

Modélisation statistique des données Aleurodicus sur agrume par des lois de Poisson pondérées

Dominique Mizère

Faculté des Sciences, Université Marien Ngouabi, BP.69, Brazzaville (Congo)

Abstract :

In this paper we have undertaken, with aid of the weighted Poisson distribution, to fit the count data collected in Republic of Congo-Brazzaville and related to a sample of spiraling whitefly *Aleurodicus dispersus Russel* (*Homoptera : aleyrodidae*) described by the preimaginal span (from eggs to adult stage) measured on days, the total number of eggs and the laying span. The linear regression issue between these three variables will be discussed as well.

Keywords : Count data, Poisson distribution, overdispersion, underdispersion, weighted Poisson model.

Resumé :

Dans ce travail nous allons, à l'aide des lois de Poisson pondérées, ajuster des données de dénombrement collectées en République du Congo-Brazzaville et relatives à un échantillon d'aleurodes *Aleurodicus dispersus Russel* (*Homoptera : aleyrodidae*) décrit par les variables : durée de développement préimaginal (de l'oeuf à l'adulte) mesurée en nombre de jours, longévité (mesurée en nombre de jours) et nombre de jours de ponte. Le problème de la régression linéaire entre ces trois variables fera aussi l'objet d'une discussion.

Mots clés : Données de dénombrement, loi de Poisson, surdispersion, sousdispersion, loi de Poisson pondérée.

1 Introduction

1.1 Matériel et présentation des données Aleurodicus

L'aleurode (ou *Aleurodicus dispersus Russel*, de l'ordre des *Homoptera*, de la famille des *aleyrodidae*) est un parasite ravageur des plantes qu'il in-

TAB. 1 – Echantillon de base de la variable DDP

valeurs de DDP	25	26	27	28	29	30	31	32	33	34	35
Effectifs observés	10	7	7	10	18	4	3	8	6	4	5

TAB. 2 – Echantillon de base de la variable LONG

Valeurs de LONG	1	2	3	4	5	6	7	8	9
Effectifs observés	29	16	22	8	2	4	0	0	1

ferste (suction de la sève, baisse de l'activité de photosynthèse, dessèchement des feuilles,...).

Des élevages expérimentaux de parasites ont été réalisés (Kiyindou, Adoumbaye, Mizère et Moussa, 1999) sur plusieurs plantes hôtes parmi lesquelles des plantes fruitières de grande consommation telles que le safoutier (*dacryodes edulis*), le manguier (*mangifera indica*), l'agrume (*citrus paradisi*) et une plante ombragère, le hura (*hura crepita*). Les plants sont âgés de 5 à 6 mois et les observations quotidiennes se font sous loupe binoculaire. Ces élevages se font à une température de 26,7°C et une humidité de 74,9%. La présente étude ne concerne que l'agrume dont le groupe comprend le citronnier, l'oranger et le pamplemoussier.

Les unités statistiques (*une unité statistique est un insecte*) sont décrites par les variables, durée de développement préimaginal (de l'oeuf à l'adulte) mesurée en nombre de jours codée DDP, longévité de l'insecte adulte mesurée en nombre de jours codée LONG et nombre de jours de ponte codée NJPO; cet insecte de couleur blanchâtre devient adulte lorsque des ailes lui poussent sur le dos. L'échantillon de l'aleurode a été partagé en deux sous-échantillons : l'un, appelé *échantillon de base*, servira à l'ajustement des modèles et l'autre, dit *échantillon test* servira à leur validation.

Quand on dispose de données nombreuses, il est conseillé de soustraire l'échantillon test à hauteur de 25% (Thria et al., 1997) et ce n'est pas le cas dans ce texte malgré la rigueur scientifique ayant prévalu à la collecte des données. Aussi, ce papier se propose-t-il de mettre en évidence les techniques statistiques de modélisation des données mais les conclusions qui en découlent n'indiquent que des tendances probantes. Les données collectées sont contenues dans les tableaux 1, 2, 3, et 4.

L'objectif des biologistes (Kiyindou et al., 1999) est d'étudier un modèle de croissance pour essayer de lutter contre ce parasite. Originaire des Îles Caraïbes et d'Amérique Centrale; il a colonisé l'Asie et l'Afrique dont le Congo-Brazzaville.

