

UNIVERSITÉ
LAVAL



Ressources naturelles
Canada

Natural Resources
Canada

ESSAI DE MAÎTRISE

Application des modèles linéaires généralisés mixtes à un problème forestier

Présenté par Oumar Samba LY

Sous la direction de Christian GENEST

Et la codirection de Michèle BERNIER-CARDOU

REMERCIEMENTS

Je tiens à remercier très sincèrement Mme Michèle Bernier-Cardou et M. Christian Genest pour avoir dirigé mes travaux de recherche.

Il m'est loisible d'adresser mes chaleureux remerciements et toutes mes reconnaissances à M. Genest, qui n'a ménagé aucun effort pour un encadrement privilégié de mes travaux. Sa disponibilité, sa courtoisie et son sens de l'humour (c'est important !) m'ont mis à l'aise et dans des conditions idéales pour mener à bien mes recherches.

Toute ma profonde gratitude à Mme Bernier-Cardou pour sa disponibilité et pour tout ce que j'ai appris auprès d'elle. Son expérience et ses conseils m'ont été d'un apport immense. Je lui dois, entre autres, l'approfondissement de la connaissance des modèles mixtes, de la programmation SAS et plus particulièrement de la macro GLIMMIX, sans oublier les bonnes conditions de travail.

De façon générale, je remercie aussi tous ceux qui, de près ou de loin, ont contribué à la réalisation de cet essai.

Une partie du financement nécessaire à la réalisation de cet essai a été fournie par le Conseil de recherches en sciences naturelles et en génie du Canada, ainsi que par le Fonds québécois de la recherche sur la nature et les technologies, par le biais de subventions de recherche individuelles et collectives attribuées à M. Genest et aux membres de son équipe. Qu'ils trouvent ici l'expression de ma gratitude.

Le service canadien des forêts-Québec a aussi contribué financièrement à la réalisation de cet essai. Tous mes remerciements.

TABLE DES MATIÈRES

REMERCIEMENTS	2
TABLE DES MATIÈRES	3
INTRODUCTION	4
I OBJECTIFS DE RECHERCHE	6
II THÉORIE DES MODÈLES LINÉAIRES GÉNÉRALISÉS MIXTES AVEC LIENS CANONIQUES BINOMIAL OU DE POISSON	7
II.1 MODÈLES LINÉAIRES GÉNÉRALISÉS	7
II.2 MODÈLES LINÉAIRES GÉNÉRALISÉS MIXTES	9
II.3 PARAMÈTRES DE DISPERSION ET FONCTION DE DÉVIANCE	12
III EXEMPLE D'APPLICATION	14
III.1 EXPÉRIENCE.....	16
<i>III.1.1 Plan d'expérience et collecte des données</i>	16
<i>III.1.2 Présentation des données</i>	18
<i>III.1.3 Manipulation des fichiers de données</i>	19
III.2 MÉTHODES.....	21
<i>III.2.1 La macro GLIMMIX de SAS</i>	21
<i>III.2.2 Modèle complet</i>	23
<i>III.2.3 Calcul des nombres de degré de liberté (d.d.l.)</i>	25
<i>III.2.4 Choix de la structure de variance pour les mesures répétées</i>	26
<i>III.2.5 Réduction de la partie aléatoire</i>	27
<i>III.2.6 Statistiques d'ajustement</i>	29
<i>III.2.7 Analyse de la partie fixe</i>	29
<i>III.2.8 Décomposition de la partie fixe à l'aide de contrastes</i>	31
<i>III.2.9 Estimation des moyennes pertinentes à l'interprétation des contrastes</i>	33
III.3 RÉSULTATS ET CONCLUSIONS	36
CONCLUSION	39
BIBLIOGRAPHIE	41
ANNEXES	42
TABLEAUX.....	43
FIGURES.....	56
PROGRAMME SAS	59

INTRODUCTION

Ce document est un essai pour l'obtention du diplôme de maîtrise en statistique. Son thème général est l'étude des modèles linéaires généralisés mixtes (GLMM) avec mesures répétées. Ces modèles seront présentés et utilisés ici pour étudier un problème de sylviculture et son impact écologique.

Les modèles linéaires généralisés sont des techniques dont la théorie est riche et qui font l'objet d'une large utilisation. Outre la foresterie, on les applique à des domaines aussi variés que la médecine, la finance, les sciences économiques et industrielles, entre autres. Ces techniques sont d'autant plus intéressantes qu'elles permettent d'analyser des effets mixtes et des mesures répétées et qu'elles se prêtent à l'emploi de différentes lois de probabilité pour les termes d'erreur du modèle, en autant que ces distributions fassent partie de la famille exponentielle. Dans le cadre de cet essai, on fera appel à ces modèles pour l'analyse de la biodiversité d'insectes de la famille des carabidés.

Au printemps 2000, un projet de recherche multidisciplinaire intitulé «Systèmes sylvicoles adaptés à la forêt mélangée» (SSAM) a été entrepris dans une sapinière à bouleau jaune de la région de Portneuf, au nord-ouest de la ville de Québec. Ce projet visait à promouvoir une gestion à la fois rentable et rationnelle de la forêt québécoise. Des méthodes sylvicoles appropriées ont donc été recherchées. Pour ce faire, il était important d'évaluer l'impact de perturbations naturelles ou artificielles susceptibles d'affecter les peuplements forestiers et leur biodiversité. Parmi ces perturbations, on peut citer les coupes forestières, les feux de forêts, l'utilisation de pratiques et de produits agricoles et/ou sylvicoles tels que les pesticides et l'aménagement des sols. De façon plus spécifique, des recherches ont été entreprises afin d'étudier l'effet de l'éclaircie forestière et la scarification des sols sur la biodiversité des insectes. Ce sont les données de cette étude qui feront l'objet d'une analyse dans cet essai. L'étude sera cependant restreinte aux spécimens de diverses espèces de la famille des carabidés, laquelle appartient à l'ordre des coléoptères.

Le document comporte quatre parties importantes. La première présente les objectifs de recherche. La deuxième est consacrée à une présentation succincte de la théorie des modèles

linéaires généralisés mixtes avec mesures répétées dont les termes d'erreur suivent une loi binomiale ou de Poisson. L'application de ces modèles à l'étude de la biodiversité des carabidés fait l'objet de la troisième partie, elle-même subdivisée en trois sections, à savoir la description de l'expérience, la présentation et l'application des méthodes aux données recueillies, ainsi que la discussion des résultats de l'analyse. Une brève conclusion constitue la quatrième partie.

I. OBJECTIFS DE RECHERCHE

Nous visons deux objectifs principaux dans cet essai, à savoir :

- 1- Présenter succinctement la théorie des modèles linéaires généralisés mixtes.
- 2- Exposer en détail une application concrète de ces modèles et s'en servir pour répondre à des questions pratiques soulevées dans le cadre du projet SSAM.

Plus spécifiquement, l'objectif de l'application considérée est de comparer certaines pratiques sylvicoles et d'évaluer leur impact sur la biodiversité et l'entomofaune en particulier. Nous verrons dans la suite comment l'emploi des modèles linéaires généralisés mixtes permet de juger en quoi l'abondance des carabidés dans la forêt d'origine est modifiée par différentes coupes forestières et la scarification des sols. Il s'agit en clair de déterminer si ces divers traitements, leur ampleur et leur interaction ont un impact sur l'abondance des carabidés. Le cas échéant, on aimerait savoir comment cet impact se manifeste au cours d'une saison de croissance.

Afin de répondre à ces interrogations, on fera d'abord appel à des modèles d'analyse de la biodiversité des carabidés. Puis, on cherchera à modéliser la probabilité de présence-absence ainsi que l'abondance de l'espèce la plus fréquente. Les questions plus précises auxquelles nous tenterons de répondre seront présentées dans la partie application.

II. THÉORIE DES MODÈLES LINÉAIRES GÉNÉRALISÉS MIXTES AVEC LIENS CANONIQUES BINOMIAL OU DE POISSON

Ces lignes suivent les présentations de Littell *et al.* (1996) et Brown et Prescott (1999).

II.1 MODÈLES LINÉAIRES GÉNÉRALISÉS

Soit Y une variable aléatoire de la famille exponentielle, et \mathbf{y} un vecteur $n \times 1$ d'observations de cette variable aléatoire tel que

$$\begin{aligned}\mathbf{y} &= \boldsymbol{\mu} + \boldsymbol{\varepsilon} \\ E(\mathbf{y}) &= \boldsymbol{\mu} \\ \text{var}(\mathbf{y}) &= \text{var}(\boldsymbol{\varepsilon}) = \mathbf{R}\end{aligned}$$

où \mathbf{R} est une matrice diagonale car on suppose les y_i mutuellement indépendants. Soit \mathbf{X} une matrice dont les lignes sont des vecteurs $p \times 1$ transposés de variables explicatives d' y_i , $i = 1, \dots, n$. Un modèle linéaire généralisé pour \mathbf{y} est

$$\boldsymbol{\eta} = g(\boldsymbol{\mu}) = \mathbf{X}\boldsymbol{\beta}$$

où $g(\cdot)$ est une fonction monotone appelée « fonction de lien », elle a un inverse que nous noterons $h(\cdot) = g^{-1}(\cdot)$, et $\boldsymbol{\beta}$ est le vecteur $p \times 1$ des paramètres associés aux p variables explicatives, respectivement. La log-vraisemblance d' \mathbf{y} est

$$l(\boldsymbol{\theta}, \boldsymbol{\phi}; \mathbf{y}) = \sum_{i=1}^n \left[\frac{y_i \theta_i - b(\theta_i)}{a(\phi_i)} + c(y_i, \phi_i) \right]$$

où θ_i est le paramètre canonique de la densité d' Y , $b(\theta_i)$, $a(\phi_i)$ et $c(y_i, \phi_i)$ sont des fonctions connues de θ_i , ϕ_i et (y_i, ϕ_i) respectivement. La fonction de lien canonique est donnée par $g(\cdot) = b^{-1}(\cdot)$. Ce sont les seules fonctions de lien que nous utiliserons dans cet essai.

La matrice $\mathbf{R} = \mathbf{A}\mathbf{B}$ où $\mathbf{A} = \text{Diag}(a(\phi_1), \dots, a(\phi_n))$ et $\mathbf{B} = \text{Diag}(b''(\theta_1), \dots, b''(\theta_n)) = \text{Diag}(g^{-1}(\mu_1), \dots, g^{-1}(\mu_n))$ où $b''(\cdot)$ dénote la dérivée seconde de $b(\cdot)$ par rapport à θ_i et $g^{-1}(\cdot)$, la dérivée de l'inverse de la fonction de lien par rapport à μ_i .
Sous forme matricielle, la log-vraisemblance s'écrit:

$$l(\boldsymbol{\theta}, \boldsymbol{\varphi}; \mathbf{y}) = \mathbf{y}' \mathbf{A}^{-1} \boldsymbol{\theta} - (\mathbf{b}_0^{1/2})' \mathbf{A}^{-1} (\mathbf{b}_0^{1/2}) + \mathbf{1}' \mathbf{c}$$

où

$$\begin{aligned} \boldsymbol{\theta} &= (\theta_1, \theta_2, \dots, \theta_n)' \\ \mathbf{b}_0 &= (b(\theta_1), b(\theta_2), \dots, b(\theta_n))' \\ \mathbf{c} &= (c(y_1, \phi_1), c(y_2, \phi_2), \dots, c(y_n, \phi_n))' \end{aligned}$$

$\mathbf{1}$ est un vecteur $n \times 1$ de 1, et l'exposant $1/2$ du vecteur \mathbf{b}_0 indique un vecteur constitué des racines carrées des éléments de \mathbf{b}_0 .

