

Modèle

Le manguier est un arbre qui présente de forts asynchronismes inter- et intra-arbres. Cette particularité induit une fenêtre d'exposition élargie des inflorescences et des fruits aux ravageurs. Parmi ceux-ci se trouve la cécidomyie qui pond dans les inflorescences, pouvant provoquer la mort de celles-ci.

1 Connaissances

1.1 Inflorescences des manguiers

Les inflorescences ont une durée de vie théorique de 50 jours. Au cours de son développement, l'inflorescence traverse plusieurs stades phénologiques : C/D, E, F/PF. Ce que l'on considère comme la naissance est en fait le débourrement qui marque également le début du stade C. Les durées des différents stades phénologiques sont les suivantes :

- C/D : 7 jours ;
- E : 9 jours ;
- F/PF : 34 jours.

Ces stades influent probablement sur l'attractivité des inflorescences sur les cécidomyies.

1.2 Cécidomyies

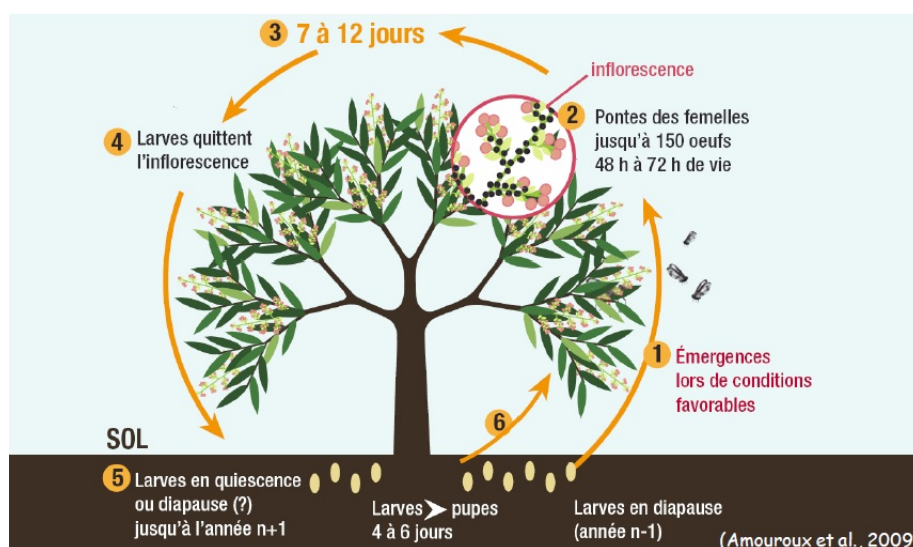


FIGURE 1 – Cycle de développement des cécidomyies

Le cycle de développement des cécidomyies est présenté sur la figure 1. Les femelles présentes dans le verger à la date t pondent des oeufs dans les inflorescences. Sept à douze jours plus tard, les oeufs se sont transformés en larves et s'éjectent des inflorescences pour aller s'enfouir dans le sol. Dans le sol, les larves peuvent soit entrer en diapause, soit entrer en pupaison. Si elles entrent en diapause, alors les femelles ne ressortent que l'année suivante. Si elles entrent en pupaison, alors le cycle de développement se poursuit et des cécidomyies adultes émergent quatre à six jours plus tard, perpétuant ainsi le cycle.

2 Données

Une expérimentation a été faite en 2017 et 2018 pour tester un éventuel effet de la modalité de couverture du sol sur l'impact des cécidomyies. Un verger a été séparé en trois, chacune des parties ayant une modalité de couverture du sol différente : enherbement ras, paillage synthétique ou enherbement haut.

Pour l'année 2017 on dispose de deux suivis sur ce verger, donnant deux jeux de données. Le *dataset 1* porte sur les inflorescences : leurs dates de débourrements et leurs dates de morts sont recensées. Cela nous permet d'avoir le nombre d'inflorescences vivantes à chaque date, mais aussi la durée de vie moyenne observée d'une inflorescence ainsi que des estimations de la proportion des différents stades phénologiques des inflorescences à chaque date.

Le *dataset 2* porte sur les cécidomyies et les inflorescences. Des pièges ont été placés sous les inflorescences pour récolter les larves qui s'en éjectent dans chacun des trois sous-vergers. On a le nombre d'inflorescences vivantes au-dessus des pièges et le nombre des inflorescences des arbres suivis et le nombre de larves piégées à chaque date.

