke, Optimal foraging: Movement patterns of bumblebees between inflorescences. *Theoretical population biology*. **13**, 72–98 (1978).

26. M. Beekman, A. L. Gilchrist, M. Duncan, D. J. Sumpter, What makes a honeybee scout? *Behavioral Ecology and Sociobiology*. **61**, 985–995 (2007).

27. K. Burnham, D. Anderson, *Model selection and multimodel inference: A practical information-theoretic approach.* (Springer New York, ed. 2, 2002).

28. A. E. Samuelson, R. Schürch, E. Leadbeater, Dancing bees evaluate central urban forage resources as superior to agricultural land. *Journal of Applied Ecology* (2021), doi:[10.1111/1365-2664.14011](https://doi.org/10.1111/1365-2664.14011).

29. M. J. Couvillon, F. C. R. Pearce, E. L. Harris-Jones, A. M. Kuepfer, S. J. Mackenzie-Smith, L. A. Rozario, R. Schürch, F. L. W. Ratnieks, Intra-dance variation among waggle runs and the design of efficient protocols for honey bee dance decoding. *Biology Open*. **1**, 467–472 (2012).

30. R. Schürch, K. Zwirner, B. J. Yambrick, T. Pirault, J. M. Wilson, M. J. Couvillon, Dismantling babel: Creation of a universal calibration for honey bee waggle dance decoding. *Animal Behaviour*. **150**, 139–145 (2019).

31. M. L. Goldstein, S. A. Morris, G. G. Yen, Problems with fitting to the power-law distribution. *The European Physical Journal B - Condensed Matter and Complex Systems 2004 41:2*. **41**, 255–258 (2004).

32. A. Clauset, C. R. Shalizi, M. E. J. Newman, POWER-law distributions in empirical data (2009) (available at <http://www.santafe.edu/~aaronc/powerlaws/.>).

33. C. Grüter, F. L. Ratnieks, Honeybee foragers increase the use of waggle dance information when private information becomes unrewarding. *Animal Behaviour*. **81**, 949–954 (2011).

34. A. E. Samuelson, E. Leadbeater, A land classification protocol for pollinator ecology research: An urbanization case study. *Ecology and Evolution*. **8**, 5598–5610 (2018).

35. L. M. Carrascal, I. Galván, O. Gordo, Partial least squares regression as an alternative to current regression methods used in ecology. *Oikos*. **118**, 681–690 (2009).

36. K. C. Baldock, M. A. Goddard, D. M. Hicks, W. E. Kunin, N. Mitschunas, H. Morse, L. M. Osgathorpe, S. G. Potts, K. M. Robertson, A. V. Scott, P. P. Staniczenko, G. N. Stone, I. P. Vaughan, J. Memmott, A systems approach reveals urban pollinator hotspots and conservation opportunities. *Nature Ecology and Evolution*. **3**, 363–373 (2019).

37. A. Arenas, V. M. Fernández, W. M. Farina, Floral odor learning within the hive affects honeybees’ foraging decisions. *Naturwissenschaften*. **94**, 218–222 (2007).

38. R. Schürch, C. Gruẗer, Dancing bees improve colony foraging success as long-term benefits outweigh short-term costs. *PLOS ONE*. **9**, e104660 (2014).

39. F. Nürnberger, I. Steffan-Dewenter, S. Härtel, Combined effects of waggle dance communication and landscape heterogeneity on nectar and pollen uptake in honey bee colonies. *PeerJ*. **2017**, e3441 (2017).

40. W. H. Kirchner, A. Grasser, The significance of odor cues and dance language information for the food search behavior of honeybees (hymenoptera: Apidae). *Journal of Insect Behavior*. **11**, 169–178 (1998).

41. B. Wild, D. M. Dormagen, A. Zachariae, M. L. Smith, K. S. Traynor, D. Brockmann, I. D. Couzin, T. Landgraf, Social networks predict the life and death of honey bees. *Nature Communications 2021 12:1*. **12**, 1–12 (2021).

42. F. Wario, B. Wild, R. Rojas, T. Landgraf, Automatic detection and decoding of honey bee waggle dances. *PLOS ONE*. **12**, e0188626 (2017).

43. W. McKinney, Pandas: A foundational python library for data analysis and statistics. *Proceedings of the 9th Python in Science Conference* (2011).

44. E. Jones, T. Oliphant, P. Peterson, SciPy: Open source scientific tools for python (2001).

45. R. C. Team, R: A language and environment for statistical computing (2020), (available at <https://www.r-project.org/>).

46. J. A. Nelder, R. Mead, A simplex method for function minimization. *The Computer Journal*. **7**, 308–313 (1965).

47. S. G. Johnson, The NLopt nonlinear-optimization package (2020), (available at <http://github.com/stevengj/nlopt>).

48. D. Eddelbuettel, R. François, Rcpp: Seamless r and c++ integration. *Journal of Statistical Software*. **40**, 1–18 (2011).

49. K. Aho, D. Derryberry, T. Peterson, Model selection for ecologists: The worldviews of aic and bic. *Ecology*. **95**, 631–636 (2014).

50. J. S. Sekhon, Multivariate and propensity score matching software with automated balance optimization: The matching package for r. *Journal of Statistical Software*. **42**, 1–52 (2011).

51. F. Bertrand, N. Meyer, M. Beau-Faller, K. E. Bayed, I.-J. Namer, M. Maumy-Bertrand, Régression bêta pls. *Journal de la Société Française de Statistique*. **154**, 143–159 (2013).

52. F. Cribari-Neto, A. Zeileis, Beta regression in r. *Journal of Statistical Software*. **34**, 1–24 (2010).