Le vecteur des paramètres $\boldsymbol{\beta}$ est estimé en maximisant cette log-vraisemblance par rapport à $\boldsymbol{\beta}$. Nelder et Wedderburn (1972) ont montré que cette maximisation était équivalente à résoudre le système d'équations linéaires suivant de façon itérative :

$$\mathbf{X}'(\mathbf{B}\mathbf{R}^{-1}\mathbf{B})\mathbf{X}\boldsymbol{\beta} = \mathbf{X}'(\mathbf{B}\mathbf{R}^{-1}\mathbf{B})\mathbf{y}^*$$

où $\mathbf{y}^* = \boldsymbol{\eta} + (\mathbf{y} - \boldsymbol{\mu})\mathbf{B}^{-1}$. À la dernière itération, $\text{Var}(\mathbf{y}^*) = \mathbf{A}\mathbf{B}^{-1}$, $\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{A}^{-1}\mathbf{B}\mathbf{X})^{-1} \mathbf{X}'\mathbf{A}^{-1}\mathbf{B}\mathbf{y}^*$ et $\text{Var}(\hat{\boldsymbol{\beta}}) = (\mathbf{X}'\mathbf{A}^{-1}\mathbf{B}\mathbf{X})^{-1}$. \mathbf{B} et \mathbf{A} sont remplacés par leurs estimations. L'inverse de la fonction de lien permet d'estimer $\boldsymbol{\mu}$ à l'aide de $\mathbf{X}\hat{\boldsymbol{\beta}}$:

$$\hat{\boldsymbol{\mu}} = \mathbf{g}^{-1}(\mathbf{X}\hat{\boldsymbol{\beta}})$$

avec variance approximative: $\text{Var}(\hat{\boldsymbol{\mu}}) = \mathbf{B}\mathbf{X}(\mathbf{X}'\mathbf{B}\mathbf{A}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{B}$.

II.2 MODÈLES LINÉAIRES GÉNÉRALISÉS MIXTES

Dans un modèle linéaire mixte normal, on pose

$$\mathbf{y} = \boldsymbol{\mu} + \boldsymbol{\varepsilon} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon}$$

où $\boldsymbol{\mu} = E(\mathbf{y} | \mathbf{u}) = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}$ est la moyenne conditionnelle d' \mathbf{y} étant donné \mathbf{u} , \mathbf{Z} est une matrice $n \times q$ d'incidence des effets aléatoires \mathbf{u} . On suppose que $\mathbf{u} \sim N(\mathbf{0}, \mathbf{G})$ où $\mathbf{0}$ est le vecteur nul de dimension $q \times 1$, \mathbf{G} est la matrice de variance-covariance de \mathbf{u} . On suppose que \mathbf{u} et $\boldsymbol{\varepsilon}$ sont indépendants. Les autres symboles sont comme dans le modèle linéaire généralisé où la fonction de lien serait la fonction identité. La variance d' \mathbf{y} conditionnellement à \mathbf{u} est $\text{Var}(\mathbf{y} | \mathbf{u}) = \mathbf{R} = \text{Var}(\boldsymbol{\varepsilon})$, cette dernière pouvant prendre une forme bloc-diagonale si la variable aléatoire Y est mesurée à plusieurs reprises sur un même sujet (Brown et Prescott, 1999, Ch. 2).

Les modèles linéaires généralisés mixtes sont une généralisation à la fois des modèles linéaires généralisés et des modèles linéaires mixtes. Dans un tel modèle, on suppose comme dans ces derniers que

$$\mathbf{y} = \boldsymbol{\mu} + \boldsymbol{\varepsilon}$$

et que $g(\boldsymbol{\mu}) = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}$ où, comme dans un modèle linéaire généralisé, $g(\cdot)$ est une fonction de lien monotone, et comme dans un modèle linéaire mixte, $\mathbf{u} \sim N(\mathbf{0}, \mathbf{G})$. $\boldsymbol{\mu}$ est la moyenne conditionnelle, $E(\mathbf{y} | \mathbf{u})$. On a encore $\text{Var}(\boldsymbol{\varepsilon}) = \mathbf{R}$.

La vraisemblance d' \mathbf{y} est obtenue en multipliant la densité conditionnelle d' $\mathbf{y} | \mathbf{u}$ par la densité marginale de \mathbf{u} . Si \mathbf{R} est diagonale, on peut maximiser cette vraisemblance. Si \mathbf{R} est bloc-diagonale, on doit avoir recours à la notion de quasi-vraisemblance.

La quasi-vraisemblance d'un vecteur \mathbf{y} est définie comme suit.

Soit

$$\begin{aligned}\mathbf{y} &= (y_1, \dots, y_n)', \\ E(y_i) &= \mu_i \\ \text{var}(y_i) &\propto V(\mu_i) \text{ où } V(\cdot) \text{ est une fonction connue de } \mu_i.\end{aligned}$$

Supposons que $\mu_i = h(\eta_i)$ où $\eta_i = \mathbf{x}'_i \boldsymbol{\beta}$ et \mathbf{x}_i est une ligne de \mathbf{X} .

La fonction de quasi-vraisemblance $Q(\mu_i, y_i)$ est définie par:

$$\frac{\delta Q(\mu_i, y_i)}{\delta \mu_i} = \frac{y_i - \mu_i}{V(\mu_i)}$$

La fonction $Q(\mu_i, y_i) = \frac{y_i \theta_i - b(\theta_i)}{a(\theta_i)} + c(y_i, \theta_i)$ satisfait cette définition.

La quasi-vraisemblance conditionnelle d' $y_i | u_i$ est

$$Q(\mu_i; y_i | u_i) = \frac{y_i \theta_i - b(\theta_i)}{a(\theta_i)}$$

L'expression matricielle de la quasi-vraisemblance jointe d' \mathbf{y} devient

$$Q(\boldsymbol{\mu}, \mathbf{u}; \mathbf{y}) = \mathbf{y}' \mathbf{A}^{-1} \boldsymbol{\theta} - (\mathbf{b}_0^{1/2})' \mathbf{A}^{-1} \mathbf{b}_0^{1/2} + \frac{1}{2} \mathbf{u}' \mathbf{G} \mathbf{u}.$$

$\boldsymbol{\theta} = \mathbf{g}(\boldsymbol{\mu})$. Tous les autres symboles sont déjà définis.

Un estimateur de $\boldsymbol{\beta}$ et de \mathbf{u} est obtenu en résolvant le système d'équations

$$\begin{bmatrix} \mathbf{X}'\mathbf{W}\mathbf{X} & \mathbf{X}'\mathbf{W}\mathbf{Z} \\ \mathbf{Z}'\mathbf{W}\mathbf{X} & \mathbf{Z}'\mathbf{W}\mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{W}\mathbf{y}^* \\ \mathbf{Z}'\mathbf{W}\mathbf{y}^* \end{bmatrix}$$

où $\mathbf{W} = \mathbf{A}^{-1}\mathbf{B}$ et \mathbf{y}^* est défini plus haut.

L'inférence et l'estimation reposeront sur des fonctions linéaires de $\boldsymbol{\beta}$ et de \mathbf{u} : $\mathbf{K}'\boldsymbol{\beta} + \mathbf{M}'\mathbf{u}$ dictées par les questions de recherche.

Si l'on pose $\mathbf{L}' = [\mathbf{K}' \mathbf{M}']$,

$$\text{var}(\mathbf{K}'\hat{\boldsymbol{\beta}} + \mathbf{M}'\hat{\mathbf{u}}) = \mathbf{L}'\mathbf{C}\mathbf{L}$$

où

$$\mathbf{C} = \begin{bmatrix} \mathbf{X}'\mathbf{W}\mathbf{X} & \mathbf{X}'\mathbf{W}\mathbf{Z} \\ \mathbf{Z}'\mathbf{W}\mathbf{X} & \mathbf{Z}'\mathbf{W}\mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix}^{-1}$$

Une statistique F approximative est obtenue par :

$$F = \frac{[\mathbf{L}'\hat{\boldsymbol{\beta}}(\hat{\mathbf{u}} - \mathbf{u})]'(\mathbf{L}'\mathbf{C}\mathbf{L})^{-1}[\mathbf{L}'\hat{\boldsymbol{\beta}}(\hat{\mathbf{u}} - \mathbf{u})]}{\text{rang}(\mathbf{L})}$$

Sous $H_0 : \mathbf{L}' \begin{bmatrix} \boldsymbol{\beta} \\ \mathbf{u} \end{bmatrix} = 0$, F suit approximativement une loi de F_{v_1, v_2} où $v_1 = \text{rang}(\mathbf{L})$ et v_2 est égal

au nombre de degrés de liberté utilisés pour estimer $\mathbf{L}'\mathbf{C}\mathbf{L}$ habituellement obtenus à l'aide d'une approximation comme celle de Satterthwaite, par exemple. L'approximation vient à la fois du calcul des degrés de liberté associés à l'estimation de la variance de la fonction linéaire et du postulat d'indépendance des variables aléatoires au numérateur et au dénominateur de la loi statistique.

Les distributions pour les termes d'erreur que nous utiliserons dans notre cas sont les lois binomiale et de Poisson. Les fonctions de liens $g(\cdot)$ sont alors respectivement les liens *logit* et logarithmique, donnant ainsi des modèles logistique et log-linéaire respectivement.

II.3 PARAMÈTRES DE DISPERSION ET FONCTION DE DÉVIANCE

Souvent il arrive que la variance des erreurs ne coïncide pas avec celle que prévoit la théorie. Elle peut être plus grande (on parle de sur-dispersion) ou plus petite (on parle de sous-dispersion). Par exemple, pour un modèle de Poisson, la théorie nous dit que la variance est égale à la moyenne μ .

Le paramètre de dispersion (noté ϕ) permet donc d'augmenter ou de diminuer la variance prévue par la densité de probabilité. On suppose que $\phi > 0$. La situation idéale serait d'avoir $\phi = 1$ ce qui signifie qu'il n'y a pas de sous ou de sur-dispersion. Si ϕ est significativement plus petit que 1, il y a sous-dispersion causant une sur-estimation des erreurs standards des paramètres et une sous-estimation des statistiques des tests. Si ϕ est significativement plus grand que 1, il y a sur-dispersion avec comme conséquence un gonflement irréaliste des statistiques des tests et une sous-estimation des erreurs standards des paramètres. Dans l'un ou l'autre cas, il faudra faire des ajustements.

Pour prendre en compte la dispersion, il suffit de multiplier la matrice de variance-covariance des résidus par ϕ . Si la variance observée des résidus est exactement égale à celle prédite, alors le paramètre de dispersion vaut 1. Toutefois, ce n'est souvent pas le cas.

Pour les distributions binomiale et de Poisson on a le résumé de quelques uns de leurs composantes au tableau 1 (voir les tableaux en annexe).

Une fonction liée à la fonction de log-vraisemblance qui joue un rôle important notamment est la fonction de déviance. La statistique de déviance sert plutôt à estimer la valeur du paramètre ϕ et à mesurer le degré d'ajustement du modèle aux données. On peut également l'utiliser pour les tests d'hypothèse. Elle est égale au double de la différence des log-vraisemblances maximisées du modèle saturé et du modèle qui nous intéresse. Elle peut être standardisée à l'aide du paramètre de dispersion ϕ .

