



Cite this article: Feng Y, Ouellette NT. 2023 Non-uniform spatial sampling by individuals in midge swarms. *J. R. Soc. Interface* **20**: 20220521. <https://doi.org/10.1098/rsif.2022.0521>

Received: 20 July 2022

Accepted: 9 January 2023

Subject Category:

Life Sciences–Physics interface

Subject Areas:

biophysics

Keywords:

collective behaviour, swarming,
spatial sampling

Author for correspondence:

Nicholas T. Ouellette

e-mail: nto@stanford.edu

Non-uniform spatial sampling by individuals in midge swarms

Yenchia Feng and Nicholas T. Ouellette

Department of Civil and Environmental Engineering, Stanford University, Stanford, CA 94305, USA

NTO, 0000-0002-5172-0361

Individual animals engaged in collective behaviour can interchange their relative positions on a wide range of time scales. In situations where some regions of the group are more desirable, it is thought that more fit individuals will preferentially occupy the more favourable locations. However, this notion is difficult to test for animal groups like insect swarms that fluctuate rapidly and display little apparent structure. Here, we study the way that individuals in mating swarms of the non-biting midge *Chironomus riparius* sample the space available to them. We use Voronoi tessellation to define different regions of the swarm in a dynamic way, and show that midges indeed sample the swarm non-uniformly. However, individuals that preferentially reside in the interior or exterior of the swarm do not display statistically distinct flight behaviour, suggesting that differences in fitness must be assessed in a different way. Nevertheless, our results indicate that midge swarms are not random configurations of individuals but rather possess non-trivial internal structure.

1. Introduction

Throughout nature, groups of social animals interact with one another and engage in so-called collective behaviour [1,2]. Although collective behaviour can take many guises, some of its most celebrated and visually striking forms are the coherent and seemingly coordinated motions of ordered groups such as bird flocks and fish schools [1,3]. On short time scales, the group-level pattern tends to be the dominant observable feature of these and other similar animal groups, and there is limited rearrangement of individuals [4,5]. Since they are also typically leaderless, it would be reasonable to conclude that the relative positioning of individuals in the group is unimportant, and is perhaps set by initial conditions rather than by any dynamical or behavioural processes. At the other extreme, in animal groups such as insect swarms that do not display coherent motion [6], individuals change their location rapidly [6,7] in a way that is often uncorrelated with their neighbours [8]—again suggesting that relative position is uncontrolled. However, upon further consideration, counterexamples to these conclusions can be found. In the well-known V-formation flight of long-range migratory birds, for example, individuals switch positions on long time scales to balance energy costs in the group [9]. In shorter-range transit flocks, individuals may preferentially locate themselves near others with whom they share social bonds [5]. And even in disordered swarms, different locations may be differently valuable [10], potentially leading to subtle but persistent biases in spatial positioning. These examples suggest that the relative positioning of animals in groups may thus be valuable to investigate.

An early and influential framework for understanding and predicting spatial arrangement in animal groups is Hamilton's 'selfish herd' theory [10], which argues that individuals in a group compete with each other to position themselves in more desirable locations. In the context of social prey animals, assuming that predators come from outside the group and focus on the nearest prey, individual animals in the outer regions of the group face a higher risk of predation, whereas those in the interior regions enjoy a safety cushion between themselves and the outside predators [11]. Hamilton thus predicted that, in

this case, more fit individuals would be found in the interior of the group, while weaker individuals would be pushed to the edges. However, central regions may not necessarily always be the most desirable [10]. For example, for schools of Atlantic silversides facing predators from above and beneath, staying in the interior of the group was found to increase the predation risk [12]. In other groups, such as mating leks, reducing the risk of predation is not the only goal; in these cases as well, staying in the middle of the group may not be the best strategy [13]. Thus, making specific predictions using selfish-herd theory for how individuals will arrange themselves thus depends on understanding the function of the group; but the notion that not all regions of the group are equivalent and that more fit individuals will occupy the most desirable locations remains influential.

