

Étude des données

1 Histogrammes sur les vitesses

1.1 Histogramme des vitesses instantanées

J’ai vérifié le calcul des magnitudes des vitesses à partir des positions et obtenu une conformité complète avec la colonne “v_mag” fournie dans le fichier. Ci-dessous l’histogramme des vitesses instantanées, capées à 50, celles au-delà de la trentaine étant des anomalies.

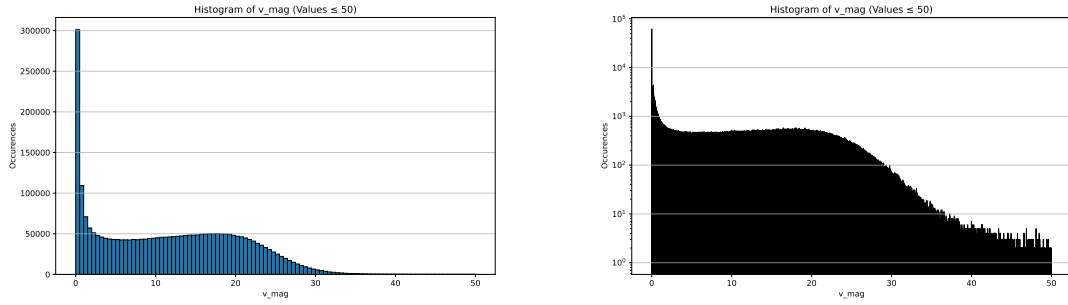


Figure 1. Histogramme de la magnitude des vitesses. À gauche : 100 bins, échelle linéaire. À droite : 10000 bins, échelle log.

1.2 Trajectoires totales

Comme il arrive fréquemment que l’ID des fourmis disparaisse, soit échangé ou réapparaisse, je me restreins pour cette section aux trajectoires que j’appelle “totales”, c’est à dire où la fourmi est suivie de son entrée sur le pont à sa sortie du pont. Le pont est représenté par un rectangle $[0, 100] \times [0, 20]$, je caractérise ces trajectoires en demandant à ce que la première occurrence de la position appartienne à $[0, 5] \times [0, 20]$ et la dernière à $[95, 100] \times [0, 20]$, ou inversement. On compte 3740 telles trajectoires (ou équivalamment d’IDs), pour un total de 23884. On voit que le “saut d’ID” est extrêmement fréquent !

Par suite, je relève pour chaque fourmi i le temps total passé sur le pont, $T_{\text{total},i}$ ainsi que son module de vitesse instantanée moyen sur toute la traversée du pont, \bar{v}_i . Je supprime les trajectoires pour lesquelles le temps total est supérieur à 300 frames (12 secondes). Sur la figure ci-dessous sont représentés les histogrammes de la vitesse moyenne pour traverser le pont, $100/T_{\text{total},i}$ avec le temps en secondes, et \bar{v}_i . C’est sans surprise que le module de vitesse instantanée moyen se trouve être supérieur au module de la vitesse moyenne de parcours du pont (notamment car la variable y n’est pas prise en compte).

Temps d’arrêt sur les trajectoires totales. Je tiens à souligner un effet de biais de la sélection de ces trajectoires. En effet, comme les “sauts d’ID” surviennent principalement lors des interactions (effet de brouillage), les trajectoires totales sont naturellement celles de fourmis ayant peu interagi lors de leur parcours (plus il y a d’interactions, plus il y a de chances pour que l’ID “saute”). Ainsi, on voit en Figure 3, où l’on représente l’histogramme des vitesses pour les trajectoires totales, que le bin 0 n’est pas tellement plus chargé que les autres, alors qu’on a vu sur la figure 1 qu’au total, il est extrêmement chargé comparé aux autres. Les trajectoires *totales*, en l’état, sont donc bien adaptées à l’étude de la vitesse de parcours et du comportement hors interaction, mais assez mal à l’étude des interactions. Il faudra pour cela “rafistoler” les ID, pour obtenir un ensemble de trajectoires totales plus représentatif.