TAB. 3 – Echantillon de base de la variable NJPO

Valeurs de NJPO	0	1	2	3
Effectifs observés	48	33	0	1

TAB. 4 – Echantillon test

LONG	DDPR	NJPO
1	28	1
3	32	1
1	31	0
3	26	1
3	25	1
3	25	1
1	32	0

1.2 Le modèle de Poisson pondéré

Considérons Y une variable aléatoire de Poisson de paramètre μ et de fonction de masse $P[Y = y] = \frac{\mu^y}{y!} e^{-\mu}$, $y \in \mathbb{N}$, $\mu > 0$; la variable aléatoire entière positive Y^ω de fonction de masse

$$P[Y^\omega = y] = \frac{w(y)}{E_\mu[w(Y)]} \frac{\mu^y}{y!} e^{-\mu}, \quad y \in \mathbb{N}, \quad (1)$$

est appelée (Patil, 2002; Mizère, 2006, pp.42-44) *version pondérée* de la variable de Poisson Y où $w(y)$ est la *fonction poids*, une fonction positive et $E_\mu[w(Y)]$ la *constante de normalisation* telle que $0 < E_\mu[w(Y)] < \infty$. La fonction $w(y) = w(y; \lambda)$ peut dépendre d'un paramètre λ qui représente le mécanisme d'enregistrement des données. Notons que $w(y) = w(y; \mu, \lambda)$ peut aussi dépendre du paramètre canonique μ . La donnée d'une fonction poids permet donc de construire une loi de Poisson pondérée.

Le paramètre μ sera estimé par la *méthode du maximum de vraisemblance* tandis que λ le sera par la *méthode du profil de la vraisemblance* ou par la *méthode des moments* (Mizère, 2006, pp.67-70).

Soit $E(Z)$ et $\text{var}(Z)$ la moyenne et la variance d'une variable aléatoire entière positive, le quotient $I(Z) = \frac{\text{var}(Z)}{E(Z)}$ est appelé *indice de dispersion de Fisher*. Lorsque $I(Z) > 1$ (resp. $I(Z) < 1$), on dit que la variable Z est surdispersée (resp. sous-dispersée). Pour la variable de Poisson Y , on a $I(Y) = 1$ et on dit qu'elle est *équidispersée*. Des études ont été menées (Kokonendji, Mizère et Balakrishnan, 2007) pour caractériser les phénomènes de surdispersion et de sous-dispersion en rapport avec la fonction poids et la constante de normalisation. Pour plusieurs données de dénombrement, la variance empirique est plus grande ou petite que la moyenne empirique, ce qui rend plausible l'hypothèse que ces observations sont surdispersées ou sous-dispersées. Quand on sait que

l'on ne peut décrire des données surdispersées (sous-dispersées) par des modèles probabilistes sous-dispersées (surdispersées), nous avons (Mizère et al., 2006 : Mizère, 2006, pp.100-149) construit des tests statistiques pour mettre en évidence les phénomènes de surdispersion et de sous-dispersion.

1.3 Le contenu de l'article

Ce travail comprend quatre sections. Dans la section 2, nous allons présenter les lois de probabilités discrètes qui vont servir à l'ajustement des données Aleurodicus. Dans la section 3, nous allons présenter et interpréter les résultats des ajustements réalisés à l'aide du test d'adéquation du Chi-deux. Dans la section 4, nous allons présenter l'étude de la régression linéaire généralisée entre la variable réponse LONG et les variables explicatives DDPN et NJPO. Les résultats de cette étude permettent d'entrevoir la lutte contre ces prédateurs.