As quantitative data on collective behaviour becomes more readily available, an additional challenge that arises is understanding the role played by time. While Hamilton's original conception of the selfish herd was one in which the group settled into a static equilibrium, real animal groups in nature are continually in motion, and both the local, relative configurations of animals and the global group structure change dynamically. The time scale on which structural rearrangements occur can vary widely. For example, rearrangements have been observed to be slow compared with both the group lifetime and the time scale of individual behavioural change in bird flocks [4,5], while insects in swarms change their relative configurations very rapidly [6,7]. In such dynamic cases, preferred regions of the groups in the spirit of selfish-herd theory can be difficult to identify, if they exist at all; it may even be that individual configurations in a dilute, dynamic group like an insect swarm are ephemeral enough that preferred regions cannot be established. Clues as to whether different regions are more or less preferable, however, may potentially be found by examining how uniformly individuals explore the space available to them: if they sample some regions of the group more than others, it may indicate that such regions are more desirable.

Here, we investigated how individual *Chironomus riparius* midges explored the available volume in laboratory mating swarms. To distinguish different regions of the swarm dynamically, we constructed time-varying Voronoi tessellations of the swarms using the three-dimensional, time-resolved trajectories of the individual midges. Based on the location of the Voronoi centroids, we divided the swarms into three distinct layers. Even though the trajectories of individual midges are erratic and there is no coherent motion, we found that individuals do not sample the swarm volume uniformly; rather, individual midges preferentially spent their time in one layer over the others. Although we did not find statistical evidence of different kinematics for midges that preferentially spent their time in any given layer, we hypothesize that the distinctions may be related to individual fitness. Our results both demonstrate that midge mating swarms are not random assemblies of individuals but rather contain non-trivial structure, and underscore the importance of biological context in determining this structure even in highly dynamic situations.

2. Methods

The swarm data analysed here is freely available, and was described and archived by Sinhuber *et al.* [14]. They imaged

swarming events of *C. riparius* midges in a laboratory setting devoid of external stimuli at a frame rate of 100 Hz and reconstructed three-dimensional, time-resolved trajectories with positions, velocities and accelerations of each individual insect in the swarm. The midge colony was kept in a $(122\text{ cm})^3$ enclosure with clear acrylic sides. Swarms comprised exclusively male midges, as females of this species do not swarm. The total volume occupied by a swarm varies with the number of individuals in the swarm, as midges tend to keep their number density relatively consistent from swarm to swarm [6]. The typical swarm diameter was about 50 cm (meaning that the swarms do not fill the enclosure), and, as swarms were positioned near the centre of the enclosure using an artificial swarm marker, the individuals always remained far from any of the enclosure walls. The length of each observation was about 20 000 frames, or approximately 3 min. Of the 19 swarming observations in the dataset, we analysed five observations with similar mean swarm sizes: Ob2, Ob3, Ob7, Ob9 and Ob14, with corresponding mean swarm sizes of $N = 68, 46, 58, 49$ and 54. Further details of the datasets are given in [14].

Because the spatio-temporal reconstruction of midge trajectories is never perfect, not all trajectories in the datasets are of equal length even when midges do not enter or leave the swarm. This does not pose a problem if one is only interested in instantaneous spatial statistics. As we are interested in how midges sample the swarm volume, however, we must also consider the temporal aspect of the data. To avoid potential bias due to differing trajectory lengths, we therefore subdivided each trajectory into segments of a uniform 4 s duration. This time is sufficiently long that an individual moving at a typical speed could traverse the swarm multiple times [14,15]. Additionally, varying the lengths of the trajectory segments over a reasonable range (up to twice as long) did not significantly affect our results.