Il me semble assez naturel de travailler, par suite, sur ces trajectoires “totales”, en gardant les autres agents uniquement comme partenaires (et non sujets). Cela permettra de travailler sur des trajectoires “propres”.

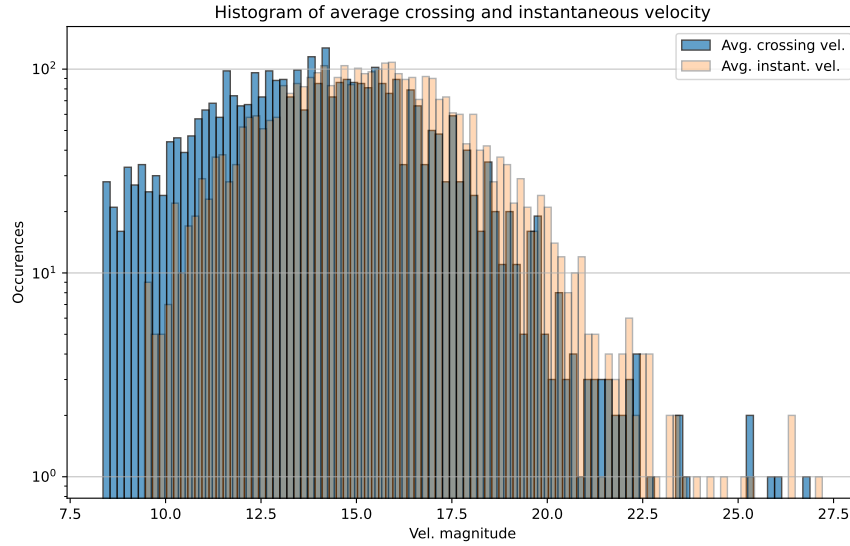


Figure 2. Histogrammes des modules de vitesses moyennes

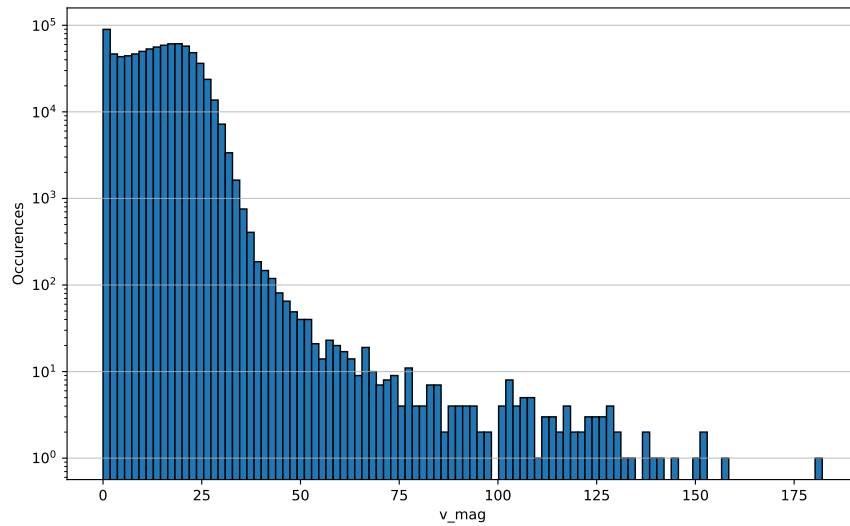


Figure 3. Histogrammes de la magnitude des vitesses *pour les trajectoires totales*.

2 Visualisation des trajectoires

Il est fondamental de pouvoir visualiser les trajectoires : c’est l’objectif du script “visualisation.py”. Pour exemple, voir la vidéo “agent_animation_30s.mp4” à récupérer depuis `Data_management/animation.zip`, dont on donne un aperçu en Figure 2 ci-dessous.