2 Une revue des lois de probabilités discrètes pour l'ajustement des données Aleurodicus

Dans notre travail antérieur (Mizère, 2006), nous avons étudié plusieurs lois de Poisson pondérées comme alternatives à la famille des lois de Poisson pour la modélisation et l'analyse statistique des données de dénombrement surdispersées ou sous-dispersées parmi lesquelles celles qui vont servir à ajuster les données Aleurodicus, notamment :

2.1 La loi $WPD_1(\mu, \lambda)$

La loi notée $WPD_1(\mu, \lambda)$ (Weighted Poisson Distribution) a pour fonction de masse

$$p_{\omega}(y; \mu, \lambda) = \frac{\left[1 + \frac{y}{\lambda+1}\right]}{\left[1 + \frac{\mu}{\lambda+1}\right]} \frac{\mu^y}{y!} e^{-\mu}; \quad y = 0, 1, 2, \dots, \lambda \geq 0. \quad (2)$$

Cette loi converge, lorsque λ tend vers l'infini, vers la loi de Poisson $P(\mu)$ de paramètre μ . C'est une loi de Poisson pondérée de fonction de masse $w(y, \lambda) = 1 + \frac{y}{\lambda+1}$ et de constante de normalisation $E_{\mu}[w(Y, \lambda)] = 1 + \frac{\mu}{\lambda+1}$; l'on a montré qu'elle est sous-dispersée (Mizère, 2006, pp.72-73).

2.2 La loi $WPD_2(\mu, \lambda)$

La loi notée $WPD_2(\mu, \lambda)$ est obtenue à partir de la loi $WPD_1(\mu, \lambda)$ par troncature en zéro, elle a pour fonction de masse

$$p_{\omega}(y; \mu, \lambda) = \frac{\left[1 + \frac{y}{\lambda+1}\right]}{\left[1 - e^{-\mu} + \frac{\mu}{\lambda+1}\right]} \frac{\mu^y}{y!} e^{-\mu}; \quad y = 1, 2, \dots, \lambda \geq 0. \quad (3)$$

Cette loi converge, lorsque λ tend vers l'infini, vers la loi de Poisson tronquée en zéro $P_o(\mu)$ de paramètre μ . C'est une loi de Poisson pondérée de fonction poids $w(y, \lambda) = \begin{cases} 0 & y = 0 \\ 1 + \frac{y}{\lambda+1} & y = 1, 2, \dots \end{cases}$ et de constante de normalisation $E_\mu[w(Y, \lambda)] = 1 - e^{-\mu} + \frac{\mu}{\lambda+1}$; l'on a montré que cette loi est sous-dispersée (Mizère, 2006, pp.76).

2.3 La loi $WPD_3(\mu, k, \lambda)$

Soit U une variable aléatoire de loi $WPD_1(\mu, \lambda)$, la variable aléatoire $V = U + k$, $k \in \mathbb{N}$ est la *translatée* de U que nous notons $WPD_3(\mu, k, \lambda)$. Tout ce que l'on a à la section 2.1 se transpose trivialement en remplaçant y par $y - k$. Cette loi converge, lorsque λ tend vers l'infini, vers la loi de Poisson translatée $PT(\mu, k)$ de paramètres μ et k . C'est une loi de Poisson pondérée de fonction poids $w(y, k, \lambda) = 1 + \frac{y-k}{\lambda+1}$; $y \geq k$ et de constante de normalisation $E_\mu[w(Y, k, \lambda)] = \mu^k \left(1 + \frac{\mu}{\lambda+1}\right)$; cette loi est sous-dispersée (Mizère, 2006, pp.75).

2.4 La loi $ZMW(\mu, \lambda, \pi_o)$

La loi notée $ZMW(\mu, \lambda, \pi_o)$ (Zero-Modified W) est la loi $WPD_1(\mu, \lambda)$ *modifiée en zéro* (Böhning, 1998, pp.841); elle a pour fonction de masse

$$p_\omega(y; \mu, \lambda, \pi_o) = \begin{cases} \pi_o & y = 0 \\ (1 - \pi_o) \frac{\left[1 + \frac{y}{\lambda+1}\right] \mu^y e^{-\mu}}{\left[1 - e^{-\mu} + \frac{\mu}{\lambda+1}\right] y!} & y = 1, 2, \dots \end{cases} \quad (4)$$

avec $0 < \pi_o < 1$. Cette loi converge, lorsque λ tend vers l'infini, vers la loi de Poisson modifiée en zéro $ZMP(\lambda, \pi_o)$. Comme loi de Poisson pondérée, la loi $ZMW(\mu, \lambda, \pi_o)$ admet pour fonction poids