La fonction de déviance non standardisée est définie comme suit :

$$d(\hat{\boldsymbol{\mu}}, \mathbf{y}) = 2[l(\mathbf{y}; \mathbf{y}) - l(\hat{\boldsymbol{\mu}}; \mathbf{y})]$$

$l(\mathbf{y}; \mathbf{y})$ est la fonction log-vraisemblance du modèle saturé. Dans ce cas, $\hat{\boldsymbol{\mu}}$ est estimé par \mathbf{y} .

$l(\hat{\boldsymbol{\mu}}; \mathbf{y})$ est la fonction log-vraisemblance du modèle d'intérêt.

La fonction de déviance standardisée s'obtient en divisant $d(\hat{\boldsymbol{\mu}}; \mathbf{y})$ par ϕ , autrement dit :

$$d^*(\hat{\boldsymbol{\mu}}; \mathbf{y}) = \frac{1}{\phi} d(\hat{\boldsymbol{\mu}}; \mathbf{y}) = 2[l(\mathbf{y}, \phi; \mathbf{y}) - l(\hat{\boldsymbol{\mu}}, \phi; \mathbf{y})]$$

L'adjectif standardisé vient du fait que l'on retrouve ϕ au dénominateur des fonctions log-vraisemblances.

III. EXEMPLE D'APPLICATION

Cette section de l'essai décrit une application des modèles linéaires généralisés mixtes à des données sylvicoles. Plus précisément, les GLMMs seront utilisés pour déterminer dans quelle mesure l'abondance des carabidés dans une sapinière à bouleau jaune est modifiée par les éclaircies et la scarification des sols. Les aspects du problème qui seront abordés dans l'analyse sont les suivants :

- 1 Effet du type de coupe sur l'abondance moyenne au cours des mois de juin, juillet, et août :
 - 1.1 témoin *vs* coupes (non scarifié seulement)
 - 1.2 coupe à blanc *vs* éclaircies
 - 1.3 2 trouées *vs* (4 ou 8)
 - 1.4 4 *vs* 8 trouées

- 2 Effet principal de la scarification sur l'abondance moyenne au cours de la saison

- 3 Effet moyen de l'interaction entre le type de coupe et la scarification au cours de la saison :
 - 3.1 Scarification \times (coupe à blanc *vs* éclaircies)
 - 3.2 Scarification \times (2 trouées *vs* (4 ou 8))
 - 3.3 Scarification \times (4 *vs* 8 trouées)

- 4 Effet principal du temps:
 - 4.1 Effet linéaire du temps
 - 4.1 Effet quadratique du temps

- 5 Effet du type de coupe sur la courbe d'évolution de la variable d'étude au cours de la saison :
 - 5.1 (témoin *vs* coupes) \times temps si non scarifié
 - 5.2 (coupe à blanc *vs* éclaircies) \times temps
 - 5.3 (2 trouées *vs* (4 ou 8)) \times temps
 - 5.4 (4 *vs* 8 trouées) \times temps

- 6 Effet moyen de l'interaction entre le type de scarification et le temps

- 6.1 Scarification × temps linéaire
 - 6.2 Scarification × temps quadratique
- 7 Effet de l'interaction entre le type de coupe et la scarification sur la courbe d'évolution de la variable d'étude au cours de la saison :
- 7.1 (coupe à blanc vs éclaircies) × scarification × temps linéaire
 - 7.2 (coupe à blanc vs éclaircies) × scarification × temps quadratique
 - 7.3 (deux trouées vs (4 ou 8)) × scarification × temps linéaire
 - 7.4 (deux trouées vs (4 ou 8)) × scarification × temps quadratique
 - 7.5 (4 vs 8 trouées) × scarification × temps linéaire
 - 7.6 (4 vs 8 trouées) × scarification × temps quadratique

Avant d'évaluer ces différents effets, il importe de préciser les conditions expérimentales dans lesquelles les données ont été collectées. Dans un premier temps, nous décrirons le dispositif expérimental. Les variables d'études seront ensuite présentées. Trois analyses seront alors proposées dans la section « méthodes. » Les deux premières analyses porteront sur l'abondance de la biodiversité des insectes et de l'espèce la plus fréquente. La troisième analyse visera à modéliser la probabilité de présence/absence de cette dernière. Dans une troisième section, nous présentons et discutons les résultats des analyses.

III.1 EXPÉRIENCE

Comme dans toute expérience, les résultats et les conclusions obtenus dépendent largement de la manière dont les données ont été collectées ; il convient donc de décrire dans cette section le plan d'expérience ainsi que la nature des données recueillies.

III.1.1 *Plan d'expérience et collecte des données*

Dans ce projet, l'étude de la diversité de l'entomofaune (c'est-à-dire la faune constituée par les insectes) vise à déterminer les effets de la coupe forestière et de la scarification des sols sur la biodiversité de cinq ordres d'arthropodes à l'aide de pièges-fosses. Les traitements sylvicoles consistent en une coupe à blanc (CB), et trois formes d'une même intensité d'éclaircie (35%), soit deux, quatre ou huit trouées que l'on a dénotées respectivement par T2, T4 et T8. Deux modes de scarification (scarifié (Y) ou non (N)) ont été croisés avec les quatre types d'éclaircie. Il y a en outre un traitement témoin sans éclaircie ni scarification. Les arthropodes capturés dans 120 pièges-fosses ont été récoltés, identifiés puis comptés toutes les semaines entre le 5 juin et le 29 août de l'an 2000, soit pendant 13 semaines. Dans la suite, on utilisera souvent indifféremment les termes parcelle ou traitement.

Plus concrètement, l'expérience s'est déroulée sur deux sites : le lac des Étangs (dans Portneuf) et le lac Poissonneux (dans la ZEC Batiscan-Neilson). Deux blocs de cinq parcelles à peu près carrées de 1.3 ha chacune ont été formés dans chaque site. Les blocs n'ont pas été choisis au hasard car il y avait des contraintes liées entre autres au relief, ainsi qu'aux cours d'eau. Les cinq types de coupe considérés ont été les suivants :

- témoin (forêt non perturbée),
- coupe à blanc,
- deux trouées,
- quatre trouées,
- huit trouées et
- témoin

Ces cinq types de coupe ont ensuite été randomisés parmi les grandes parcelles de chaque bloc sur les deux sites. La coupe à blanc consiste à couper 70% de la superficie de la parcelle en prenant soin de laisser une bande d'arbres non abattus tout au tour. Les coupes partielles (deux, quatre ou huit trouées) consistent à couper des zones circulaires dont la surface constitue 35% de la surface totale de la parcelle. Évidemment, les trouées sont plus ou moins grandes, dépendamment de leur nombre, mais la surface coupée reste la même. Enfin la parcelle témoin ne subit aucune coupe.

Une fois les traitements de coupe assignés, on a procédé de même avec la scarification des sols, laquelle consiste à retourner la terre. On considère chaque grande parcelle divisée en deux sous-parcelles. L'une est scarifiée et l'autre non, sauf pour la parcelle témoin, ce qui cause un déséquilibre dans le plan expérimental. En réalité, pour les parcelles avec quatre et huit trouées, deux et quatre respectivement des quatre et huit trouées ont été choisies au hasard et scarifiées. Dans la parcelle avec deux trouées, une moitié choisie au hasard de chaque trouée a été scarifiée. La parcelle coupée à blanc a été divisée en quatre carrés et une paire de carrés de la diagonale choisie au hasard a été scarifiée. La parcelle témoin n'a pas été scarifiée.

Des pièges-fosses circulaires contenant 100 ml d'éthanol (75%) ont été ensuite installés dans les parcelles et les sous-parcelles. On a disposé six pièges-fosses dans chaque grande parcelle à raison de trois pièges-fosses dans la partie scarifiée et trois dans la partie non scarifiée, hormis le témoin où les six pièges sont dans la partie non scarifiée. Il y avait ainsi 30 pièges-fosses par bloc, 60 par site, soit un total de 120 pièges-fosses. Une distance minimale de 15 m a été laissée entre deux pièges. On suppose que les pièges sont ainsi suffisamment éloignés les uns des autres pour être indépendants.

Chaque semaine, les insectes capturés dans chaque piège ont été recueillis. Chaque journée complète a été consacrée à l'un des sites seulement et le lendemain à l'autre, dans un ordre précis, pour compter les captures des 60 pièges de chaque site. Donc deux jours ont été nécessaires pour parcourir les 120 pièges. Il est arrivé que quelques pièges soient inondés et que

leur contenu ne puisse être prélevé. Une fois au laboratoire, l'identification de chacun des ordres des insectes a été effectuée, ainsi que le décompte.

Le plan expérimental est un plan à unités divisées («*split-plot*») à quatre blocs aléatoires complets (deux dans chaque site). Les grandes parcelles correspondent aux traitements sylvicoles et les sous-parcelles à la scarification. La structure des traitements est factorielle 4×2 , plus le témoin. Le plan inclut également trois pièges par sous-parcelle, tous situés dans les trouées sauf ceux de la parcelle témoin dans laquelle il y a six pièges (pas de sous-parcelles). Originellement, un décompte répété a été effectué de façon hebdomadaire pour chaque piège.

III.1.2 Présentation des données

Le fichier de données brutes contient une ligne par piège par semaine par espèce de carabidé. Les variables sont les blocs (A, B, C ou D), les traitements sylvicoles (T8, T4, T2, CB ou TE), la présence ou non de scarification (S ou NS), le numéro de piège (1, ..., 120), la semaine (1, ..., 13), l'identification de l'espèce et le nombre de spécimens de l'espèce. La variable d'analyse est soit le nombre d'espèces ou de spécimens. Nous n'avons pas tenu compte des sites dans l'analyse.

Au cours de la saison, soit pendant 13 semaines, les 120 pièges ont capturé en tout 1285 insectes de 40 espèces de la famille des carabidés. L'espèce la plus abondante a été *Pterostichus coracinus* (Newman), dont 407 spécimens ont été piégés. Viennent ensuite, dans l'ordre, les espèces *Pterostichus adstrictus* (Eschscholtz) et *Synuchus impunctacus* (Say) dont les fréquences respectives sont de 286 et 110 spécimens. Ces fréquences sont réparties dans les 120 pièges et les 13 semaines.

Un souci majeur avec les données est la prévalence d'un nombre élevé de zéros dans les dénombrements d'insectes. Sur les 1560 observations, résultats de la combinaison des 13 semaines par 120 pièges, 741 observations de la variable nombre de spécimens de l'espèce *P. coracinus* se sont avérées nulles, soit un pourcentage de 47.5%. Ce pourcentage très élevé de fréquences nulles étant susceptible de poser un problème de convergence dans l'estimation des

paramètres du modèle, plusieurs tentatives de sommation des fréquences ont été entreprises en veillant à préserver suffisamment la structure expérimentale pour atteindre les objectifs d'origine.

III.1.3 Manipulation des fichiers de données

Le fichier de départ a été manipulé et réorganisé à coups de regroupements et d'agrégations afin d'apprêter les données pour les analyses envisagées, à savoir les évolutions hebdomadaire et mensuelle de la biodiversité des espèces capturées par piège et par sous-parcelle.

L'analyse de l'abondance des espèces utilise comme variable réponse le nombre d'espèces différentes par semaine et par sous-parcelle. Elle nécessite une agrégation des données par voie de sommation des fréquences des espèces sur les pièges situés dans une sous-parcelle scarifiée et non scarifiée respectivement.

Nous avons également dû créer une variable *offset* pour tenir compte du déséquilibre causé par la parcelle témoin qui est non scarifiée et dans laquelle sont placés six pièges au lieu de trois. C'est une variable de pondération. Le résultat de cette opération est de ramener le pourcentage de zéros à près de 28%. En effet, sur les 468 observations, 131 fréquences nulles sont comptabilisées.

