

## ESTIMATION DE LA DISSEMINATION EFFICACE DU POLLEN DE COLZA DANS DES DISPOSITIFS DISCONTINUS

### *EFFECT OF A CROP DISCONTINUITY ON THE RAPESEED POLLEN DISPERSAL FUNCTION*

**S. HUET, H. POILLEUX**

INRA, Unité de Biométrie, Domaine de Vilvert, 78352 Jouy-en-Josas cedex

#### RESUME

La majorité des chercheurs s'accorde à penser aujourd'hui qu'il est impossible de confiner un transgène strictement dans les parcelles cultivées, même si certaines hybridations ne se produisent qu'à des fréquences très faibles. Ainsi, si l'on veut créer une filière de production agro-alimentaire sans OGM, il faut quantifier la pollution génétique d'un champ

vers un autre afin de mettre en place des consignes de culture garantissant une pureté de production satisfaisant les normes choisies. L'objectif de notre travail est d'estimer la fonction de dispersion individuelle du pollen de colza, en présence d'une discontinuité dans la culture.

#### SUMMARY

*In order to predict the pollen dispersal of a plant species under several environmental conditions, it is necessary to determine its individual pollen dispersal function. Several parametric models have already been proposed for homogeneous crop experiments. However in an agricultural landscape, a discontinuity effect on pollen flows in a cultivated area (e.g. a road crosses a field) has*

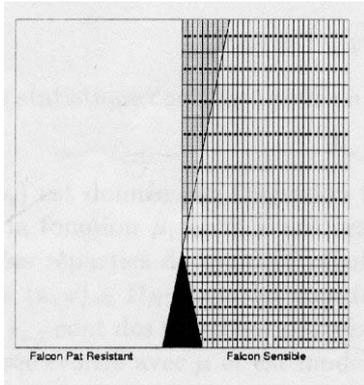
*to be taken into account. This effect is modelled and estimated: according to the size of the discontinuity, it may correspond to a significant acceleration of the pollen flow. Graphical diagnosis methods show that the modelling of the individual pollen dispersal distribution and of the discontinuity effect, best fits the data when using constant piecewise functions.*

#### DESCRIPTION DU DISPOSITIF EXPERIMENTAL

Les données dont nous disposons proviennent d'une expérimentation menée à Rennes<sup>1</sup> durant l'été 2000. Le champ est approximativement un carré de 100 m sur 100 m, partagé en 2 rectangles identiques. On note  $D_T$  le rectangle sur lequel du colza Falcon-Pat résistant (transgène pat résistant au glufosinate) a été cultivé et  $D_{NT}$ , le rectangle sur lequel du colza Falcon sensible a été cultivé. Une discontinuité en forme de triangle (base de 12 m, hauteur de 30 m) a été créée au bas de la jonction de ces deux rectangles. La partie du champ non cultivée en colza est nue.

---

<sup>1</sup> INRA Station de Génétique et Amélioration des plantes



**Figure 1** : Dispositif expérimental de Rennes

La Figure 1 illustre ce dispositif, la discontinuité étant représentée en noir. Le domaine DNT est découpé en  $N_p = 416$  secteurs de base 1.5 m et de hauteur 5 ou 10 m. Pour chaque secteur  $s$ , on observe  $k_s = 2, 3$  ou 4 répétitions de la fréquence de graines résistantes  $R_{s,j} / N_{s,j}$  où  $R_{s,j}$  est le nombre de graines résistantes parmi  $N_{s,j}$  graines prélevées dans le secteur  $s$  et semées,  $j$  variant de 1 à  $k_s$ .