En se basant sur le *dataset 2*, il a également été possible de reproduire la dynamique d'inflorescences vivantes, en la simulant. Cela nous permettra de créer des dynamiques dites attractives.

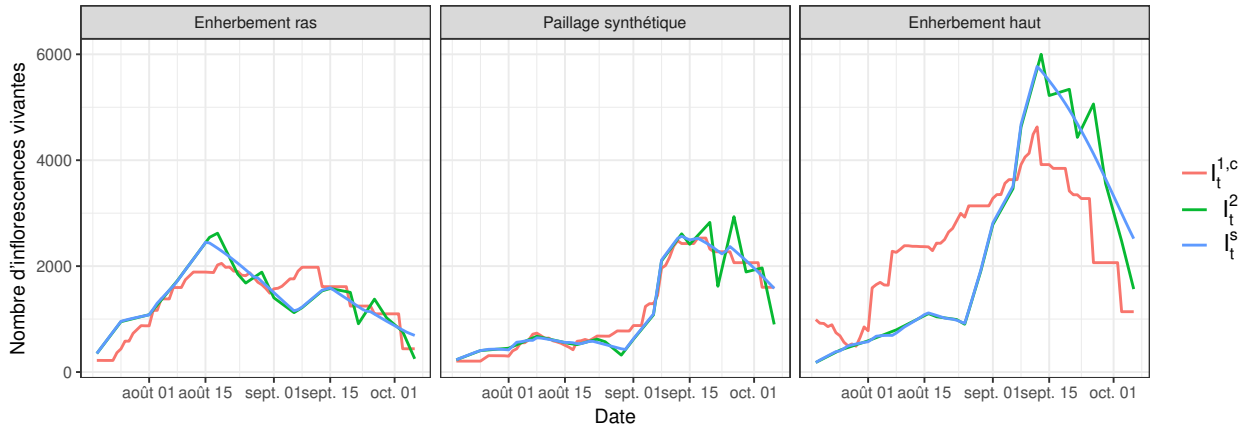


FIGURE 2 – Comparaison des trois dynamiques d'inflorescences obtenues : $I_t^{1,c}$ est issue du *dataset 1*, I_t^2 est issue du *dataset 2* et I_t^s est simulée.

3 Hypothèses

Pour modéliser le phénomène, on émet des hypothèses.

- **Les femelles ne vivent qu'un jour** : les femelles pondent et meurent le jour de leur émergence ou de leur venue dans le verger ;
- **aucune cécidomyie n'émerge du sous-bloc bachelé** : les larves ne peuvent s'enfouir dans le sol lorsque qu'une bache est présente, les individus en diapause ne peuvent pas émerger lorsqu'une bache est présente ;
- **les femelles exogènes viennent proportionnellement aux inflorescences présentes** ;
- **les individus en diapause sont confondus avec les individus exogènes** : il n'est (pour l'instant) pas possible de faire la distinction entre les individus qui émergent de diapause et ceux qui arrivent de l'extérieur ;
- **il y a des échanges de femelles entre les trois sous-blocs** : on considère cependant que les femelles préfèrent rester dans le sous-bloc duquel elles émergent. Si elles quittent le sous-blocs, elles s'en vont dans le sous-bloc limitrophe (le sous-bloc bachelé, en l'occurrence) ;
- **les manguiers ne sont pas impactés par la modalité de couverture du sol, le vol des cécidomyies non plus** : on considère que la modalité de couverture du sol n'impacte que la survie des individus lorsqu'il s'enfouissent dans le sol puis quand ils en émergent.

4 Modèle

Les phénomènes et hypothèses susmentionnés sont modélisés par des équations. Ainsi, le nombre de larves qui s'éjectent des inflorescences dans le sous-bloc i le jour t est déterminé par

$$L_{t,i} = N_{t-d_\ell,i} \times R \times E_0 \times \mu_\ell,$$

où $N_{t-d_\ell,i}$ représente le nombre de femelles présentes dans le sous-bloc i une durée de larvation auparavant, R le coefficient de disponibilité en ressources, E_0 le nombre d'œufs pondus par une femelle et μ_ℓ la probabilité de survie des œufs. Et on a $R = \max(1, kI_{t,i}/N_{t,i})$, avec le coefficient k à déterminer.