Nous explorerons également l'abondance par piège et par groupe de quatre semaines. Pour cela, la treizième semaine n'a pas été prise en compte. Les quatre premières semaines constituent la période 1, les quatre semaines suivantes, la période 2, et les quatre dernières semaines, la période 3.

Par ailleurs, un regroupement par piège et par mois a également été tenté afin de tenir compte de l'effet de piège, sans variable *offset*.

Ces différentes opérations sont délicates et requièrent de veiller à ne pas compter une espèce plus d'une fois dans une sous-parcelle.

Quant à l'étude de l'abondance de l'espèce la plus fréquente (*P. coracinus* (Newman)), la variable modélisée est le nombre de spécimens. Comme pour l'étude de la biodiversité des espèces, il y a beaucoup de zéros et même davantage dans les dénombrements par piège et par semaine. Des agrégations sont donc nécessaires. Une première agrégation pour tenir compte de l'effet aléatoire des pièges a été effectuée, ce qui a ramené le pourcentage de zéros à 39%. Nous avons alors décidé d'aller plus loin dans l'agrégation des données en payant le prix par la perte de finesse dans l'analyse. Ainsi, nous avons fait la somme des fréquences, pour chaque grande parcelle, sur les pièges des sous-parcelles scarifiées d'une part, et sur les pièges des sous-parcelles non scarifiées d'autre part, et ce pour les trois périodes. Une variable *offset* a alors été créée. Ceci a eu pour effet de réduire le pourcentage de zéros à 13%.

Pour modéliser la probabilité de présence-absence de l'espèce *P. coracinus* (Newman), une variable binomiale a été créée. Cette variable prend la valeur 1 si l'espèce est présente dans un piège à une période donnée (autrement dit, si la variable nombre de spécimens est différente de 0); sinon, elle prend la valeur 0 (c'est-à-dire si le nombre de spécimens est nul).

À la première période, chacun des pièges de chaque bloc situé, dans les parcelles témoins d'une part, et dans les sous-parcelles des parcelles avec huit trouées d'autre part, a capturé au moins un spécimen de l'espèce *P. coracinus* (Newman). Cela pose un problème dû au fait que certaines probabilités observées de présence de l'espèce *P. coracinus* (Newman) étaient égales à un quelle que soit la répétition. La probabilité observée de présence de cette espèce est donc de 1, dans ces combinaisons des trois facteurs. Pour éviter le problème de convergence qui s'ensuit, on a donc choisi d'enlever les données correspondant à ces traitements, mais seulement pour la première période. Ceci constitue certes une perte d'information et une source de déséquilibre mais il y va de la convergence de l'algorithme d'estimation des paramètres du modèle.

Dans tous les cas, les données comprennent les variables suivantes : les blocs (A, B, C et D), les traitements (T8, T4, T2, CB et TE), la présence ou non de scarification (N ou Y). L'analyse hebdomadaire de l'abondance des espèces par piège nécessite d'utiliser, en plus, les variables piège (1, ..., 120) et semaine (1, ..., 13). Le regroupement par sous-parcelle et par période de quatre semaines exige de créer respectivement une variable *offset* pour la modélisation de la

biodiversité des espèces. De plus, la variable réponse pour l'analyse de la probabilité de présence ou d'absence de cette espèce est binomiale.

Afin d'explorer diverses applications des modèles linéaires généralisés mixtes, nous étudierons les effets ci-dessus sur trois variables :

- le nombre d'espèces différentes dénombrées par sous-parcelle et par période de quatre semaines ;
- l'abondance de l'espèce *P. coracinus* (Newman) modélisée à l'aide d'une loi de Poisson. Deux modèles seront étudiés, la première prend en compte l'effet aléatoire de piège et la seconde une variable *offset*.
- la présence/absence de *P. coracinus*, une variable binomiale.

III.2 MÉTHODES

L'analyse des trois variables précitées a été effectuée en six grandes étapes. Un modèle complet valable pour les trois variables a été spécifié avant de choisir la structure de variance pour les mesures répétées. Ensuite, nous avons réduit la partie aléatoire du modèle, lorsque nécessaire. L'analyse de la partie fixe, qui est très importante, fait l'objet de la quatrième étape. Par la suite, une décomposition de la partie fixe a été entreprise à l'aide de contrastes. Enfin, les moyennes pertinentes à l'interprétation des contrastes ont été estimées.

III.2.1 La macro GLIMMIX de SAS

Il n'y a pas encore de procédure pour ajuster tous les types de modèles linéaires généralisés mixtes. La procédure GENMOD ne permet d'ajuster que certains types de modèles linéaires généralisés mixtes avec l'énoncé REPEATED. La méthode Generalized Estimating Equations (G.E.E.) est ainsi utilisée pour ce faire. Par ailleurs, la procédure MIXED est indiquée pour l'ajustement de modèles linéaires mixtes mais ne permet pas de spécifier de distribution autre

que la loi normale pour les effets résiduels. Cette procédure offre beaucoup de possibilités avec ses nombreux énoncés et options.

Seule la macro GLIMMIX de SAS (Wolfinger & O'Connell, 1993) permet d'ajuster le type de modèle que nous avons avec des effets aléatoires à plusieurs niveaux, des mesures répétées et une loi de Poisson ou binomiale. Cette macro est appropriée pour tous les types de modèles linéaires généralisés. Elle ajuste des modèles linéaires généralisés mixtes en maximisant une pseudo-vraisemblance. À chaque itération, la macro calcule une pseudo variable linéaire et ajuste un modèle mixte pondéré avec PROC MIXED. Un paramètre de dispersion est calculé par défaut mais on peut choisir de le garder fixe avec l'option PARMs de PROC MIXED.

La plupart des options de PROC GENMOD et de PROC MIXED peuvent être utilisées. On peut inclure à la fois des effets aléatoires et des mesures répétées et choisir la loi que suit la variable réponse (loi binomiale ou de Poisson dans ce cas) ainsi que la fonction de lien.

Même si la macro GLIMMIX fonctionne bien dans l'ensemble, elle n'a pas encore été totalement testée par ses concepteurs. Sa fiabilité n'est donc pas aussi grande que celle des procédures SAS éprouvées. Notamment, on ne connaît pas la valeur de l'approximation linéaire utilisée à chaque itération. On l'utilisera tout de même pour ajuster notre modèle car c'est le seul outil dont nous disposons.

Dans un premier temps, le critère d'Akaike (AIC) a été utilisé pour choisir la structure de covariance des mesures répétées la plus appropriée. Le critère de Schwartz (BIC) nous a ensuite permis de confirmer ce choix. Les deux critères, rappelés ci-dessous, sont sensibles au nombre de paramètres du modèle :

$$AIC = -2l + 2d$$

$$BIC = -2l + d \log(m)$$

Ici, l est la valeur maximale du logarithme de la vraisemblance, d est le nombre de paramètres du modèle et m est le nombre d'observations. Plus petite est la valeur du critère, meilleur est le

modèle. L'emploi de ces critères facilite considérablement le choix de la structure de covariance, qui n'est généralement pas aisé.

III.2.2 Modèle complet

Le modèle (1) ci-dessous est un modèle complet valable pour les analyses 1, 2.a), 2.b) et 3 à quelques différences près.

1. Analyse de la biodiversité des espèces: c'est un modèle de Poisson avec variable *offset*
2. Analyse de l'abondance de l'espèce *P. coracinus*:
 - 2.a) à l'aide d'un modèle de Poisson avec effet aléatoire de piège
 - 2.b) à l'aide d'un modèle de Poisson avec variable *offset*
3. Analyse de la probabilité de présence/absence de l'espèce *P. coracinus*

$$\eta_{ijklt} = \mu + a_i + \alpha_j + b_{ij} + \beta_k + (\alpha\beta)_{jk} + c_{ijk} + \delta_t + (\alpha\delta)_{jt} + (\beta\delta)_{kt} + (\alpha\beta\delta)_{jkt} + f_{ijkt} + d_{ijklt} \quad (1)$$

Pour tous les modèles, on a :

η_{ijklt} : moyenne de $g(\mu_{ijklt})$ à la période t ($t=1, 2, 3$), dans le piège l , $l=1, \dots, n_{ijkt}$, n_{ijkt} est le nombre de pièges où on a prélevé des insectes au temps t dans la sous-parcelle k ($k=NS, S$), du traitement j ($j=CB, T2, T4, T8, TE$), du bloc i ($i=A, B, C, D$).

μ : est un paramètre de référence

a_i : est l'effet aléatoire du bloc i

α_j : est l'effet fixe du traitement j

b_{ij} : est l'effet aléatoire de la grande parcelle (i, j) (combinaison du bloc i avec le traitement j)

β_k : est l'effet fixe de la scarification k

$(\alpha\beta)_{jk}$: est l'effet fixe de l'interaction entre le niveau j de traitement et le niveau k de scarification

c_{ijk} : est l'effet aléatoire de sous-parcelles (combinaison du niveau i de bloc, du niveau j de traitement et du niveau k de scarification)

δ_t : est l'effet fixe de la période t

$(\alpha\delta)_{jt}$: est l'effet fixe de l'interaction entre le niveau j de traitement et du niveau t de période

$(\beta\delta)_{kt}$: est l'interaction fixe entre le niveau k de scarification et du niveau t de période

$(\alpha\beta\delta)_{jkt}$: est l'effet fixe de l'interaction entre le niveau j de traitement, du niveau k de scarification et du niveau t de période

f_{ijkt} : est l'effet aléatoire de la sous-parcelle (i, j, k) à la période t .

Dans les modèles 1 et 2.b), $d_{ijklt} = \log(n_{ijkt})$ et ne dépend pas de l . C'est la variable *offset*. Elle a deux valeurs possibles en général : 3 pour les sous-parcelles avec deux, quatre, huit trouées ou coupées à blanc et 6 pour la parcelle témoin.

Dans les modèles 2.a) et 3, d_{ijklt} correspond à l'effet aléatoire du piège l dans la sous-parcelle (i, j, k), $l = 1, \dots, n_{ijkt}$ et ne dépend pas de t , en général.

Pour les modèles 1 et 2.b),

$$\eta_{ijklt} = E\left[\log(y_{ijklt})\right]$$

est la moyenne conditionnelle (étant donné les effets aléatoires) du logarithme du nombre d'espèces pour la combinaison (ou sous-parcelle) (i, j, k), du bloc i ($i = A, B, C, D$), du traitement j ($j = CB, T2, T4, T8, TE$), de la scarification k ($k = NS, S$) où sont placés n_{ijkt} pièges ($n_{ijkt} = 3$ si $j = CB, T2, T4$ ou $T8$ et $n_{ijkt} = 6$ si $j = TE$) à la période t ($t = 1, 2, 3$). L'indice l n'est pas nécessaire pour ces deux modèles.

Pour le modèle 2.a), on a la même signification de η_{ijklt} , sauf que l représente un numéro de piège.

Pour le modèle 3,

$$\eta_{ijklt} = E \left[\log \left(\frac{p_{ijklt}}{1 - p_{ijklt}} \right) \right]$$

est l'espérance du *logit* de p_{ijklt} , la probabilité de présence de l'espèce *P. coracinus* dans l'un des pièges de la sous-parcelle (i, j, k) à la période t . L'indice l n'est pas nécessaire. Cette probabilité est obtenue à partir de la variable binomiale y_{ijklt} définie comme suit :

$y_{ijklt} = 1$ si au moins un spécimen de l'espèce *P. coracinus* est présente dans la sous-parcelle (i, j, k) , à la période t

$y_{ijklt} = 0$ si aucun spécimen de l'espèce n'est présent.