## FONCTION DE DISPERSION INDIVIDUELLE ET COMPOSITION DU NUAGE POLLINIQUE

Soit  $(X, Y)$  la variable aléatoire désignant les coordonnées d'une plante pollinisée par du pollen émis par une source située en  $(0, 0)$ . On note  $g(x, y)$  la fonction de densité de ce couple, appelée aussi *dispersion efficace individuelle de pollen*. La *dispersion globale de pollen*,  $\mu(x, y)$ , i.e. la composition du nuage pollinique, est le résultat des dispersions individuelles de plusieurs plantes transgéniques ou non. Plus précisément, c'est la probabilité qu'une graine située au point  $(x, y)$  soit résistante. Nous supposons que toutes les plantes du champ ont la même fonction de dispersion individuelle de pollen. La relation entre la dispersion globale  $\mu$  et la dispersion individuelle  $g$  s'écrit alors :

$$\mu(x, y) = \frac{\sum_{(u,v) \in D_T} g(x-u, y-v)}{\sum_{(u,v) \in D_T} g(x-u, y-v) + \sum_{(u,v) \in D_{NT}} g(x-u, y-v)} \quad (1)$$

où les couples  $(u, v) \in D_T$  correspondent aux coordonnées des plantes de  $D_T$ .

Nous avons procédé en deux étapes. Dans un premier temps, nous avons testé l'ajustement de plusieurs fonctions  $g$  en restreignant les données aux secteurs du domaine  $D_{NT}$  de la zone grisée de la Figure 1. Ces secteurs sont tels que pour tout  $(u, v)$  dans  $D_T$ , les droites joignant  $(u, v)$  aux centres de chacun de ces secteurs n'interceptent pas la discontinuité représentée par le triangle noir. Dans un deuxième temps, nous modélisons l'influence de la discontinuité sur la fonction de dispersion.

## MODELISATION DE LA DISPERSION INDIVIDUELLE G SUR LE CHAMP RESTREINT

Il y a un grand choix de fonctions  $g$  possibles. Pour identifier la plus pertinente, nous nous fondons sur les résultats d'Etienne Klein (2000) et sur les données du champ restreint. Nous supposons que la fonction de dispersion individuelle est isotrope. Après avoir testé une grande diversité de fonctions et vérifié la qualité de l'ajustement par le graphique des résidus réduits, nous avons retenu le modèle suivant. La fonction de dispersion individuelle  $g$  présente une décroissance géométrique et est définie (à une constante normalisatrice près) par :

$$g(x, y) = \left( \sqrt{x^2 + y^2} + b \right)^{-a} \text{ avec } a > 2 \text{ et } b > 0 \quad (2)$$

Le modèle statistique considéré pour un secteur  $s$  de centre  $(x_s, y_s)$  et une répétition  $j$  est le suivant :

$$R_{s,j} = N_{s,j} \mu(x_s, y_s) + \varepsilon_{s,j}$$

où  $\mu(x_s, y_s)$  est donnée par l'équation 1 et la fonction  $g$  par l'équation 2. Afin de simplifier le calcul de la fonction  $\mu$ , nous avons restreint l'ensemble des couples  $(u, v)$  de  $D_T$  aux couples de coordonnées réparties de manière régulière selon un pas horizontal de 0.75 m et vertical de 0.5 m ; les couples  $(u, v) \in D_{NT}$  sont les coordonnées symétriques des couples précédents dans le domaine  $D_{NT}$ . Les  $\varepsilon_{s,j}$  sont des variables aléatoires supposées indépendantes et centrées. La variance des  $\varepsilon_{s,j}$  est supposée croître avec  $\mu$  et est modélisée par :

$$Var(\varepsilon_{s,j}) = \sigma^2 (\mu(x_s, y_s))^{0.5}$$

Les paramètres  $a, b$  décrivant la fonction  $g$  et le paramètre  $\sigma^2$  sont estimés simultanément par la méthode de quasi-vraisemblance à l'aide du logiciel **nls2** ([www.bia.inra.fr](http://www.bia.inra.fr)). Les résultats obtenus sont donnés dans le tableau suivant.

	Estimation	Ecart-type
$\hat{a}$	2.9150	0.11483
$\hat{b}$	0.2114	0.003295
$\hat{\sigma}^2$	4.174	

**Tableau 1** : paramètre de la fonction  $g$

Cette première analyse nous a permis de constater que plusieurs fonctions de décroissance sont de bons candidats pour modéliser  $g$  : la géométrique définie ci-dessus, la logistique, la fonction constante par morceaux. La décroissance exponentielle ajuste moins bien les données. Concernant la modélisation de la variance des erreurs, il apparaît très nettement que cette variance croît avec l'espérance des observations et qu'il est nécessaire d'en tenir compte dans le modèle statistique.