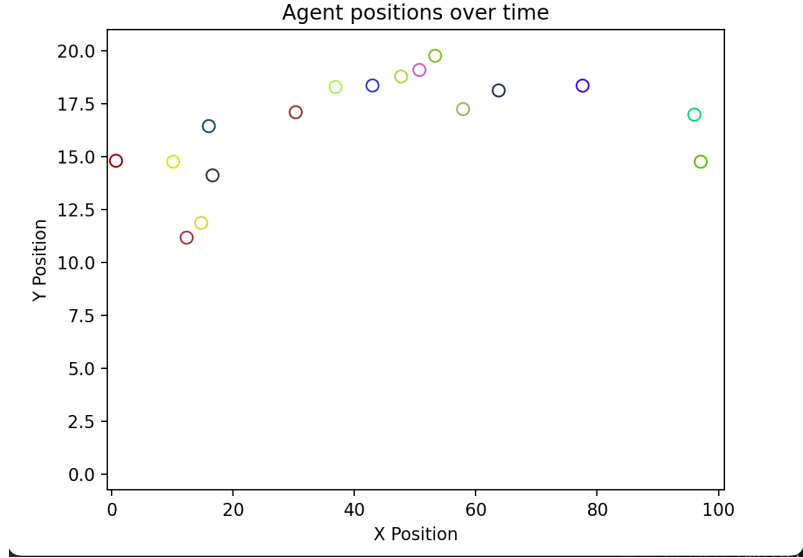


Figure 4. Visualisation des trajectoires

3 Écart-type de la vitesse

Pour répondre à la question de la vitesse constante ou non, on récupère **pour chaque trajectoire** l'écart-type de la variable de vitesse sur toute la trajectoire divisée par l'espérance de la vitesse, en retirant les arrêts : on filtre donc les vitesses inférieures à un seuil choisi. On trace l'histogramme de ce que l'on a récupéré :

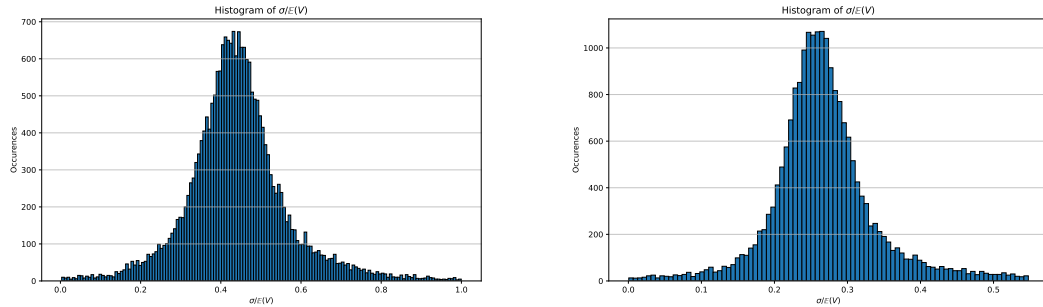


Figure 5. Histogramme de l'écart-type de la vitesse le long d'une trajectoire divisé par l'espérance de la vitesse sur cette même trajectoire. A gauche, le seuil est 3, à droite le seuil est 10.

Pas de doute, ce sont des gaussiennes !

Celle de gauche, (on a retiré les valeurs où $v_{\text{mag}} < 3$) est de **moyenne 0.44** et d'**écart-type 0.12**.

Celle de droite, (on a retiré les valeurs où $v_{\text{mag}} < 10$) est de **moyenne 0.27** et d'**écart-type 0.07**.

Faibles valeurs de seuil:

On observe clairement une gaussienne quand le seuil n'est pas trop faible (par tâtonnement, environ 1.5). Lorsqu'on passe en dessous de ces valeurs, la distribution est une gaussienne modifiée avec une surcharge des plus hautes valeurs, voir la figure ci-dessous.