Le modèle s'appuie sur les postulats suivants :

$$a_i \sim N(0, \sigma_a^2)$$

$$b_{ij} \sim N(0, \sigma_b^2)$$

$$c_{ijk} \sim N(0, \sigma_c^2)$$

$$f_{ijkl} \sim N(0, \sigma_f^2)$$

$$d_{ijklt} \sim N(0, \sigma_d^2), \text{ seulement pour le modèle 2.a) avec effet aléatoire de piège.}$$

Ces effets aléatoires sont également supposés indépendants.

III.2.3 Calcul des nombres de degré de liberté (d.d.l.)

Ce calcul dont les résultats sont résumés au Tableau 1, en annexe, est effectué pour le modèle 1. Le principe est le même pour les autres modèles.

Les quatre blocs (Bloc), les cinq types de traitements sylvicoles (Trait), les deux niveaux de scarification (Scari) ainsi que les trois périodes (Periode) donnent respectivement 3, 4, 1 et 2 degrés de liberté. Les combinaisons des niveaux de blocs et des traitements sylvicoles correspondent à $3 \times 4 = 12$ degrés de liberté. À cause du déséquilibre causé par la parcelle témoin (TE), seuls quatre niveaux de Trait sont combinés avec les deux niveaux de Scari, donnant ainsi 3 degrés de liberté pour une interaction Trait \times Scari. Les erreurs de sous-parcelles ont 12 degrés de liberté (Bloc \times Scari correspond à 3×1 d.d.l. ajoutés aux $3 \times 3 \times 1 = 9$ d.d.l. provenant des combinaisons Bloc \times Trait \times Scari). Les combinaisons de Trait \times Periode d'une part, et de Scari \times Periode d'autre part, donnent respectivement $(5 - 1) \times (3 - 1) = 8$ d.d.l. et $(2 - 1) \times (3 - 1) = 2$ d.d.l. respectivement. L'interaction Trait \times Scari \times Periode a $(4 - 1) \times (2 - 1) \times (3 - 1) = 6$ d.d.l. à cause du déséquilibre dû au traitement témoin. Les 54 d.d.l. correspondent aux $3 \times 1 \times 2 = 6$ d.d.l. de Bloc \times Scari \times Periode, aux $3 \times 3 \times 1 \times 2 = 18$ d.d.l. de Bloc \times Trait \times Scari \times Periode, aux $3 \times 4 \times 2 = 24$ d.d.l. de Bloc \times Trait \times Periode et enfin aux $3 \times 2 = 6$ d.d.l. de Bloc \times Periode.

On peut remarquer que le nombre de degrés de liberté pour la source de variation « Bloc » est assez petit pour que celle-ci soit considérée comme étant aléatoire. Les blocs peuvent être considérés comme fixes ou aléatoires (Littell et coll., 1996). Toutefois, en référence au livre Littell et coll. (1996, p. 453), nous avons décidé de considérer les effets des blocs comme aléatoires.

III.2.4 Choix de la structure de variance pour les mesures répétées

L'ajustement d'un modèle avec mesures répétées implique le choix d'une structure de covariance appropriée pour les mesures répétées. Ceci est important non seulement pour obtenir de bonnes estimations des erreurs standard des effets fixes mais également pour interpréter les composantes

de variance. Selon la structure de covariance choisie, les corrélations entre les dénombrements d'insectes d'une même sous-parcelle à des périodes différentes se comportent différemment. Plusieurs structures de covariance ont été testées et les résultats consignés dans le tableau 3 en annexe.

Ces structures de covariance ont toutes conduit à la convergence de la macro GLIMMIX de SAS pour le modèle d'analyse de la biodiversité des espèces. L'ajustement des données avec la structure de covariance *VC* semble être approprié au regard des critères d'Akaike et de Schwartz. En effet, la valeur du critère d'Akaike correspondante est la plus petite et le critère de Schwartz le confirme. Ceci a conforté notre décision d'adopter ce modèle par la suite.

Lorsque les structures de covariance UN et CS sont utilisées, la macro GLIMMIX ne converge pas pour le modèle d'analyse avec effet aléatoire de piège pour l'étude de l'abondance de l'espèce *P. coracinus*. En revanche, les structures VC, AR(1) et TOEP conduisent à la convergence de l'algorithme. Parmi les modèles pour lesquels il y a convergence, aussi bien le critère d'Akaike que celui de Schwartz conduisent au choix de celui avec la structure AR(1).

Les types de structures de covariance testés ont tous permis la convergence de la macro pour le modèle avec variable *offset* pour l'analyse de l'abondance de l'espèce *P. coracinus*. La structure *VC* semble néanmoins suffire pour modéliser les données, vu que l'AIC correspondant est le plus petit. C'est la structure la plus simple. Puisque les corrélations sont nulles, il ne semble pas y avoir d'association entre les observations d'une même sous-parcelle d'une période à l'autre.

La macro GLIMMIX a convergé seulement avec les structures VC et AR(1) pour le modèle d'analyse de la probabilité de présence/absence de l'espèce *P. coracinus*. Le critère d'Akaike suggère de choisir la première structure.

III.2.5 Réduction de la partie aléatoire

Les modèles d'analyse peuvent être allégés de certains effets aléatoires si cela s'avère nécessaire. Le processus de réduction se fait en partant du modèle complet pour lequel les paramètres de

covariance sont estimés. L'effet aléatoire le plus imbriqué est d'abord enlevé du modèle. On recalcule la valeur du critère d'Akaike et les nouvelles valeurs des paramètres de covariance. On enlève ensuite le deuxième effet aléatoire le plus imbriqué, on recalcule l'AIC et les estimations des paramètres. On continue ainsi le processus jusqu'à ce que l'effet aléatoire le moins imbriqué soit enlevé du modèle. On observe alors les valeurs d'AIC et les estimations des paramètres de covariance pour choisir un modèle final. Les différentes étapes de réduction de la partie aléatoire pour chacun des quatre modèles sont rapportées dans les tableaux 4, 5, 6 et 7 en annexe.

Le tableau 4 se rapporte au modèle d'analyse de la biodiversité des espèces et indique que σ_a^2 et σ_b^2 sont estimés presque nuls. On devrait donc enlever les effets aléatoires de Bloc et de Bloc \times Trait du modèle. Toutefois, les critères d'Akaike et de Schwartz valant respectivement 88.6 et 87.3 suggèrent le modèle obtenu à l'étape 0 pour lequel aucun effet aléatoire n'est enlevé, c'est-à-dire le modèle complet.

Pour l'analyse de l'abondance de l'espèce *P. coracinus* avec effet aléatoire de piège, seuls les effets aléatoires f_{ijkt} et d_{ijkl} restent dans le modèle réduit (tableau 5).

Au tableau 6 sont rapportées les différentes étapes de la réduction de la partie aléatoire du modèle d'analyse de l'abondance de l'espèce *P. coracinus* avec variable *offset*. Les modèles obtenus aux étapes 0 et 3 donnent strictement les mêmes estimations des paramètres de covariance, ce qui est normal du reste. En effet, les composantes de variance enlevées étaient estimées à 0. Il ressort aussi de ce tableau que le facteur aléatoire « Bloc » n'a pas d'effet sur le logarithme moyen du nombre de spécimens de *P. coracinus*. Il est estimé à zéro. Si l'on se fie aux critères d'Akaike et de Schwartz, le modèle où seul l'effet aléatoire de Bloc \times Trait est présent semble convenir.

Le modèle d'analyse de la probabilité de présence/absence contient tous les effets aléatoires de départ sauf d_{ijkl} , compte tenu de la petitesse des valeurs d'AIC et de BIC comparativement aux autres modèles.

III.2.6 Statistiques d'ajustement

Des statistiques telles la déviance et l'estimation du paramètre de dispersion sont fournies avec la macro GLIMMIX de SAS. Elles permettent de juger de la qualité d'ajustement des modèles et sont produites dans le tableau 8 en annexe.

Pour le modèle d'analyse de la biodiversité des espèces, il n'y a ni sur- ni sous-dispersion puisque le paramètre de même nom ϕ est estimé à 1.00. La déviance est estimée à 26.51, ce qui donne ainsi une déviance standardisée de 26.51 obtenue en divisant la déviance par la dispersion ($26.51/1.00 = 26.51$).

Lorsque l'on prend en compte l'effet aléatoire de piège pour analyser l'abondance de l'espèce *P. coracinus*, la déviance est estimée à 278.21 et la déviance standardisée à près de 297. L'estimation 0.94 du paramètre de dispersion étant très proche de 1, on conclut que la dispersion est bien celle d'une loi de Poisson, étant donnés les effets aléatoires.

Par ailleurs, il n'y a ni sous- ni sur-dispersion lorsque l'analyse de l'abondance de l'espèce *P. coracinus* se fait avec un modèle de Poisson avec variable *offset*. En effet, le paramètre de dispersion est estimé à 1. La déviance vaut 164.762.

Il n'y a pas non plus de sous- ou de sur-dispersion pour le modèle d'analyse de la probabilité de présence/absence de l'espèce *P. coracinus*.

En résumé, il n'y a donc pas de souci majeur quant à l'ajustement. Ces modèles serviront à l'analyse de la partie fixe.

III.2.7 Analyse de la partie fixe

Tous les effets fixes sont maintenus dans les différents modèles, car ce sont eux qui permettent d'atteindre les objectifs de recherche. Les résultats des tests de Fisher de type 3 correspondants aux modèles sont donnés au tableau 9.

Biodiversité des espèces : Les traitements sylvicoles ont un effet significatif sur le nombre d'espèces par piège et par période. En effet, la statistique du test de Fisher correspondant est égale à $F = 3.74$; elle comporte 4 et 12 degrés de liberté et le seuil observé correspondant est égale à 0.0336. Le test est donc significatif au seuil nominal de 5%. Il en est de même pour la scarification ($F = 7.01$ avec 1 et 12 d.d.l. et $P = 0.0213$) et la période ($F = 48.94$ avec 2 et 53 d.d.l. et $P < 0.0001$). L'interaction entre la scarification et la période a également un effet significatif sur l'abondance moyenne des espèces ($F = 3.93$, $P = 0.0256$). Aucun autre effet ne semble influencer la biodiversité ($P > 0.05$).

Abondance de l'espèce *P. coracinus* (effet aléatoire de pièges) : Les résultats suggèrent que les traitements sylvicoles ont un effet significatif sur l'abondance de *P. coracinus* ($F = 12.07$, $P < 0.0001$). La période aussi a un effet significatif sur l'abondance de cette espèce ($F = 3.59$, $P = 0.0344$) mais celui-ci ne serait pas le même selon que la zone éclaircie a été scarifiée ou non ($F = 4.19$ pour l'interaction Scari \times Période, $P = 0.0204$).

Abondance de l'espèce *P. coracinus* (variable *offset*) : Il semble y avoir un effet des traitements sylvicoles sur l'abondance de *P. coracinus* ($F = 12.90$, $P < 0.0001$). Ça semble être également le cas pour l'effet de la période selon le niveau de scarification Scari \times Période ($F = 3.21$, $P = 0.0469$).

Probabilité de présence/absence de *P. coracinus* : Seul l'effet principal des traitements sylvicoles est significatif ($F = 6.48$, $P = 0.0051$).