Since midge swarms are dilute, local densities within the swarm tend to experience fast fluctuations, and the global distribution of midges in the swarm volume can change significantly on short time scales. Therefore, the relative desirability of locations in the swarm should not be referenced to a fixed spatial reference frame but rather to the dynamic configurations of the individuals. In previous work, we considered subdividing swarms into components using a topological approach [16]; however, as with all topological schemes, this method produced a fairly coarse segmentation. For our purposes here, we turned instead to a geometric method to describe the configurations of the individuals in the swarm. We constructed instantaneous Voronoi tessellations of the swarms. We associated each midge with a unique Voronoi polyhedron, which bounds the volume in space that is everywhere closer to the associated midge than to any other midges in the swarm. The centroids of these Voronoi polyhedra are the most empty locations in the swarm, in that they are the points farthest from all the nearby midges. In a swarm of finite size, Voronoi centroids can be partitioned into three classes based on their location relative to the swarm: those that are located an infinite distance from the swarm centre, those that are located at a finite distance but lie outside the swarm and those that lie inside the swarm. Voronoi centroids located at infinity are associated with midges that make up the convex hull of the swarm: Voronoi polyhedra associated with these midges are unbounded and extend outward from the swarm to infinity. Voronoi centroids located not at infinity but outside the swarm are associated with midges that are positioned in the bulk of the swarm but close to the convex hull. Voronoi centroids located inside the swarm are associated with midges that are deeper in the swarm. Based on these three classes of Voronoi centroids, we thus define three regions of the swarm: an outer layer, a middle layer and an inner layer. Because we recompute the Voronoi tessellation at every time step based on the

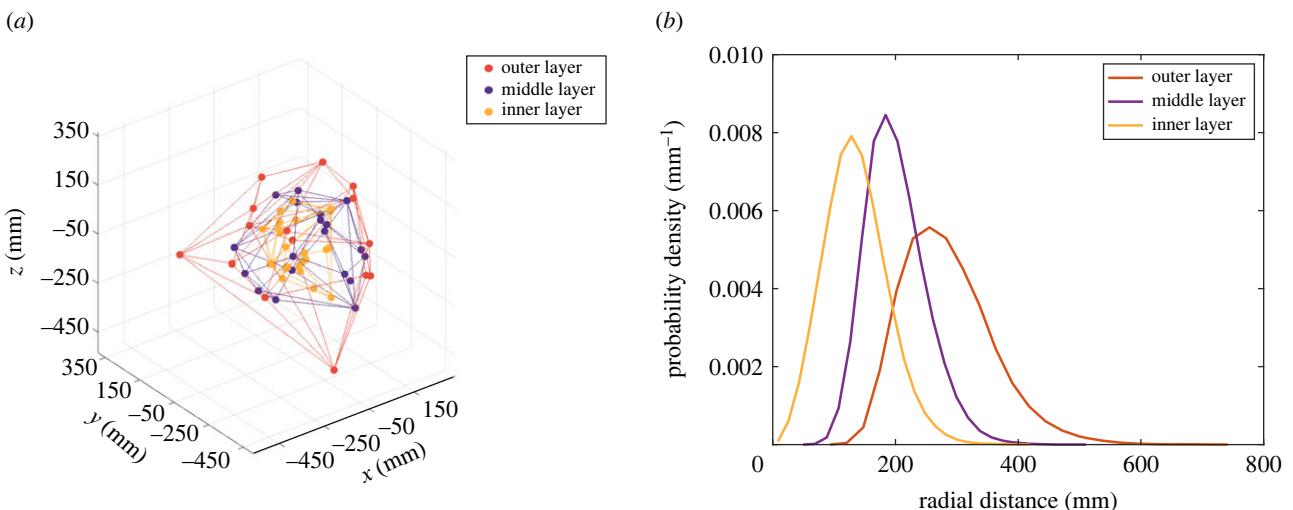


Figure 1. (a) Partitioning of a swarm of 59 individuals at a single instant into the three layers described in the text. Data are taken from dataset OB2. As expected, the three layers are spatially nested, but are not spherical. (b) Probability density functions (PDFs) of the radial distance of individuals in each layer from the swarm centre of mass. Although individuals in the outer and middle layers are generally farther from the swarm centre than individuals in the inner layer, all three PDFs overlap, showing that radial distance alone is not a sufficient indicator of the local environment of an individual midge.