$$w(y; \mu, \lambda, \pi_o) = \left(1 + \frac{y}{\lambda+1}\right) \left\{ \frac{\pi_o}{1 - \pi_o} \frac{1 - e^{-\mu} + \frac{\mu}{\lambda+1}}{e^{-\mu}} \right\}^{\delta_o(y)}, \quad y = 0, 1, 2, \dots$$

qui dépend du paramètre canonique μ et pour constante de normalisation

$$E_\mu[w(Y, \mu, \lambda, \pi_o)] = \frac{1 - e^{-\mu} + \frac{\mu}{\lambda+1}}{1 - \pi_o}$$

avec $\delta_o(y) = \begin{cases} 1 & y = 0 \\ 0 & y \neq 0 \end{cases}$ la fonction indicatrice de $\{0\}$. En posant $p(0; \mu, \lambda) = \frac{e^{-\mu}}{1 + \frac{\mu}{\lambda+1}}$ et

$$p_{ZMW}(0; \mu, \lambda) = \frac{1 - p(0; \mu, \lambda)}{(\lambda + \mu + 2)^2} + p(0; \mu, \lambda), \quad (5)$$

l'on a montré (Mizère, 2006, pp.78-80) que cette loi est surdispersée (resp. sous-dispersée) lorsque $\pi_o > p_{ZMW}(0; \mu, \lambda)$ (resp. $\pi_o < p_{ZMW}(0; \mu, \lambda)$).

3 Ajustement des données aux lois de probabilité discrètes

3.1 Graphes des lois empiriques et théoriques

Nous avons réalisé dans un premier temps les diagrammes en bâtons comparatifs des lois empiriques et théoriques (cf. figure 1). Il ressort de ces graphes la présomption que les données des variables DDP, LONG et NJPO sont compatibles avec les lois respectives de Poisson translatée, de Poisson tronquée en zéro et de Poisson.

3.2 Ajustement des données à l'aide du test d'adéquation du Chi-deux

3.2.1 La variable NJPO

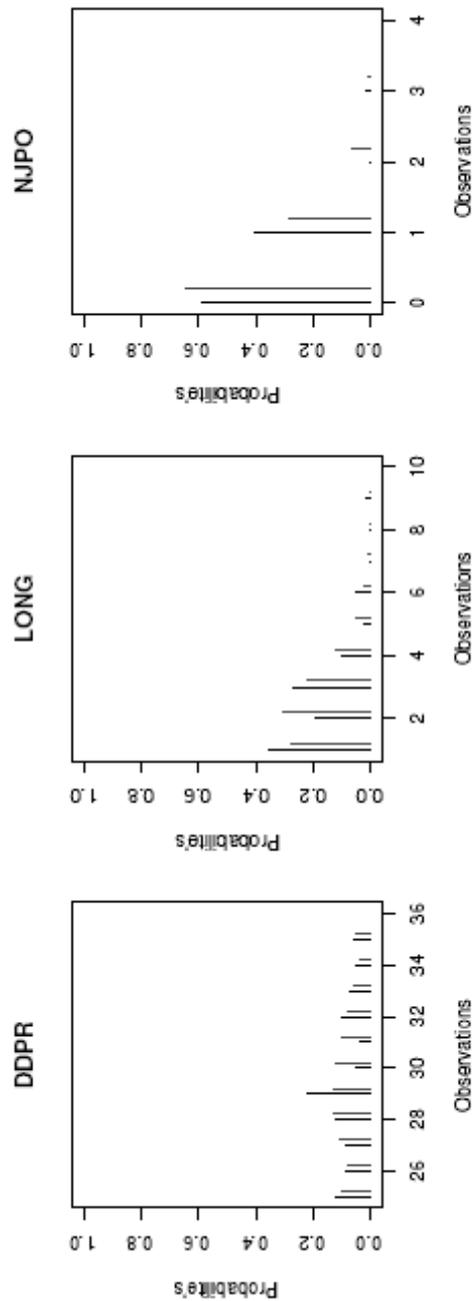
Les observations de la variable NJPO appartiennent à l'ensemble $\{0, 1, 2, 3\}$ contenant zéro. Elles ont pour statistiques élémentaires $\bar{y} = 0.4390244$ et $s'^2 = 0.3233966$ et pour indice de dispersion Fisher $I(Y^\omega) = 0.73 < 1$. Nous avons la présomption que ces observations sont sous-dispersées ; elles ne peuvent donc être ajustées que par des lois sous-dispersées de support égal à l'ensemble \mathbb{N} . C'est le cas de la loi $WPD_1(\mu, \lambda)$ et de sa loi limite, la loi de Poisson $P(\mu)$. L'estimateur du profil de la vraisemblance de λ est égal à $\hat{\lambda} = 0$. Le nombre de degré de liberté de la statistique χ^2 noté d.d.l est égal à $l - m - 1$ où l est le nombre de classes et m le nombre de paramètres estimés par la méthode du maximum de vraisemblance (Saporta, 1978).