Le nombre de femelles émergentes du sous-bloc i au jour t est donnée par

$$N_{t,i}^{\text{emer}} = L_{t-d_p,i} \times p_p \times \mu_p \times \mu_{\text{sol}} \times \frac{1}{1 + SR},$$

avec $L_{t-d_p,i}$ le nombre de larves présentes une durée de pupaison auparavant, p_p la probabilité pour une larve d'entrer en pupaison, μ_p la probabilité de survivre à la phase de pupaison, μ_{sol} la probabilité de survivre à la couverture du sol et SR le *sex-ratio*.

La population de femelles dans le sous-bloc i au jour t est donnée par

$$N_{t,i} = N_{t,i}^{\text{endo}} + N_{t,i}^{\text{exo}} + N_{t,i}^{\text{side}},$$

où $N_{t,i}^{\text{endo}}$ représente le nombre de femelles émergentes du sous-bloc i au jour t et qui y restent, $N_{t,i}^{\text{exo}}$ le nombre de femelles exogènes au bloc venant dans le sous-bloc i le jour t et $N_{t,i}^{\text{side}}$ le nombre de femelles ayant émergées dans le(s) sous-bloc(s) j et/ou k le jour t et venant dans le sous-bloc i . Plus précisément, on a

$$N_{t,i}^{\text{exo}} = \gamma \times I_{t,i}$$

(avec γ à déterminer),

$$N_{t,i}^{\text{endo}} = \left(\frac{I_{t,i}}{I_{t,i} + I_{t,\text{lim}}} + (1 - p_m) \frac{I_{t,\text{lim}}}{I_{t,i} + I_{t,\text{lim}}} \right) N_{t,i}^{\text{emer}},$$

avec $I_{t,\text{lim}}$ et $N_{t,\text{lim}}$ désignant respectivement le total des inflorescences et des femelles présentes dans le(s) sous-bloc(s) limitrophe(s), et p_m est à déterminer. Enfin

$$N_{t,i}^{\text{side}} = p_m \frac{I_{t,j}}{I_{t,i} + I_{t,j}} N_{t,j}^{\text{emer}} \mathbf{1}_{\text{lim}(i,j)} + p_m \frac{I_{t,k}}{I_{t,i} + I_{t,k}} N_{t,k}^{\text{emer}} \mathbf{1}_{\text{lim}(i,k)},$$

où $\mathbf{1}_{\text{lim}(i,j)}$ vaut 1 si les sous-blocs i et j sont limitrophes, et 0 sinon.

5 Calibration

Après avoir implémenté le modèle, il reste 5 paramètres à calibrer :

- γ : le coefficient relatif à l'arrivée des individus exogènes ;
- p_m : le paramètre régulant les échanges entre les sous-blocs ;
- μ_{ER} : la probabilité de survie des larves à l'enherbement ras ;
- μ_{EH} : la probabilité de survie des larves à l'enherbement haut ;
- k : le coefficient relatif à la capacité d'accueil des inflorescences.

Il faut aussi noter que le modèle a besoin des inflorescences en entrée. Il a été décidé de ne pas donner les inflorescences vivantes, mais les inflorescences attractives. On émet l'hypothèse que vers la fin de sa vie, l'inflorescence n'est plus considérée comme attractive par les cécidomyies, on considère que c'est soit dû au fait que les inflorescences ont déjà trop de piqures, soit dû au fait que l'inflorescence a atteint la fin de son développement phénologique.