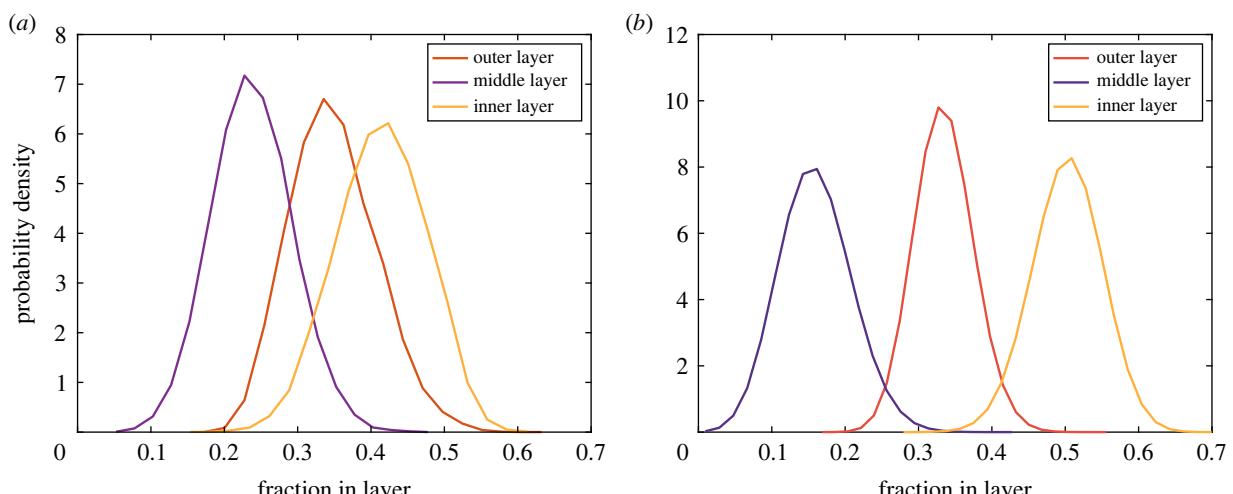


Figure 2. PDFs of the instantaneous fraction of individuals in each layer for (a) the swarm datasets and (b) the Monte Carlo uniform sampling described in the text. The fraction of individuals in the middle layer is enhanced at the expense of the fraction of individuals in the inner layer in the swarm data as compared with the uniform sampling case.

instantaneous midge locations, the locations, sizes and composition of these layers are dynamic. Note too that even though we define these layers based on a mathematical construction, it is plausible that individual midges would be aware to some degree of where they are located relative to the rest of the swarm (that is, nearer to the edge or nearer to the middle).

3. Results

In figure 1a, we show the instantaneous configuration of a swarm with individuals located in the outer, middle and inner layers indicated. As expected, these layers are concentric, though they are not spherical shells. The relative size, position and composition of these layers also display significant fluctuations, given that swarms tend to be dilute [7]. To illustrate this complexity, in figure 1b we show the probability density functions (PDFs) of the radial distance of each midge from the centre of mass of the swarm, conditioned on the layer in which the midge is instantaneously

located. Although the mean values of these three distributions are certainly distinct and are ordered as one would expect, there is substantial overlap. This result underscores the distinction between our Voronoi-based partitioning of the swarm and a simple distance-based segmentation.

Our primary goal in segmenting the swarm into layers is to study how the midges arrange themselves in the available volume so as to understand whether swarms possess any persistent internal structure. If they do, we would expect that midges would sample the swarm non-uniformly. One simple way to begin to address this question is simply to count the number of midges in each layer at any given time. In figure 2a, we show the PDFs of the instantaneous fraction of the total number of midges in the swarm residing in each of the layers. The middle layer typically contains the fewest midges, and the inner layer typically contains the most. As with the distance from the swarm centre, however, the fraction of midges in each layer also displays significant fluctuations as the configuration of the midges in the swarm changes.

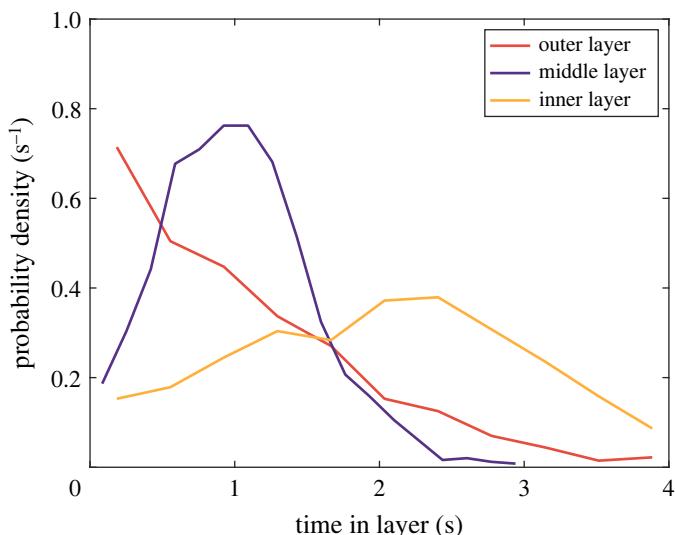


Figure 3. PDFs of the time spent by midges in each layer of the swarm along the 4 s trajectory segments analysed. The PDFs of the time spent in the inner and outer layers are very wide, indicating that some midges spend nearly all of their time in only one layer.