Au risque $\alpha = 0.05$ de première espèce, il ressort du tableau 5, que les modèles $WPD_1(\mu, \lambda)$ et $P(\mu)$ sont rejetées, ce qui infirme notre présomption sur l'examen des graphes des distributions empiriques et théoriques (cf. figure 1). Ce rejet est très certainement dû à une *déflation de zéro* qui caractérise le phénomène de sous-dispersion. En effet, pour les lois $ZMP(\mu, \pi_o)$ et $ZMW(\mu, 30, \pi_o)$, la proportion observée de zéros $\hat{\pi}_o = \frac{48}{82} = 0.585$ est inférieure respectivement à la probabilité estimée de zéro $e^{-\bar{y}} = 0.644$ et à $p_{ZMW}(0; \bar{y}, 30) = 0.636$ (cf. expression 5). Pour la loi $ZMW(\mu, 0, \pi_o)$, on a $p_{ZMW}(0; \bar{y}, 0) = 0.540 \simeq \hat{\pi}_o$. Toujours au risque $\alpha = 0.05$ de première espèce, il ressort du tableau 6, que l'on ne peut rejeter l'idée que les données proviennent des lois $ZMP(\mu, \hat{\pi}_o)$ et $ZMW(\mu, 30, \hat{\pi}_o)$.

3.2.2 La variable LONG

Les observations de la variable LONG appartiennent à l'ensemble $\{k \in \mathbb{N}/1 \leq k \leq 9\}$. Elles ont pour statistiques élémentaires $\bar{y} = 2.463415$ et $s'^2 = 2.424571$ et pour indice de dispersion de Fisher $I(Y^\omega) = 0.98 < 1$. Nous avons la présomption que ces observations sont sous-dispersées. par définition même, la longévité ne peut être que strictement positive, c'est pour cela que ces données ne peuvent être ajustées que par des lois sous-dispersées de support égal à l'ensemble \mathbb{N}^* .

FIG. 1 – Distributions empiriques (bâtons de gauche) et distributions théoriques (bâtons de droite)



TAB. 5 – Ajustement des données aux lois $WPD_1(\mu, \lambda)$ et $P(\mu)$

NJPO	Eff. obs	$WPD_1(\mu, 0)$	$WPD_1(\mu, 30)$	$P(\mu)$
0	48	51.707	52.857	52.862
1	33	25.163	23.215	23.207
2	0	4.592	5.093	5.094
3	1	0.536	0.833	0.835
	$\hat{\mu}$	0.243	0.425	0.439
	χ^2	6.030	8.665	8.676
	d.d.l	1	1	1
	p-valeur	0.014	0.003	0.003



TAB. 6 – Ajustement des données aux lois $ZMW(\mu, \lambda, \hat{\pi}_o)$ et $ZMP(\mu, \hat{\pi}_o)$

NJPO	Eff. obs	$ZMW(\mu, 0, \hat{\pi}_o)$	$ZMW(\mu, 30, \hat{\pi}_o)$	$ZMP(\mu, \hat{\pi}_o)$
0	48	48	48	48
1	33	32.067	32.075	32.075
2	0	1.866	1.851	1.851
3	1	0.065	0.073	0.071
	$\hat{\mu}$	0.077	0.111	0.115
	χ^2	0	0	0
	d.d.l	1	1	1
	p-valeur	1	1	1

C'est le cas de la loi $WPD_2(\mu, \lambda)$ et de sa loi limite $P_o(\mu)$. L'estimateur du profil de la vraisemblance de λ est égal à $\hat{\lambda} = 30$.