## MODELISATION DE L'EFFET DE LA DISCONTINUITÉ

Dans le cas où la droite reliant la plante pollinisée située en  $(x, y)$  à la plante émettrice située en  $(u, v)$  traverse une discontinuité de largeur  $L$  à la distance  $r_0$  de la source, on modélise alors la dispersion individuelle de la façon suivante :

$$\gamma(x, y) = \begin{cases} (\sqrt{x^2 + y^2} + b)^{-a} & \text{si } 0 \leq \sqrt{x^2 + y^2} < r_0 \\ 0 & \text{si } r_0 \leq \sqrt{x^2 + y^2} < r_0 + L \\ (\sqrt{x^2 + y^2} - \alpha L + b)^{-a} & \text{si } r_0 + L \leq \sqrt{x^2 + y^2} \end{cases} \quad (3)$$

Si le paramètre  $\alpha$  est égal à 0, la discontinuité n'a pas d'effet ; s'il est négatif, la discontinuité "ralentit" la dispersion du pollen, par contre s'il est positif, la discontinuité "accélère" la dispersion du pollen. On généralise ce modèle en supposant que  $\alpha$  varie avec la largeur du trou,  $L$ , de la façon suivante :

$$\alpha = \begin{cases} \alpha_1 & \text{si } 0 < L \leq 4 \\ \alpha_2 & \text{si } 4 < L \leq 8 \\ \alpha_3 & \text{si } 8 \leq L \end{cases} \quad (4)$$

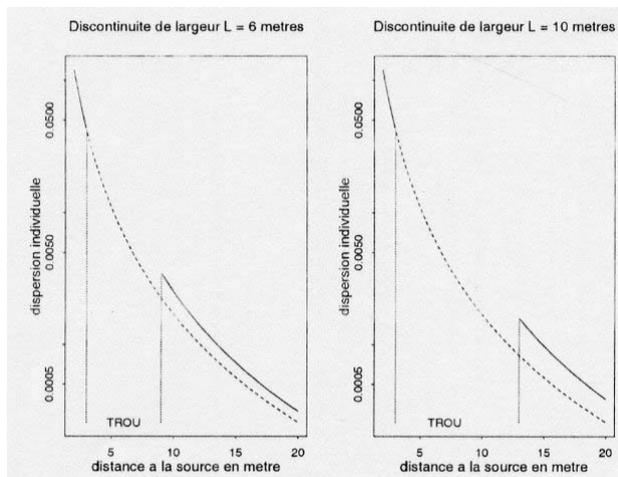
Les intervalles choisis  $[0; 4[$ ,  $[4; 8[$ ,  $[8; 30[$  différencient grossièrement les petites, moyennes et grandes discontinuités.

On considère le modèle statistique

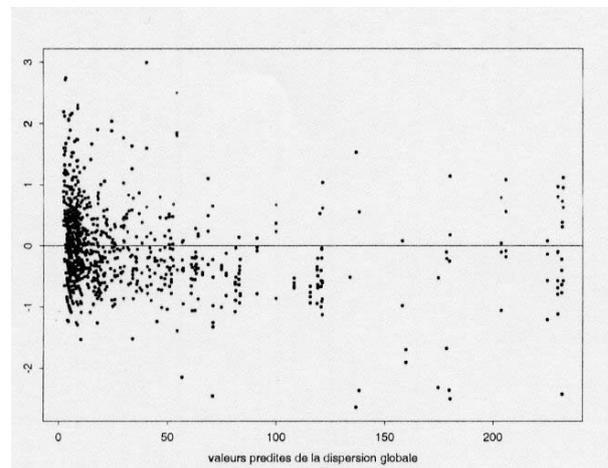
$R_{s,j} = N_{s,j} \mu(x_s, y_s) + \varepsilon_{s,j}$  où  $Var(\varepsilon_{s,j}) = \sigma^2 (N_{s,j} \mu(x_s, y_s))^{1,5}$  et  $\mu(x_s, y_s)$  est donnée par l'équation 1, la fonction  $g$  étant remplacée par la fonction  $\gamma$  définie par l'équation 3 avec  $\alpha$  vérifiant l'équation 4. On estime alors simultanément les paramètres  $a$ ,  $b$ ,  $\alpha_1$ ,  $\alpha_2$  et  $\alpha_3$ . Le tableau ci-dessous contient les résultats des estimations et leurs écarts-types.