Tous les modèles suggèrent que les traitements sylvicoles ont une grande importance pour expliquer, d'une part, la biodiversité des espèces de carabidés en général et, d'autre part, la présence et l'abondance de l'espèce *P. coracinus* en particulier.

Les résultats de ces tests conduisent à reconsidérer les questions posées dans les objectifs de recherche. Seuls les contrastes associés aux effets significatifs présentent de l'intérêt.

Dans la sous-section suivante, nous présentons les résultats des estimations de certains effets pertinents et de ceux des contrastes qui ont de l'intérêt et ce, pour tous les modèles.

III.2.8 Décomposition de la partie fixe à l'aide de contrastes

Comme on peut le constater au tableau 10, seuls quelques contrastes présentent finalement un intérêt. Si un effet fixe n'est pas statistiquement significatif, il n'y a pas d'intérêt à faire des contrastes sur ses niveaux, qu'il s'agisse d'un effet principal ou d'une interaction. Par ailleurs, lorsqu'une interaction est statistiquement significative, on gagne plus à l'étudier plutôt que l'effet principal lui-même. Rappelons que les déséquilibres liés à la structure des données sont pris en compte. Il y a deux façons de faire les contrastes pour le modèle d'analyse de la probabilité de présence/absence de *P. coracinus*. Lorsque l'on compare les parcelles témoins aux parcelles coupées, seuls les compartiments non scarifiés et les périodes 2 et 3 sont à prendre en compte. Les autres contrastes relatifs à l'effet principal des traitements sylvicoles peuvent se faire en considérant, soit seulement les périodes 2 et 3 avec les deux niveaux de scarification, soit les trois périodes mais seulement pour les sous-parcelles scarifiées. Seuls les résultats des contrastes pertinents sont présentés dans le tableau 10.

Biodiversité des espèces : Il semble que l'abondance moyenne d'espèces est presque la même entre les parcelles témoins et les parcelles coupées non scarifiées ($F = 3.78$ avec 1 et 53 degrés de liberté, $P = 0.0571$). Les parcelles avec deux trouées d'une part et celles avec quatre ou huit trouées d'autres part diffèrent quant à la biodiversité des espèces qu'on y trouve ($F = 10.34$, $P = 0.0074$). Il semble également y avoir un effet linéaire différent de la période selon le niveau de scarification ($F = 6.91$, $P = 0.0112$). En d'autres termes, la différence entre le nombre moyen d'espèces de carabidés en début et en fin de saison ne serait pas la même selon que les parcelles sont scarifiées ou non. L'abondance d'espèces diminue au cours de la saison, tant dans les parcelles traitées scarifiées que dans les non scarifiées, mais cette diminution serait plus rapide dans les sous-parcelles scarifiées (figure 1).

Abondance de l'espèce *P. coracinus* (effet aléatoire de pièges) : L'effet principal de la période est significatif, tout comme l'effet de l'interaction entre la période et la scarification. Les

contrastes linéaire et quadratique de l'effet principal de la période ont donc moins d'importance que ceux de l'interaction entre la scarification et la période. S'il y a une tendance linéaire ou quadratique du temps, on aimerait savoir comment elle se comporte selon que la scarification est présente ou non.

L'abondance de *P. coracinus* dans les parcelles témoins diffère de celle dans les parcelles éclaircies non scarifiées ($F = 24.41, P < 0.0001$). Il semble également que l'abondance lorsque deux trouées sont pratiquées dans une grande parcelle n'est pas la même que lorsque quatre ou huit trouées y sont dégagées ($F = 24.58, P < 0.0001$). En outre, l'effet linéaire du temps varie selon le niveau de scarification ($F = 7.18, P = 0.0098$) mais rien n'indique que son effet quadratique diffère selon le niveau de scarification ($P = 0.4196$). L'abondance semble diminuer au cours de la saison dans les parcelles non scarifiées, comme on peut le voir à la figure 2. En revanche, lorsque les parcelles sont scarifiées, l'abondance reste presque la même durant toute la saison.

Abondance de l'espèce *P. coracinus* (variable *offset*) : Le nombre de spécimens de l'espèce *P. coracinus* diffère significativement selon que le peuplement forestier de la parcelle non scarifiée est coupé (partiellement ou à blanc) ou pas du tout ($F = 27.62, P < 0.0001$). De même, on trouve un nombre moyen différent de spécimens dans les parcelles à deux trouées comparativement à celles où quatre et huit trouées sont dégagées ($F = 18.46, P = 0.0006$). Comme pour le modèle précédent, l'effet linéaire de la période semble être différent selon le niveau de scarification ($F = 5.08, P = 0.0276$). On constate également une tendance à la baisse de l'abondance de l'espèce au cours de la saison au niveau des parcelles non scarifiées (Figure 3).

Probabilité de présence/absence de *P. coracinus* : La probabilité moyenne de trouver l'espèce *P. coracinus* ne semble pas être la même dans les parcelles témoins et éclaircies ($F = 9.63, P = 0.0032$). Elle ne serait pas la même non plus entre les parcelles à deux trouées et celles à quatre ou huit trouées ($F = 13.04, P = 0.0007$).

Pour mieux appréhender et interpréter le sens de ces différences, nous estimerons les moyennes importantes dans la sous-section suivante.

III.2.9 Estimation des moyennes pertinentes à l'interprétation des contrastes

L'estimation des moyennes ne se fait pas de façon directe à cause des fonctions de lien. Il est nécessaire de convertir les valeurs trouvées dans les sorties SAS par l'application inverse des fonctions de lien utilisées. En effet, SAS donne les estimations des moyennes des moindres carrés et leurs différences sur l'échelle des fonctions de lien. Si un contraste est significatif, il y a intérêt à calculer les estimations des membres du contraste et du contraste lui-même.

Lorsque le lien est logarithmique, les moyennes des moindres carrés s'obtiennent en prenant l'exponentielle des estimations. L'exponentielle des différences donne un rapport des moyennes des moindres carrés. Nous donnons ci-dessous un exemple de calcul, la démarche étant la même pour les autres contrastes.

Pour le modèle d'analyse de l'abondance de l'espèce *P. coracinus*, on aimerait évaluer la différence entre la moyenne de l'abondance de l'espèce dans les parcelles témoins et la moyenne dans les parcelles avec éclaircies mais dans les secteurs non scarifiés. Le calcul se fait comme suit :

$$\begin{aligned}\log(\hat{\mu}_{TE}) - \log(\hat{\mu}_{CB+T2+T4+T8}) &= \log\left(\frac{\hat{\mu}_{TE}}{\hat{\mu}_{CB+T2+T4+T8}}\right) \\ &= 1.1074 - (-0.1242) \\ &= 1.2316\end{aligned}$$

Il suffit alors de prendre l'exponentielle de la différence estimée pour obtenir le rapport de moyennes :

$$\exp(1.2316) = 3.4267$$

Ce rapport signifie que *P. coracinus* est 3.4 fois plus abondant dans les parcelles témoins que dans les secteurs non scarifiés des parcelles traitées.

Si le lien est *logit* et que l'estimation trouvée est, disons x , alors la moyenne des moindres carrés de la probabilité $\hat{\pi}$ s'obtient par la formule suivante :

$$\hat{\pi} = \frac{\exp(x)}{1 + \exp(x)}.$$

Si l'on prend l'exponentielle de la différence estimée, disons z , des moyennes des moindres carrés de $\text{logit}(\pi_1)$ et de $\text{logit}(\pi_2)$, on obtient le rapport de cotes (*odds ratio*) des effets en question. En d'autres termes, on a :

$$\log\left(\frac{\pi_1}{1-\pi_1}\right) - \log\left(\frac{\pi_2}{1-\pi_2}\right) = \log\left(\frac{\frac{\pi_1}{1-\pi_1}}{\frac{\pi_2}{1-\pi_2}}\right) = z \Leftrightarrow \text{rapport des cotes}(\pi_1, \pi_2) = \exp(z)$$

Biodiversité des espèces : En moyenne, on estime le nombre d'espèces différentes respectivement à 1.027 et 1.440 dans les parcelles à deux trouées et celles à quatre ou huit trouées (tableau 11), soit un rapport de 1.402 (tableau 12). Cela signifie qu'en moyenne, le nombre d'espèces différentes dans les parcelles à quatre ou huit trouées représente près de 1.4 fois ce que l'on trouve dans les parcelles avec deux trouées.

On sait en plus que l'effet linéaire du temps est décroissant et que la biodiversité est plus riche en début de saison qu'à la fin de l'été pour les parcelles scarifiées ou non : le nombre d'espèces de carabidés en août est environ 60% de ce qu'il est en juin dans les sous-parcelles non scarifiées, et 38% dans les sous-parcelles scarifiées (tableau 12).

Abondance de l'espèce *P. coracinus* (effet aléatoire de pièges) : Le nombre moyen de spécimens de l'espèce *P. coracinus* est de plus de 3 par piège dans les parcelles témoins. Il n'est que de 0.883 dans les parcelles traitées non scarifiées (tableau 11), conduisant à un rapport moyen de 3.4 (tableau 12). Il y a donc en moyenne près de 3.4 fois plus de spécimens de cette espèce dans les parcelles témoins que dans les parcelles avec éclaircies non scarifiées.

L'espèce est 4.5 ($\cong 1/0.223$; tableau 12) fois plus abondante dans les parcelles avec quatre ou huit trouées que dans les parcelles à deux trouées, puisque les premières ont un nombre moyen de spécimens égal à 1.468, contre 0.327 pour les dernières. Pour les sous-parcelles scarifiées excluant donc la parcelles témoin, on a dénombré plus de spécimens en début de saison qu'à la fin. En août, le nombre de spécimens est réduit de 25% relativement à l'abondance observée sur juin (tableau 12). Lorsque l'on considère les secteurs non scarifiés des parcelles, cette tendance s'estompe (rapport = 0.905 \approx 1, figure 2).

Abondance de l'espèce *P. coracinus* (variable *offset*) : Selon ce modèle, les nombres moyens de spécimens de *P. coracinus* dans les parcelles témoins et les parcelles avec éclaircies non scarifiées sont estimés respectivement à 3.425 et 0.969, soit un rapport d'environ 3.534. On peut dire alors que les parcelles témoins ont en moyenne près trois fois et demie plus de spécimens que les parcelles traitées sans scarification.

Il a été capturé en moyenne 1.862 spécimens dans les parcelles avec deux trouées contre 1.736 dans les parcelles à quatre ou huit trouées, soit un rapport de 0.207. On trouve donc, en moyenne, presque 5 fois plus de spécimens de l'espèce dans les parcelles à quatre ou huit trouées que dans les parcelles à deux trouées (rapport = 0.207 \approx 1/5). Lorsque les parcelles sont non scarifiées, l'abondance est en moyenne presque la même en début et en fin de saison (rapport = 0.997 \approx 1). En revanche, s'il y a scarification, les spécimens sont plus abondants en début qu'en fin de saison (rapport entre l'abondance en août et en juin = 0.866, figure 3).

Probabilité de présence/absence de *P. coracinus* : En moyenne, la probabilité de présence de l'espèce *P. coracinus* dans les parcelles témoins est estimée à 0.898. Cette estimation chute à 0.507 dans les parcelles coupées non scarifiées, soit un rapport de cotes de 8.2. Il y a alors près de 8 fois plus de chances de capturer au moins un spécimen de l'espèce dans les parcelles témoins que dans les parcelles éclaircies non scarifiées. Quand on compare les parcelles à deux trouées à celles avec quatre ou huit trouées, on trouve un rapport de cotes de 9.5. En plaçant les pièges dans les parcelles à quatre ou huit trouées, on augmente d'environ dix fois la probabilité de capturer au moins un spécimen de l'espèce *P. coracinus* par rapport à des pièges placés dans des parcelles à deux trouées.