To determine whether these PDFs are evidence of non-uniform sampling, we need to compare them with what would be expected for the case of uniform sampling. If the swarm layers were simple spherical shells, we could directly compute expectations for the fraction of midges in each layer analytically. However, because they are highly non-spherical and fluctuating, we must take a different approach. We used a Monte Carlo method to estimate the PDFs of the fraction of individuals expected in each layer given uniform sampling. To do so, we first calculated the largest minimum bounding sphere of swarm Ob7, a swarm containing a mean number of individuals of $N=58$ that was close to the average number of individuals in all the swarms considered ($N_{\text{avg}}=55$). We then placed N_{avg} particles randomly inside this sphere, and computed the Voronoi tessellation and thus the inner, middle and outer layers for this random configuration. We repeated this process 5 million times and constructed the PDFs of the fraction of points in each layer, shown in figure 2b. Although the relative fraction of individuals in each layer is ordered in the same way for the actual midges (figure 2a) and the Monte Carlo estimate (figure 2b), in that individuals are most likely to be found in the inner layer, then the outer layer and then the middle layer, the quantitative results are different (as confirmed by a two-sample Kolmogorov–Smirnov test). In particular, the fraction of individuals in the middle layer is enhanced for the real midges relative to uniform sampling, at the expense of the inner layer. This result suggests both that midges indeed do not sample the swarm uniformly and that the outer parts of the swarm may be more favourable than the innermost regions.

More compelling evidence that individual midges do not explore the swarm volume uniformly can be seen not by considering snapshots of their relative positions alone but instead their time history. To that end, in figure 3 we show the PDFs of the time spent in each layer by individual midges, recalling that (for consistency) we broke each midge trajectory into uniform 4 s segments. Most of the midges spend roughly a quarter of their time in the middle layer, with a fairly narrow distribution. The outer-layer PDF peaks near zero, indicating that many midges spend essentially no time on the edges of the swarm, while the inner-layer PDF peaks at longer times. However, the widths of these PDFs are much larger than the

middle-layer PDF, showing that the time spent by midges in these layers is more variable: some midges appear to primarily be inner-layer midges, while others are primarily outer-layer midges. Although we cannot compare these results against a suitable null model, as there is no accepted biologically accurate swarm model [2,3], when combined with the instantaneous results shown in figure 2, these PDFs provide strong evidence that midges do not sample the swarm space uniformly, but rather that individuals have distinct preferences as to where they spend their time.

It is reasonable to ask whether midges that preferentially spend their time in one layer are different in some measurable way from those that spend their time in a different layer. The simplest physical metrics that may distinguish these individuals are their kinematic statistics—that is, their speed and acceleration. We thus computed the PDFs of these kinematic quantities conditioned on which layer the midges occupied. We conditioned the PDFs in two ways. First, as shown in figure 4, we conditioned the speed and acceleration PDFs on the instantaneous layer occupied by the individual midges, regardless of which layer those midges spent most of their time in. But because we found that midges tend to display a preference for one layer over the others, we also conditioned the PDFs based on the layer in which the midges spent most of their time, as shown in figure 5. For these latter calculations, we assigned the kinematic information along each entire trajectory to one of three statistical ensembles based on the layer in which the midge spent the plurality of its time, and then computed the statistics of those ensembles. Regardless of how we did the conditioning, however, there is little difference between the kinematic statistics for midges in different layers of the swarm. Thus, although it stands to reason that midges that prefer one layer over another ought to be different in some way, those differences are not reflected in their kinematic flight properties.