	Estimation	Ecart-type
$a$	2.7518	0.0218
$b$	0.1657	0.0071
$\alpha_1$	-0.0714	0.1190
$\alpha_2$	0.2270	0.0790
$\alpha_3$	0.2767	0.0504

**Tableau 2 : Estimation des paramètres**



**Figure 2 : Effets de la discontinuité**



**Figure 3 : Graphique des résidus réduits**

La *Figure 2* représente les effets sur la queue de la fonction de dispersion individuelle de deux discontinuités de largeurs respectives 6 et 10 m à 3 m de la source ; ces courbes sont dessinées à partir des valeurs estimées du tableau 2. Il ne nous a pas semblé opportun de représenter l'effet d'une discontinuité de largeur inférieure à 4 m compte tenu de la faible valeur de l'estimation de  $\alpha_1$  relativement à l'écart-type estimé. La courbe en trait plein représente la dispersion individuelle estimée juste avant le trou et au-delà de  $r_0 + L$ . La courbe en pointillé représente la dispersion individuelle estimée pour les mêmes distances et en l'absence de discontinuité.

Le graphique de la *Figure 3* représente les résidus réduits

$$\frac{R_{s,j} - N_{s,j} \hat{\mu}(x_s, y_s)}{\hat{\sigma}(N_{s,j} \hat{\mu}(x_s, y_s))^{0.75}}$$

en fonction des valeurs prédites  $N_{s,j} \hat{\mu}(x_s, y_s)$  où  $\hat{\mu}$  est la valeur de  $\mu$  quand les paramètres de la fonction  $\gamma$  ont été remplacés par leurs estimateurs.

Si le modèle est correctement choisi, le graphique des résidus réduits en fonction des valeurs prédites ne doit présenter aucune structure particulière. Le graphique obtenu est, à première vue, satisfaisant. Cependant, il met en évidence une légère sous-estimation de la fonction  $\mu$  pour les faibles valeurs de  $\mu$  et surestimation pour les fortes valeurs. Cette remarque invite à généraliser la démarche que nous avons suivie dans plusieurs directions :

- estimer l'effet de la discontinuité pour un découpage plus fin de l'intervalle de variation de la largeur du trou,
- généraliser la modélisation de l'effet de la discontinuité en introduisant un paramètre  $\delta$  qui influe sur la décroissance de la dispersion, par exemple :

$$\gamma(x, y) = \begin{cases} (\sqrt{x^2 + y^2} + b)^a & \text{si } 0 \leq \sqrt{x^2 + y^2} < r_0 \\ 0 & \text{si } r_0 \leq \sqrt{x^2 + y^2} < r_0 + L \\ (\sqrt{x^2 + y^2} - \alpha L + b)^{a-\delta} & \text{si } r_0 + L \leq \sqrt{x^2 + y^2} \end{cases}$$

- reconsidérer la modélisation de la fonction de dispersion individuelle  $g$  en choisissant des modèles plus "souples". Considérer, par exemple, les fonctions constantes par morceaux revient à choisir un découpage de l'intervalle de variation de la distance à la source ( $[0;140]$  pour le champ considéré) et à supposer que  $g$  est constante sur chacun des morceaux du découpage. Ainsi les estimations de chaque palier de  $g$  sont indépendantes les unes des autres.

## REFERENCES

Klein E., 2000. *Estimation de la fonction de dispersion du pollen. Application à la dissémination de transgènes dans l'environnement*. Thèse de l'Université Paris XI – Orsay, 80 p + annexes.