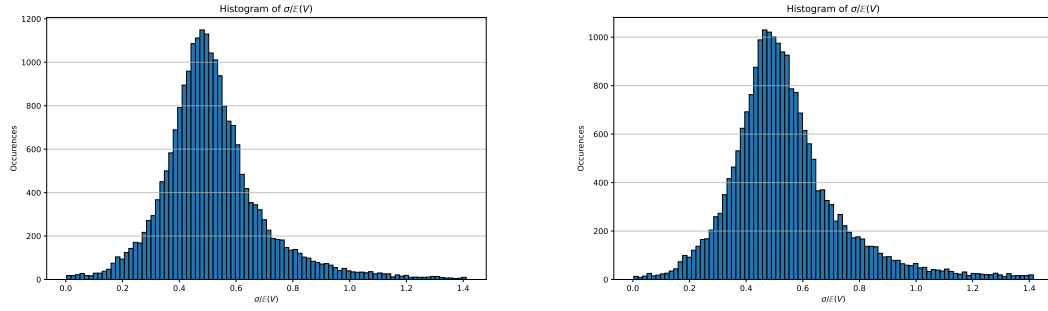


Figure 6. Cas des seuils 1 (à gauche) et 0 (à droite).

Pour le seuil égal à 1, la gaussienne est de **moyenne 0.5 (!)** et d'**écart-type 0.15**.

Sans seuil, la distribution quasi-gaussienne est de **moyenne 0.54** et d'**écart-type 0.2**.

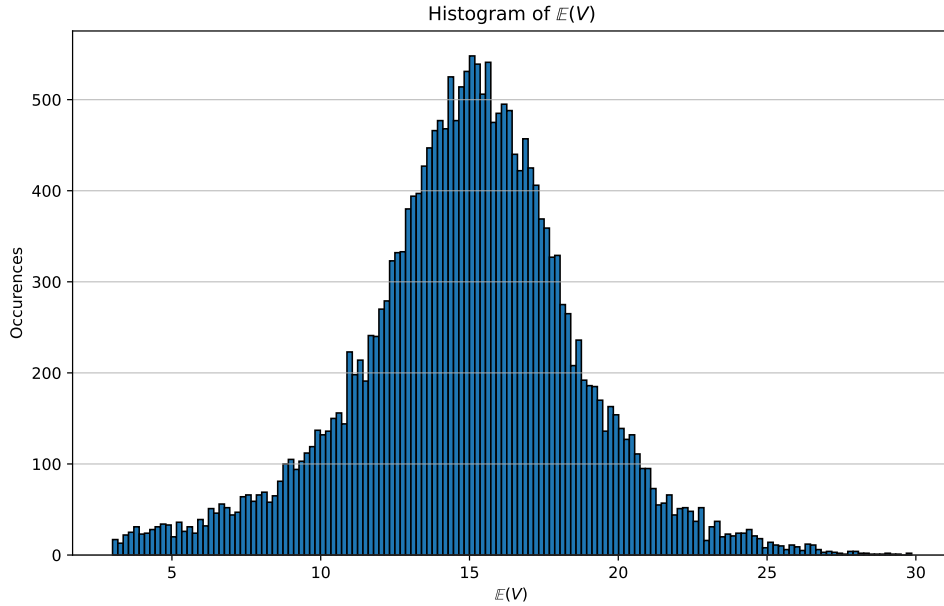


Figure 7. Histogramme de l'espérance de la variable de vitesse le long d'une trajectoire, où le seuil est 3.

On obtient (presque) une gaussienne, de **moyenne 14,9** et d'**écart-type 3.77**. On note une légère surcharge au niveau des faibles valeurs comparativement à une vraie gaussienne, probablement due aux arrêts.

Conclusion Ces données nous permettent d'affirmer que la vitesse prend *un continuum de valeurs* le long de la trajectoire, et que, *le long d'une trajectoire typique*, on a que pour un certain $\bar{v} \in [10, 20]$,

$$\left[\frac{1}{2} \bar{v}, \frac{3}{2} \bar{v} \right] \subset \text{Im}(t \mapsto v(t)).$$

Les fractions $\frac{1}{2}$ proviennent ici du cas du seuil égal à 1 (la moyenne de la gaussienne est 0.5).