4. Discussion and conclusion

By dynamically constructing Voronoi tessellations at every time instant, we have shown that swarms can be dynamically segmented into three distinct and non-trivial layers in a time-

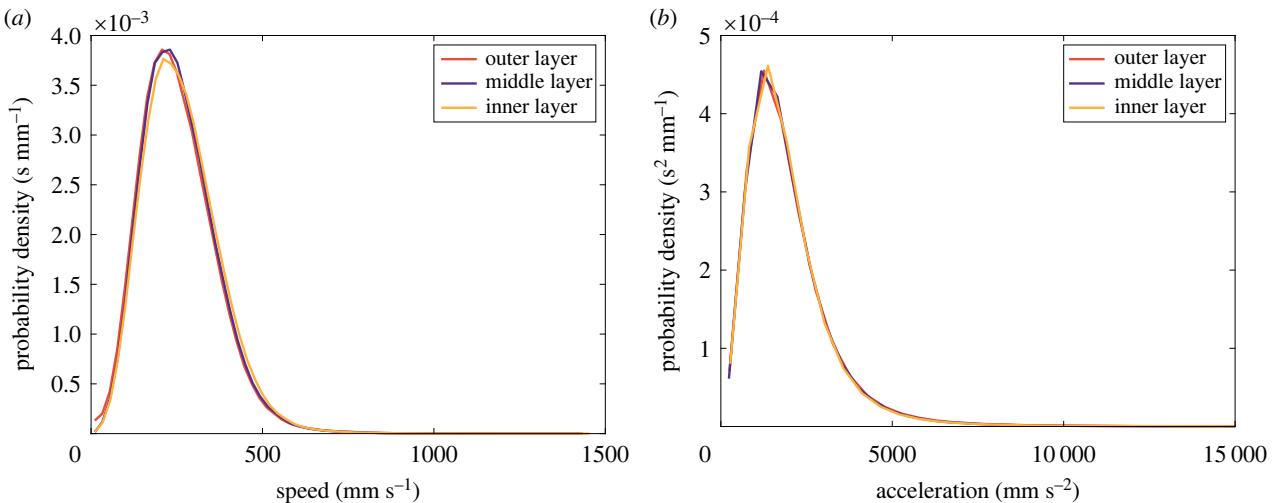


Figure 4. PDFs of the (a) speed and (b) acceleration magnitude conditioned on the instantaneous layer assignment of a midge. Midges in different layers do not display different kinematic statistics.