Comment interprète-t-on ces résultats? C'est la question à laquelle on tente d'apporter des éléments de réponse dans la prochaine et dernière section.

III.3 RÉSULTATS ET CONCLUSIONS

Un constat qui semble ressortir de cette étude est que les sites où s'est déroulée l'expérience ne regorgent pas d'insectes de la famille des carabidés. Au plus, cinq espèces différentes d'insectes de cette famille sont capturées dans un piège à une semaine donnée. Ce fait peut relever de plusieurs causes. Les conditions liées à la saison, qu'elles soient naturelles ou exceptionnelles, peuvent être défavorables à la reproduction des carabidés (ceci serait à confirmer auprès des entomologistes). Peut-être aussi la récolte n'a-t-elle pas été fructueuse, eu égard aux circonstances qui prévalaient lors des passages. On sait, en effet, que certains pièges ont été inondés.

Qu'à cela ne tienne, la biodiversité des espèces de carabidés peut être liée aux traitements sylvicoles. Il en va de même de la probabilité de présence ou d'absence de l'espèce *P. coracinus* et de son abondance. C'est donc dire que la coupe forestière affecte la biodiversité, même si le degré de perturbation dépend, tant soit peu, de la nature et de l'ampleur de la coupe.

La biodiversité semble être la même dans les parcelles témoins et dans les zones coupées non scarifiées. En considérant, non pas la biodiversité mais plutôt l'espèce la plus abondante, on trouve que les parcelles témoins abritent plus de spécimens, ce qui est confirmé par la probabilité de capturer ces spécimens dans ces zones. Cela voudrait peut-être dire que, contrairement à la biodiversité en général, l'espèce *P. coracinus* aime vivre dans son milieu naturel.

On sait que les parcelles témoins sont des zones non perturbées. Il est donc plus vraisemblable que les spécimens de cette espèce soient plus fréquents dans ces zones proches de leur milieu naturel que dans des zones avec éclaircies où il y a présence de risques : présence accrue des rayons solaires qui assèchent le sol, vulnérabilité aux prédateurs, entre autres.

Qu'il s'agisse de la biodiversité des insectes ou seulement de celle de l'espèce *P. coracinus*, la préférence va aux parcelles avec quatre ou huit trouées relativement à celles avec deux trouées. Si l'on considère les parcelles à quatre ou huit trouées comme étant des traitements intermédiaires entre la pratique de deux trouées et la coupe totale, on pourrait dire que les insectes ne préfèrent pas des traitements extrêmes mais plutôt un niveau de traitement intermédiaire. Plutôt que de se retrouver dans de grandes ouvertures dans la forêt, le *P. coracinus* préfère de petites ouvertures, même plus nombreuses. D'après le Docteur J. Klimaszewski, du Centre de Foresterie des Laurentides, les insectes rampants ont besoin d'humidité. La scarification et les trouées assèchent leur habitat. L'analyse montre donc que la taille des éclaircies a son importance sur l'abondance.

Sans scarification, la période a aussi son importance aussi bien sur la biodiversité que sur la présence/absence et l'abondance de l'espèce *P. coracinus*. En début de saison, où l'espèce est plus prolifique, la chance d'en capturer des spécimens est plus grande. Plus la saison avance, moins abondantes sont les captures. Avec la scarification, l'évolution du nombre d'espèces est quasiment nulle entre le début et la fin de la saison. Cela porte donc à croire que le début et la fin de la saison ont quelque chose à voir sur le rythme d'abondance des insectes. Est-ce là un début de liaison avec le cycle de vie de ces insectes ? Nous pensons que l'interprétation peut relever de facteurs intrinsèques aux espèces de la famille de carabidés. La réponse est donc aux entomologistes, qui devraient entériner ou réfuter ces résultats.

Notons tout de même que les modèles donnent des résultats assez semblables, à quelques exceptions près. Si les modèles d'analyse de la biodiversité et de l'abondance de l'espèce *P. coracinus* sont comparables (dans le sens qu'ils utilisent une variable *offset*), il n'en demeure pas moins que le modèle d'analyse de l'abondance de l'espèce *P. coracinus* avec effet aléatoire de piège a l'avantage de donner des résultats avec plus de finesse. En effet, l'agrégation supplémentaire des données peut engendrer des pertes d'information. Par ailleurs, les abondances et les probabilités moyennes sont affectées par des effets aléatoires différents selon le modèle. Quoiqu'il en soit, les résultats de ce projet ne sont pas directement comparables à ceux des études précédentes, dans la mesure où l'analyse dans ce document est restreinte à la

famille des carabidés et les données sont agrégées. Une consultation auprès des entomologistes serait nécessaire pour entériner nos résultats et les comparer à ce qui s'observe habituellement.

CONCLUSION

Les modèles linéaires généralisés mixtes, nous l'avons dit, sont très usités et nous avons eu l'occasion de le vérifier, en partie, dans cet essai. Parmi les atouts majeurs de ces modèles, signalons la possibilité qu'ils offrent de spécifier une distribution et une structure de covariance appropriées pour les termes d'erreur, ainsi que celle de prendre en compte des mesures répétées.

Outre qu'ils possèdent de nombreux avantages, les modèles GLMMs s'avèrent très pratiques. Ainsi avons-nous pu, grâce à eux, étudier un problème on ne peut plus d'actualité et aussi important que l'impact de la sylviculture sur la biodiversité des carabidés. À cet égard, nos résultats font nettement ressortir l'influence des pratiques sylvicoles sur l'abondance et la diversité des espèces. Nos conclusions montrent la pertinence de pratiquer des coupes forestières qui soient respectueuses de la forêt, de l'entomofaune et de l'environnement en général. Nous l'avons vu, l'impact de ces coupes varie significativement selon le nombre et la taille des trouées.

Au-delà de l'intérêt intrinsèque de l'application considérée, nos analyses mettent en lumière le fait que derrière une gamme d'avantages certains se dressent parfois des difficultés importantes dans la mise en œuvre des modèles linéaires généralisés mixtes. En particulier, la fiabilité de la macro GLIMMIX que nous avons utilisée pour nos analyses n'est pas totalement garantie. Il est cependant possible que cette lacune soit corrigée dans la version 9 de SAS, qui venait à peine d'être mise en marché au moment de la finalisation de cet essai.

Une difficulté récurrente à laquelle nous nous sommes butés concernait la convergence des algorithmes d'estimation, surtout en présence d'une surabondance de zéros dans les données. À plusieurs reprises, nous avons dû recourir à des agrégations de données qui ne semblaient pas souhaitables du point de vue scientifique, vu la perte d'information engendrée. Les modèles dits de Poisson avec excès de zéros (*ZIP : Zero-inflated Poisson*, en anglais) offrent toutefois une solution de rechange intéressante à ce problème de surabondance de zéros. Développé dans les années 1990, le modèle ZIP a été d'abord motivé par des problèmes en industrie manufacturière

où il est très appliqué. L'article de Diane Lambert paru en 1992 dans le *Journal of the American Statistical Association* est une référence clef sur le sujet.

Des études similaires à la nôtre ont été conduites par le passé en faisant appel au modèle ZIP. Yip (1988), par exemple, l'a utilisé pour modéliser le nombre d'insectes par feuille. Les hypothèses distributionnelles sous-tendues par le modèle ZIP peuvent toutefois s'avérer invalides si la présence d'excès de zéros n'est pas bien modélisée. Pour de plus amples informations sur ces modèles, consulter la bibliographie.

BIBLIOGRAPHIE

Abdel, H. S. & Piegorsch, W. W. (2002). Zero-inflated data. *Encyclopedia of Environmetrics*, 4, 2387-2391.

Brown, H. & Prescott, R. (1999). *Applied Mixed Models in Medicine*. Chichester: Wiley

Cameron, A. C. & Trivedi, P. K. (1986). Econometric models based count data: comparisons and applications of some estimators and tests. *Journal of Applied Econometrics*, 1, 29-53.

Dobson, A. J. (1990). *An Introduction to Generalized Linear Models*. London: Chapman and Hall.

Gurmu, S., Rilstone, P. & Stern, S. (1999). Semiparametric estimation of count regression models. *Journal of Econometrics*, 88, 123-150.

Lambert, D. (1992). Zero-inflated Poisson regression with an application to defects in manufacturing. *Technometrics*, 33, 1-14.

Li, C. S., Lu, J.-C., Park, J., Kim, K., Brinkley, P. A. & Peterson, J. P. (1999). Multivariate zero-inflated Poisson models and their applications. *Technometrics*, 41, 29-38.

Littell, R. C., Milliken, G. A., Stroup, W. W. & Wolfinger, R. D. (1996). *SAS System for Mixed Models*. Cary, NC: SAS Institute Inc., 633 pp.

Verbeke, G. & Molenberghs, G. (1997). *Linear Mixed Models in Practice*. New-York: Springer.

Wolfinger, R. & O'Connell, R. (1993). Generalized linear mixed models: A pseudo-likelihood approach. *Journal of Statistical Computation and Simulation*, 48, 233-243.

ANNEXES

TABLEAUX

Tableau 1 : Quelques caractéristiques des distributions binomiale et de Poisson

	Poisson	Binomiale
Densité	$f_y(y, \mu) = \exp(-\mu) \frac{\mu^y}{y!}$ $= \exp\left\{\frac{y \log(\mu) - \mu}{1} + (-1) \log(y!)\right\}$	$f(y, n, \pi) = \binom{n}{\pi} \pi^y (1-\pi)^{n-y}$ $= \exp\left\{\frac{y \log\left(\frac{\pi}{1-\pi}\right) + \log(1-\pi)^n}{1} + \log\binom{n}{y}\right\}$
θ	$\log(\mu)$	$\log\left(\frac{\pi}{1-\pi}\right)$
$b(\theta)$	μ	$n \log(1-\pi)$
$a(\phi)$	1	$1/n$
$c(y, \phi)$	$-\log(y!)$	$\log\binom{n}{y}$
moyenne	μ	$n\pi$
variance	μ	$n\pi(1-\pi)$

Tableau 2: Table d'analyse de la variance du modèle 1

N°	Sources de variation	Nombre de degré de liberté	Type d'effet
1	Blocs (Bloc)	3	Fixe ou aléatoire (aléatoire dans ce cas)
2	Traitements sylvicoles (Trait)	4	Fixe
3	Bloc x Trait (erreur grandes parcelles)	12	aléatoire
4	Scarification (Scari)	1	Fixe
5	Scari x Trait	3	Fixe
6	Bloc x Scari + Bloc x Trait x Scari (erreurs sous-parcelles)	12	aléatoire
7	Période (Periode)	2	Fixe
8	Periode x Trait	8	Fixe
9	Scari x Periode	2	Fixe
10	Trait x Scari x Periode	6	Fixe
11	Bloc x Scari x Periode + Bloc x Trait x Scari x Periode + Bloc x Trait x Periode + Bloc x Periode	54	aléatoire

Tableau 3 : Résultats obtenus en utilisant différentes structures de covariance pour les quatre modèles d'analyse