Au risque $\alpha = 0.05$ de première espèce, il ressort du tableau 7, que l'on ne peut rejeter l'idée que les données proviennent des lois $WPD_2(\mu, 30)$ et $P_o(\mu)$.

3.2.3 La variable DDPR

Les observations de la variable DDPR sont très grandes et appartiennent à l'ensemble $\{k \in \mathbb{N}/25 \leq k \leq 35\}$. Elles ont pour statistiques élémentaires $\bar{y} = 29.28049$ et $s'^2 = 8.870973$ et pour indice de dispersion de Fisher $I(Y^\omega) = 0.30 < 1$. Nous avons la présomption que ces observations sont sous-dispersées; elles ne peuvent donc être ajustées que par des modèles censurées à gauche et sous-dispersées. C'est le cas de la loi $WPD_3(\mu, k, \lambda)$ et de sa loi limite $PT(\mu, k)$. Les estimations par la méthode des moments du paramètre k et par la méthode du maximum de vraisemblance de μ de la loi de Poisson translatée, donnent les résultats suivants : $\hat{k} = 20$ et $\hat{\mu} = 9.28049$. Tandis que l'estimateur du profil de la vraisemblance de k est égal à $\hat{k} = 21$ avec $\hat{\mu} = 8.280488$, résultats qui diffèrent de ceux obtenus par la méthode des moments; l'estimateur du profil de la vraisemblance de λ est égal à $\hat{\lambda} = 30$.

Au risque $\alpha = 0.05$ de première espèce, il ressort du tableau 8 que l'on ne peut rejeter l'idée que les données proviennent des lois $WPD_3(\mu, 20, 30)$ et $PT(\mu, 20)$. On peut conclure qu'à partir du rang $\lambda_o = 30$, le modèle $WPD_3(\mu, 20, 30)$ est une approximation de $PT(\mu, 20)$. Toujours au risque $\alpha = 0.05$ de première espèce, il ressort du tableau 9 que toutes les lois sont rejetées. En réalité, ce qui est rejeté, c'est la valeur $\hat{k} = 21$ qui ne permet pas de valider au moins l'une des lois. l'estimateur des moments de k est meilleur que l'estimateur du profil de la vraisemblance, dans le sens où il permet de valider au moins un modèle.

La valeur \hat{k} est interprétée par les biologistes comme la durée minimale pour un parasite de devenir adulte; la plante hôte ayant la plus petite valeur \hat{k} (à supposer que l'étude porte sur plusieurs plantes hôtes (Mizère et al., 2007) est celle qui est la plus favorable au développement du parasite.

4 Etude de la dépendance linéaire

4.1 Test d'indépendance entre les variables

4.1.1 Variables DDPR et NJPO

En regroupant les valeurs de DDPR dans deux classes $a1 = \{25, 26, 27, 28, 29\}$ et $a2 = \{30, 31, 32, 33, 34, 35\}$ et celles de NJPO dans deux classes $b1 = \{0\}$ et $b2 = \{1, 2, 3\}$, les tris croisés des deux variables sont contenus dans le tableau 10. Le test d'indépendance du Chi-deux donne les résultats suivants : $\chi^2 = 0.2438$, d.d.l = 1 et p - valeur = 0.6215. Au risque $\alpha = 0.05$ de première espèce, les variables DDPR et NJPO sont indépendantes.

TAB. 7 – Ajustement des données aux lois $WPD_2(\mu, \lambda)$ et $P_o(\mu)$

LONG	Eff. obs	$WPD_2(\mu, 0)$	$WPD_2(\mu, 30)$	$P_o(\mu)$
1	29	20.887921	22.658856	22.678580
2	16	26.065148	24.808966	24.797292
3	22	19.274424	18.092105	18.075961
4	8	10.021556	9.886782	9.882341
5	2	4.001752	4.318731	4.322233
6	4	1.294641	1.570872	1.575344
7	0	0.351679	0.489397	0.421479
8	0	0.082283	0.133318	0.134531
9	1	0.020594	0.040972	0.041569
	$\hat{\mu}$	1.663810	2.123424	2.186847
	χ^2	8.101609	6.13703	6.123134
	d.d.l	3	3	3
	p-valeur	0.04395777	0.1051304	0.1057708