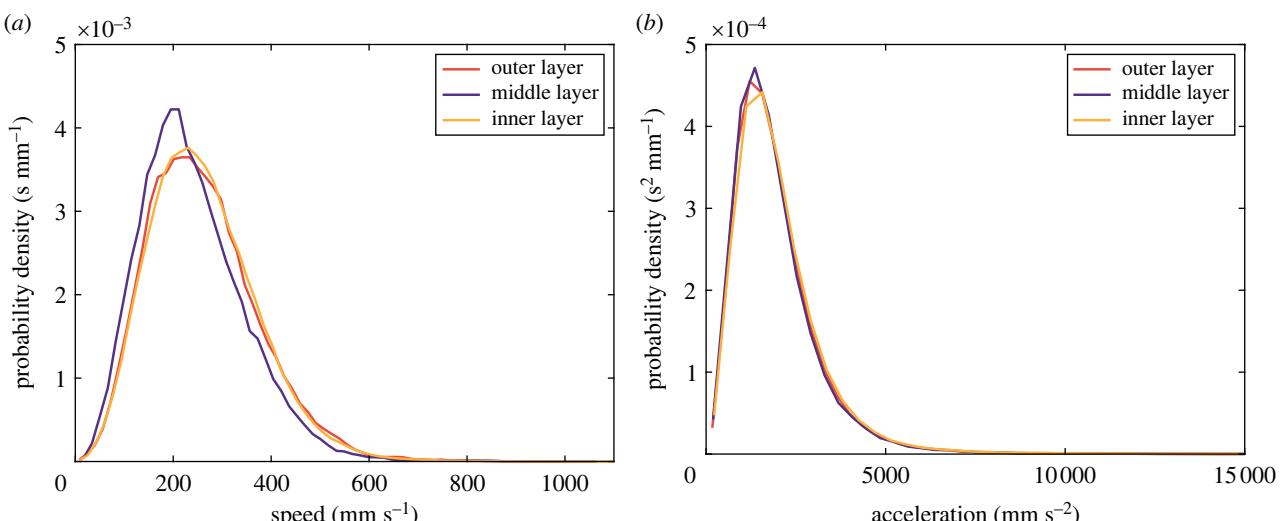


Figure 5. PDFs of the (a) speed and (b) acceleration magnitude conditioned on the layer in which a midge spends the plurality of its time along the trajectory segment analysed. With the exception of midges that spend most of their time in the middle layer, which move slightly more slowly, midges that spend their time in different layers also do not display different kinematic statistics.

dependent fashion. Individual midges do not, however, sample these layers uniformly. Rather, even though swarming midges are continually in motion and rearranging, individuals tend to spend the bulk of their time in only one of the layers.

Given that midges do not sample the swarm volume randomly but rather preferentially spend their time either on the exterior or in the interior of the swarm, selfish-herd theory suggests that these different regions of the swarm ought to have different costs and benefits to individuals, and that more fit individuals ought to occupy the preferred locations [10]. Applying this idea, however, proves challenging for at least two distinct reasons.

At the level of individual midges, we lack sufficient information to determine which individuals are more fit. A reasonable initial hypothesis is that more fit individuals may be able to execute more rapid or dynamic flight manoeuvres so as to be able to chase and catch females more effectively when they enter the swarm [17]. However, as shown in figures 4 and 5, the observed flight kinematics of inner-layer, middle-layer and outer-layer midges are not

statistically distinguishable. It is possible that differences in flight performance may only be observable in the presence of females or predators; without these kinds of external cues, there is not necessarily a good reason for an individual midge to over-exert itself. But it may also be that signatures of fitness are only evident in other biological characteristics (such as appearance, sensory capability or body size [18], for example) of the midges and not in their motion ability [19]. Neither of these scenarios is measurable in our experiments as currently constituted.

We also lack full information about the purpose of the swarming behaviour, so that we cannot *a priori* argue which regions of the swarm ought to be most desirable and thus occupied by the most fit individuals. It is generally agreed that swarming is an essential component of midge mating [17,18,20], with leks of males presenting a target for disparate females to visit to find mates. Since females come from outside the swarm, it would then stand to reason that outer regions might be more desirable, as those male midges on the swarm edges would get a head start on chasing incoming females [10,17]. This kind of argument is indirectly supported

by our observation that the outer regions of the swarm contain higher fraction of the midges than would be expected given uniform sampling (figure 2). However, this simple picture is not without complications. For example, even though a male on the edge of the swarm could be in just the right spot to encounter an incoming female, this location could also in other circumstances be the *least* desirable spot, if the female approached the swarm from its other side. Since females may approach the swarm from any direction, it is then plausible that in fact the interior regions of the swarm are more desirable as they would give a male there an equal likelihood of catching a female entering the swarm from any direction, and that then the under-representation of inner-layer midges relative to the uniform estimate could be a manifestation of particularly fit individuals there driving away others [19]. As additional support for this hypothesis, we note that midge predators *also* reside outside the swarm, and so midges on the exterior may be at a higher risk of predation [13]. Thus, without further information to distinguish these scenarios, we cannot use the observable spatial positioning of the midges as a proxy for their fitness.

Our results thus underscore the necessity of combining physics-based analysis and biological insight for understanding the dynamics of collective behaviour [2]. The Voronoi analysis we have described demonstrates that midges do not sample the swarm uniformly, even though it is not *a priori* obvious; but without additional biological knowledge, the full ramifications of this finding cannot be assessed. Future studies of collective behaviour should continue to embrace this cross-disciplinary spirit—and in this vein, we close by noting that the Voronoi analysis we use here can be straightforwardly applied to any animal group for which the trajectories of individuals are known.

Data accessibility. All data used for this work can be found in [14].

Authors' contributions. Y.F.: investigation, writing—original draft; N.T.O.: conceptualization, supervision, writing—review and editing. All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. No funding has been received for this article.