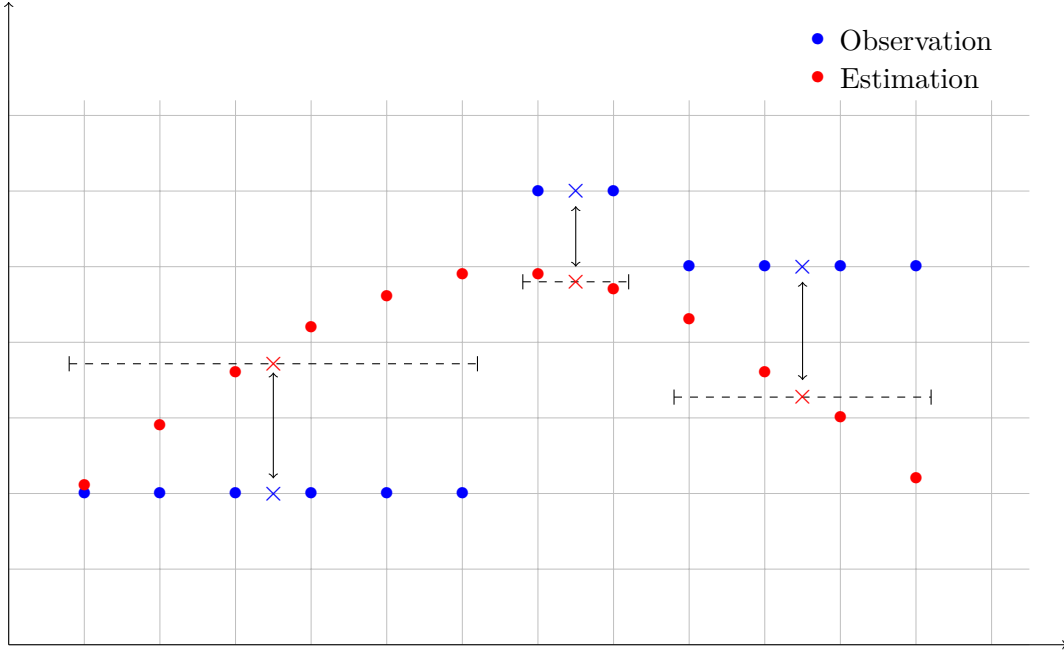


FIGURE 3 – Schéma illustrant le fonctionnement de la fonction objectif. Pour chaque période correspondant à un relevé (segments en pointillés), on calcule la moyenne des valeurs estimées (les croix rouges). On compare ensuite les moyennes ainsi calculées avec les valeurs observées associées (les croix bleues).

Fonction objectif Il y a 80 jours entre le premier jour de relevé et le dernier. Il n’y a eu cependant que 20 relevés. Le nombre de jours entre les relevés étant variable, il a été décidé de n’appliquer la fonction de coût que sur les moyennes des prédictions correspondant à un relevé (donc sur 20 valeurs). Ce faisant, on évite d’attribuer plus de poids à des observations qui découleraient d’un relevé plus espacé dans le temps avec les autres. La figure 3 illustre le fonctionnement.

Sur ces 20 valeurs, on utilise la fonction de coût MAE définie par

$$MAE(y, \hat{y}) = \frac{1}{n} \sum_n |y - \hat{y}|.$$

On le fait distinctement pour chacun des trois sous-blocs. Cela donne trois valeurs à minimiser.

Algorithme d’optimisation L’algorithme d’optimisation choisi est l’algorithme génétique multi-critère NSGA-II. Nous avons trois critères, qui correspondent à la prédiction sur chacun des trois sous-blocs.

NSGA-II ne renvoie pas une unique solution, mais un ensemble de solutions non-dominées (*i.e.* pour toutes les solutions proposées, il n’existe pas une autre solution qui soit meilleure sur les trois critères). Ainsi on choisira parmi les choix possibles le jeu de paramètre qui minimisera la somme des trois critères.

Résultats En utilisant les inflorescences attractives simulées en entrée du modèle, les paramètres trouvés sont

γ	p_m	μ_{ER}	μ_{EH}	k
0.073	0.373	0.223	0.000	10.788

Les résultats produits sont visibles sur la figure 4.

6 Analyse de sensibilité

Une analyse de sensibilité a été effectuée sur le modèle actuel. La méthode Sobol (par décomposition de la variance) a été choisie. L’implémentation utilisée est celle décrite dans le livre *Global*