TAB. 8 – Ajustement des données aux lois $WPD_3(\mu, 20, \lambda)$ et $PT(\mu, 20)$

DDPR	Eff. obs	$WPD_3(\mu, 20, 0)$	$WPD_3(\mu, 20, 30)$	$PT(\mu, 20)$
25	10	7.208	8.120	8.174
26	7	6.733	6.775	6.783
27	7	9.220	9.000	8.993
28	10	10.875	10.454	10.432
29	18	11.260	10.787	10.757
30	4	10.388	10.011	9.983
31	3	8.640	8.441	8.423
32	8	6.542	6.521	6.514
33	6	4.545	4.647	4.650
34	4	2.917	3.074	3.082
35	5	3.660	4.165	4.204
	$\hat{\mu}$	8.387	9.054	9.280
	χ^2	15.012	13.992	13.952
	d.d.l	7	7	7
	p-valeur	0.035	0.051	0.052

TAB. 9 – Ajustement des données aux lois $WPD_3(\mu, 21, \lambda)$ et $PT(\mu, 21)$

DDPR	Eff. obs	$WPD_3(\mu, 21, 0)$	$WPD_3(\mu, 21, 30)$	$PT(\mu, 21)$
25	10	5.925	6.894	6.942
26	7	6.621	6.730	6.741
27	7	9.527	9.307	9.303
28	10	11.509	11.025	11.005
29	18	11.976	11.420	11.391
30	4	10.941	10.508	10.480
31	3	8.905	8.696	8.678
32	8	6.535	6.538	6.532
33	6	4.365	4.504	4.507
34	4	2.676	2.862	2.871
35	5	3.015	3.511	3.545
	$\hat{\mu}$	7.399	8.073	8.280
	χ^2	17.797	15.517	15.425
	d.d.l	7	7	7
	p-valeur	0.012	0.029	0.030

TAB. 10 – Tableau de contingence entre les variables DDPR et NJPO

	b1	b2
a1	32	20
a2	16	14

4.1.2 Variables DDPR et LONG

En regroupant les valeurs de LONG dans deux classes $c1 = \{1, 2\}$ et $c2 = \{3, 4, 5, 6, 7, 8, 9\}$, les tris croisés des deux variables sont contenus dans le tableau 11. Le test d'indépendance du Chi-deux donne les résultats suivants : $\chi^2 = 10.2931$, d.d.l = 1 et p – valeur = 0.001335. Au risque $\alpha = 0.05$ de première espèce, les variables DDPR et LONG sont dépendantes.

4.1.3 Variables NJPO et LONG

Le tableau 12 contient les tris croisés des variables NJPO et LONG. Le test d'indépendance du Chi-deux donne les résultats suivants : $\chi^2 = 10.3981$, d.d.l = 1 et p – valeur = 0.001261. Au risque $\alpha = 0.05$ de première espèce, les variables DDPR et LONG sont dépendantes.

TAB. 11 – Tableau de contingence entre les variables DDPR et LONG

	c1	c2
a1	36	16
a2	9	21

TAB. 12 – Tableau de contingence entre les variables NJPO et LONG

	c1	c2
b1	34	14
b2	11	23

4.2 Etude de la variable LONG par régression

4.2.1 Définition et ajustement du modèle

Il ressort de l'étude sur l'indépendance entre les variables DDPR, NJPO et LONG, ce qui suit : la variable LONG est dépendante des variables DDPR et NJPO tandis que ces dernières sont indépendantes. Comme la variable LONG suit une loi de Poisson tronquée en zéro (cf. section 3.2.2) d'espérance mathématique égale à