References

- Parrish JK, Edelstein-Keshet L. 1999 Complexity, pattern, and evolutionary trade-offs in animal aggregation. *Science* **284**, 99–101. (doi:10.1126/science.284.5411.99)
- Ouellette NT, Gordon DM. 2021 Goals and limitations of modeling collective behavior in biological systems. *Front. Phys.* **9**, 687823. (doi:10.3389/fphy.2021.687823)
- Ouellette NT. 2022 A physics perspective on collective animal behavior. *Phys. Biol.* **19**, 021004. (doi:10.1088/1478-3975/ac4bef)
- Mora T, Walczak AM, Del Castillo L, Ginelli F, Melillo S, Parisi L, Viale M, Cavagna A, Giardina I. 2016 Local equilibrium in bird flocks. *Nat. Phys.* **12**, 1153–1157. (doi:10.1038/nphys3846)
- Ling H, McIvor GE, van der Vaart K, Vaughan RT, Thornton A, Ouellette NT. 2019 Costs and benefits of social relationships in the collective motion of bird flocks. *Nat. Ecol. Evol.* **3**, 943–948. (doi:10.1038/s41559-019-0891-5)
- Kelley DH, Ouellette NT. 2013 Emergent dynamics of laboratory insect swarms. *Sci. Rep.* **3**, 1073. (doi:10.1038/srep01073)
- Puckett JG, Kelley DH, Ouellette NT. 2014 Searching for effective forces in laboratory insect swarms. *Sci. Rep.* **4**, 4766. (doi:10.1038/srep04766)
- Ni R, Ouellette NT. 2015 Velocity correlations in laboratory insect swarms. *Eur. Phys. J. Spec. Top.* **224**, 3271–3277. (doi:10.1140/epjst/e2015-50077-5)
- Voelkl B, Portugal SJ, Unsöld M, Usherwood JR, Wilson AM, Fritz J. 2015 Matching times of leading and following suggest cooperation through direct reciprocity during V-formation flight in ibis. *Proc. Natl Acad. Sci. USA* **112**, 2115–2120. (doi:10.1073/pnas.1413589112)
- Hamilton WD. 1971 Geometry for the selfish herd. *J. Theor. Biol.* **31**, 295–311. (doi:10.1016/0022-5193(71)90189-5)
- Krause J. 1993 The effect of 'Schreckstoff' on the shoaling behaviour of the minnow: a test of Hamilton's selfish herd theory. *Anim. Behav.* **45**, 1019–1024. (doi:10.1006/anbe.1993.1119)
- Parrish JK. 1989 Re-examining the selfish herd: are central fish safer? *Anim. Behav.* **38**, 1048–1053. (doi:10.1016/S0003-3472(89)80143-5)
- Neems RM, Lazarus J, McLachlan AJ. 1992 Swarming behavior in male chironomid midges: a cost-benefit analysis. *Behav. Ecol.* **3**, 285–290. (doi:10.1093/beheco/3.4.285)
- Sinhaber M, van der Vaart K, Ni R, Puckett JG, Kelley DH, Ouellette NT. 2019 Three-dimensional time-resolved trajectories from laboratory insect swarms. *Sci. Data* **6**, 190036. (doi:10.1038/sdata.2019.36)
- Puckett JG, Ni R, Ouellette NT. 2015 Time-frequency analysis reveals pairwise interactions in insect swarms. *Phys. Rev. Lett.* **114**, 258103. (doi:10.1103/PhysRevLett.114.258103)
- Sinhaber M, Ouellette NT. 2017 Phase coexistence in insect swarms. *Phys. Rev. Lett.* **119**, 178003. (doi:10.1103/PhysRevLett.119.178003)
- Sullivan RT. 1981 Insect swarming and mating. *Fla. Entomol.* **64**, 44. (doi:10.2307/3494600)
- McLachlan A, Neems R. 1995 Swarm-based mating systems. In *Insect reproduction* (ed. SR Leather). Boca Raton, FL: CRC Press.
- Fiske P, Rintamäki PT, Karvonen E. 1998 Mating success in lekking males: a meta-analysis. *Behav. Ecol.* **9**, 328–338. (doi:10.1093/beheco/9.4.328)
- Downes JA. 1969 The swarming and mating flight of Diptera. *Annu. Rev. Entomol.* **14**, 271–298. (doi:10.1146/annurev.en.14.010169.001415)