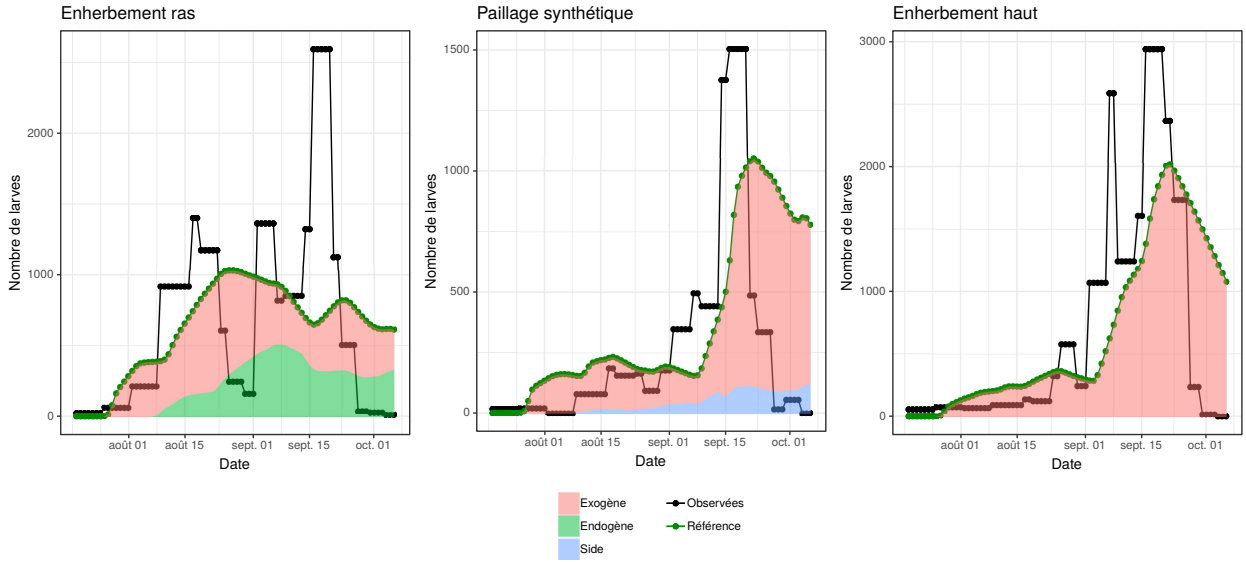


FIGURE 4 – Comparaison des nombres de larves piégées observés et simulés. La provenance des femelles ayant pondus les larves est disponible pour la simulation

Sensitivity Analysis. The Primer (Saltelli *et al.*, 2007). Les résultats sont visibles sur la figure 5. L'effet principal caractérise la part de variance attribuée au paramètre seul et l'effet total prend en compte l'impact du paramètre et de ses interactions. Il en ressort que le paramètre le plus influant sur la variance du modèle est γ (du moins pour le paillage synthétique et l'enherbement haut). Viennent ensuite les paramètres liés à la modalité de couverture du sol (cela varie logiquement selon les différents sous-blocs). Le paramètre k n'est quant à lui actif qu'à travers ses interactions.

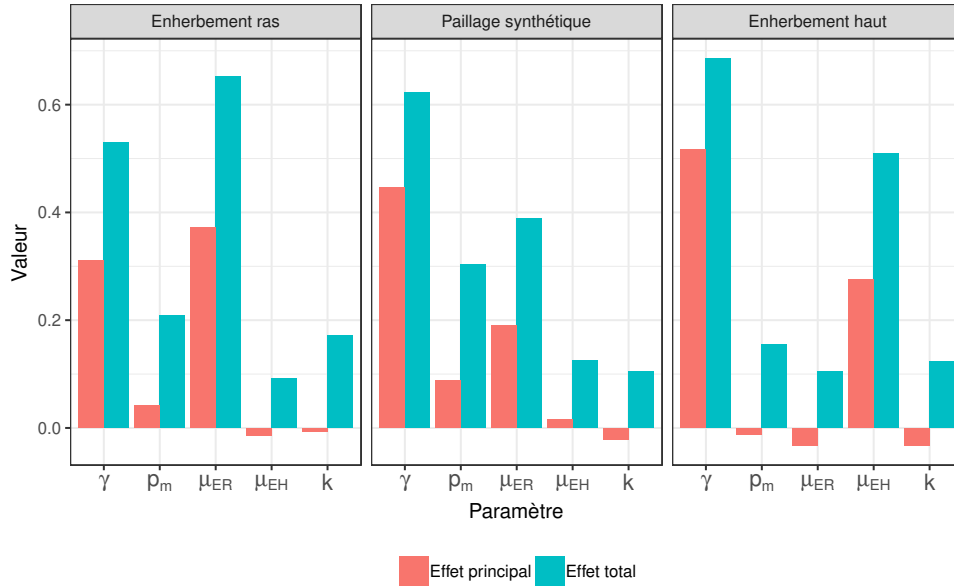


FIGURE 5 – Analyse de sensibilité du modèle. Méthode Sobol.

Pour l'analyse de sensibilité on a utilisé $N = 10000$ jeux de paramètres, avec :

- $\gamma \sim \mathcal{U}[0; 1]$;
- $p_m \sim \mathcal{U}[0; 1]$;
- $\mu_{ER} \sim \mathcal{U}[0; 1]$;
- $\mu_{EH} \sim \mathcal{U}[0; 1]$;
- $k \sim \mathcal{U}[1; 50]$.