$$E(\text{LONG}) = \frac{\mu}{1 - e^{-\mu}}, \quad (6)$$

nous proposons une modélisation du paramètre μ suivant l'équation

$$\ln(\mu) = \beta_1 + \beta_2 \text{DDPR} + \beta_3 \text{NJPO} \quad (7)$$

dans laquelle $\ln(\cdot)$ est appelée *fonction lien* (Mc Cullag et Nelder, 1989). La taille de l'échantillon de base est égale à $n = 82$, le nombre de paramètres à $p = 3$ et le nombre de degré de liberté à d.d.l = $n - p = 79$. Les estimations $\hat{\beta}_j$ des paramètres du modèle par la méthode du maximum de vraisemblance sont contenues dans le tableau 13. Et au risque $\alpha = 0.05$ de première espèce, il ressort du tableau 13 que tous les coefficients du modèles sont significativement non nuls : les variables DDPR et NJPO ont de l'effet sur la variable LONG. La statistique de Pearson est égale à $X^2 = 81.36255$ et le quotient $\frac{X^2}{n-p} = 1.02 \approx 1$ permettent de conclure à la forte présomption que ce modèle dit *modèle linéaire généralisé* (Mc Cullag et Nelder, 1989 ; Mizère, 2006, pp.95-99) est compatible avec les données.

4.2.2 Validation du modèle

Les valeurs prédites μ_i du paramètre μ sur l'échantillon test à travers l'expression (7) sont contenues dans le tableau 14, leur moyenne $\bar{\mu} = 2.191847$ est sensiblement égale à l'estimation $\hat{\mu} = 2.186847$ de μ obtenue sur l'échantillon de

TAB. 13 – Tableau des coefficients du modèle

Variabes	$\hat{\beta}$	$s_{\hat{\beta}}$	$t = \frac{\hat{\beta}}{s_{\hat{\beta}}}$	$P (> t)$
Constante	-1.45343	0.814738	-1.783923	0.0391
DDPR	0.06539	0.027387	2.387629	0.0096
NJPO	0.50705	0.111151	4.561812	$9.15e^{-06}$

TAB. 14 – Valeurs prédites μ_i de μ sur l'échantillon test

μ_i
2.42187
3.1459
1.774751
2.124975
1.990468
1.990468
1.89468
$\bar{\mu}=2.191847$

base de la variable LONG suivant la loi de Poisson tronquée en zéro : le modèle peut donc être considéré comme validé.

4.2.3 La lutte anti-parasite

Etant donné que sur l'agrume, la durée de développement minimale du parasite est de $\hat{k} = 20$ jours, la lutte contre l'aleurode doit se focaliser sur sa non reproduction.

Références

- [1] Böhning D. (1998). Zero-Inflated Poisson Models and C.A.M.A.N : a Tutorial Collection of Evidence. *Biometrical Journal*, 40,7,833-843,
- [2] Kiyindou A., Adoumbaye I.P., Mizère D. et Moussa J.B. (1999). Influence de la plante hôte sur le développement et la reproduction de l'aleurode Aleurodicus dispersus Russel (Hom : Aleyrodidae) en République du Congo. *Fruit*, vol.54, 115-122,
- [3] Kokonendji C.C., Mizère D. et Balakrishnan N. (2007). Connections of the Poisson weight function to overdispersion and underdispersion. Accepté pour publication dans *J.Statist. Plann. Inference*, doi : 10.1016/j.jspi.2007.05.028,
- [4] Mc Cullagh P. et Nelder T.A. (1989). Generalized linear model. *second Edition, Chapman & Hall*, London,

- [5] Mizère D. (2006). Contribution à la modélisation et à l'analyse statistique des données de dénombrement. *Thèse de Doctorat* présentée à l'université de Pau et des Pays de l'Adour (UPPA), France, le 27 janvier ,
- [6] Mizère D., Kokonendji C.C. et Dossou-Gbété S. (2006). Quelques tests de la loi de Poisson contre des alternatives générales basés sur l'indice de dispersion de Fisher. *Rev. statistique Appliquée*, LIV(4),59-82,
- [7] Mizère D., Kissita G. et Kiyindou A. (2007). Etude statistique de la plante hôte sur le développement de l'aleurode A. Dispersus par la loi de Poisson translatée. *Prépublication* Université Marien Ngouabi,
- [8] Patil G.P. (2002). Weighted distributions. *In Encyclopedia of Environmetrics* (eds A.H. El-Shaarawi and W.W. Piegorsch), vol4,2369-2377,
- [9] Saporta G. (1978). Théorie et méthodes de la statistique. *Edition Technip*, Paris,
- [10] Thria S., Lechevallier Y., Gascuel O. et Canu S. (1997). Statistique et méthodes neuronales. *Dunod*.