Types de matrice de variance-covariance	MODÈLES D'ANALYSE							
	Biodiversité des espèces		Abondance de l'espèce P. coracinus (effet aléatoire de piège)		Abondance de l'espèce P coracinus (avec variable offset)		Probabilité de présence/absence de l'espèce P. coracinus	
	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
1 : type=UN	67.6	64.0	1300.7	1295.1	247.2	242.3	-	-
2 : type=VC	68.5	67.2	1133.7	1131.2	240.5	238.7	1650.6	1647.6
3 : type=CS	68.3	67.1	968.5	965.4	242.3	239.8	1678.3	1675.3
4 :type=AR(1)	67.3	66.1	1123.3	1120.2	240.6	238.1	1792.8	1789.1
5 :type=TOEP	69.1	67.3	1166.6	1163.6	241.5	238.5	1521.2	1517.5

Tableau 4 : Les différentes étapes de réduction de la partie aléatoire pour le modèle d'analyse de la biodiversité

		Étapes			
		0*	1	2	3
Effets aléatoires enlevés		aucun	σ_c^2	σ_b^2	σ_a^2
AIC		88.6	92.5	90.6	88.9
BIC		87.3	90.6	89.4	90.4
Estimés des paramètres de covariance (Erreur standard)	σ_a^2	0.00	0.001 (0.004)	0.001 (0.004)	-
	σ_b^2	0.00	0.001 (0.008)	-	-
	σ_c^2	0.017 (0.015)	-	-	-
	σ_d^2	0.370 (0.074)	0.435 (0.076)	0.441 (0.071)	0.448 (0.071)

* Le modèle le modèle complet) est le modèle d'analyse de la biodiversité des espèces

Tableau 5 : Les différentes étapes de réduction de la partie aléatoire du modèle complet avec effet aléatoire de pièges pour l'analyse de l'abondance de l'espèce *P. coracinus*

		Étapes					
		0	1	2	3	4*	5
Effets aléatoires enlevés		Aucun	σ_f^2	σ_d^2	σ_c^2	σ_b^2	σ_a^2
AIC		1123.3	1133.8	1137.7	1123.3	1121.1	1123.3
BIC		1120.2	1131.3	1135.9	1120.2	1118.6	1128.2
Estimations des paramètres de covariance (erreur standard)	σ_a^2	0.000 (.)	0.000 (.)	0.000 (.)	0.000 (.)	0.000 (.)	-
	σ_b^2	0.004 (0.077)	0.415 (0.067)	0.000 (.)	0.004 (0.077)	-	0.004 (0.077)
	σ_c^2	0.000 (.)	0.000 (.)	0.000 (.)	-	0.000 (.)	0.000 (.)
	σ_d^2	0.376 (0.097)	0.313 (0.089)	-	0.376 (0.097)	0.377 (0.096)	0.376 (0.097)
	σ_f^2	0.207 (0.080)	-	0.127 (0.063)	0.207 (0.080)	0.207 (0.080)	0.207 (0.797)
	AR(1) subject= σ_d^2	0.102 (0.114)	0.112 (0.099)	0.424 (0.061)	0.102 (0.114)	0.104 (.110)	0.102 (0.114)
	Résidus	0.935 (0.116)	1.260 (0.133)	1.925 (0.177)	0.935 (0.116)	0.937 (0.115)	0.935 (0.116)

*Le modèle obtenu à l'étape 4 est le modèle final d'analyse

Tableau 6: Étapes de réduction de la partie aléatoire du modèle avec variable *offset* pour l'analyse de l'espèce *P. coracinus*

		Étapes			
		0	1	2	3
Effets aléatoires enlevés		aucun	σ_c^2	σ_b^2	σ_a^2
AIC		240.5	238.8	239.3	240.5
BIC		238.7	237.6	238.1	238.7
Estimés des paramètres de covariance (erreur standard)	σ_a^2	0.000 (-)	0.000 (-)	0.000 (-)	-
	σ_b^2	0.065 (0.082)	0.090 (0.067)	-	0.065 (0.082)
	σ_c^2	0.049 (0.073)	-	0.098 (0.068)	0.049 (0.073)
	Résidus	1.924	2.036	1.992	1.926
	Periode (subject= σ_c^2)	(0.362)	(0.357)	(0.369)	(0.362)

Tableau 7 : processus de réduction de la partie aléatoire pour le modèle complet d'analyse de la probabilité de présence/absence de l'espèce *P. coracinus*

		Étapes					
		0	1	2	3	4	5
Effets aléatoires enlevés		Aucun	σ_f^2	σ_d^2	σ_c^2	σ_b^2	σ_a^2
AIC		1650.6	1483.1	1465.5	1650.6	1643.7	1649.1
BIC		1647.6	1480.0	1462.5	1647.6	1641.3	1653.1
Estimations des paramètres de covariance (Erreur standard)	σ_a^2	0.126 (0.403)	0.096 (0.238)	0.074 (0.204)	0.126 (0.403)	0.157 (0.375)	-
	σ_b^2	0.197 (0.708)	0.389 (0.451)	0.326 (0.394)	0.197 (0.708)	-	0.314 (0.693)
	σ_c^2	0 (.)	0.024 (0.397)	0.067 (0.362)	-	0 (.)	0 (.)
	σ_d^2	3.083 (0.925)	1.029 (0.460)	-	3.083 (0.925)	3.082 (0.906)	3.090 (0.925)
	σ_f^2	2.702 (0.955)	-	0.567 (0.438)	2.702 (0.955)	2.705 (0.934)	2.715 (0.959)
	AR(1) subject= σ_d^2	0.489 (0.071)	0.776 (0.079)	0.873 (0.082)	0.489 (0.071)	0.491 (0.071)	0.488 (0.070)

Tableau 8 : Quelques statistiques d'ajustement des modèles

	Biodiversité des espèces	Abondance de <i>P. coracinus</i> (effet de piège)	Abondance de <i>P. coracinus</i> (variable offset)	Présence/Absence de <i>P. coracinus</i>
Déviance	26.51	278.212	164.762	299.111
Déviance standardisée	26.51	296.987	164.762	299.111
Dispersion	1.00	0.937	1.000	1.000

Tableau 9: Analyse des effets fixes des quatre modèles

effet	Biodiversité des espèces				<i>P. coracinus</i> (effet aléatoire de pièges)				<i>P. coracinus</i> (variable <i>offset</i>)				Présence/Absence <i>P. coracinus</i>			
	Num DF	Den DF	F value	Pr>F	Num DF	Den DF	F value	Pr>F	Num DF	Den DF	F value	Pr>F	Num DF	Den DF	F value	Pr>F
Trait	4	12	3.74	0.0336	4	54	12.07	<0.0001	4	15	12.90	<0.0001	4	12	6.48	0.0051
Scari	1	12	7.01	0.0213	1	54	0.25	0.6225	1	66	0.03	0.8651	1	12	0.55	0.4713
TraitxScari	3	12	1.61	0.2392	3	54	0.27	0.8475	3	66	0.15	0.9275	3	12	0.28	0.8379
Periode	2	53	48.94	<0.0001	2	54	3.59	0.0344	2	66	2.69	0.0753	2	48	0.17	0.8415
TraitxPeriode	8	53	1.45	0.1994	8	54	1.00	0.4466	8	66	1.05	0.4104	7	48	0.43	0.8806
Scari x Periode	2	53	3.93	0.0256	2	54	4.19	0.0204	2	66	3.21	0.0469	2	48	0.38	0.6890
TraitxScarixPeriode	6	53	1.42	0.2232	6	54	0.88	0.5171	6	66	0.48	0.8240	5	48	1.37	0.2528

Note : En gras, les p-valeurs < 5% (les effets fixes correspondants sont significatifs)

Tableau 10 : Résultats des contrastes pour les 4 modèles d'analyse

	Biodiversité des espèces				P. coracinus (effet de piège)				P. coracinus (variable offset)				Présence/Absence P. coracinus			
	Num DF	Den DF	F value	P value	Num DF	Den DF	F value	P value	Num DF	Den DF	F value	P value	Num DF	Den DF	F value	P value
Trait	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
TE vs CB T2 T4 T8 (NS)	1	53	3.78	0.0571	1	54	24.41	<.0001	1	66	27.62	<.0001	1	48	9.63	0.0032
CB vs T2 T4 T8	1	12	0.04	0.8390	1	54	0.97	0.3283	1	15	1.07	0.3176	1	48	0.00	0.9581
T2 vs (T4, T8)	1	12	10.34	0.0074	1	54	24.58	<.0001	1	15	18.46	0.0006	1	48	13.04	0.0007
T4 vs T8	1	53	0.02	0.9012	1	54	0.94	0.3355	1	15	0.25	0.6254	1	48	1.57	0.2164
Scari	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
TraitxScari	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
(CB vs T2 T4 T8)xScari	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
(T2 vs (T4, T8))xScari	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
(T4 vs T8)xScari	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Note : en gras les contrastes significatifs au seuil de 5%.

Tableau 10 : Résultats des contrastes pour les 4 modèles d'analyse (suite)

	Biodiversité des espèces				P. coracinus (effet de piège)				P. coracinus (variable offset)				Présence/Absence P. coracinus			
	Num DF	Den DF	F value	P value	Num DF	Den DF	F value	P value	Num DF	Den DF	F value	P value	Num DF	Den DF	F value	P value
Periode	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Effet linéaire	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Effet quadratique	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
TraitsPeriode	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
(TE vs CB T2 T4 T8 (NS))xPeriode	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
(CB vs T2 T4 T8)xPeriode	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
(T2 vs (T4, T8))xPeriode	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
(T2 vs (T4, T8))xPeriode	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
(T4 vs T8)xPeriode	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ScarixPeriode	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
ScarixPeriode linéaire	1	53	6.91	0.0112	1	54	7.18	0.0098	1	66	5.08	0.0276				
ScarixPeriode quadratique	1	53	2.38	0.1289	1	54	0.66	0.4196	1	66	0.59	0.4438				

Note : en gras les contrastes significatifs au seuil de 5%.

Tableau 10 : Résultats des contrastes pour les 4 modèles d'analyse (fin)

	Biodiversité des espèces				P. coracinus (effet de piège)				P. coracinus (variable offset)				Présence/Absence P. coracinus			
	Num DF	Den DF	F value	P value	Num DF	Den DF	F value	P value	Num DF	Den DF	F value	P value	Num DF	Den DF	F value	P value
TraitxScarixPeriode	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
(CB vs T2 T4 T8)xScarixPeriode linéaire	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
(CB vs T2 T4 T8)xScarixPeriode quadratique	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
(T2 vs (T4, T8))xScarixPeriode linéaire	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
(T2 vs (T4, T8))xScarixPeriode quadratique	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
(T4 vs T8)xScarixPeriode linéaire	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
(T4 vs T8)xScarixPeriode quadratique	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-

Tableau 11 : Estimation des moyennes des moindres carrés pertinents

Effets	Modèles			
	Biodiversité des espèces (nombre d'espèces/piège)	<i>P. coracinus</i> (effet aléatoire de pièges) nombre moyen de spécimens /piège)	<i>P. coracinus</i> (variable offset) nombre moyen de spécimens/piège	Présence/Absence <i>P. coracinus</i> Probabilité moyenne de présence de présence de l'espèce par piège
TE	-	3.026	3.425	0.898
CB T2 T4 T8@NS	-	0.883	0.969	0.507
T2	1.027	0.327	1.862	0.184
T4 T8	1.440	1.468	1.736	0.683

Tableau 12: Estimations des moyennes des contrastes significatifs

Effets	Modèles			
	Biodiversité des espèces (Rapport des nombres moyens d'espèces par piège)	<i>P. coracinus</i> (effet aléatoire de pièges) (Rapport des nombres moyens de spécimens de l'espèce par piège)	<i>P. coracinus</i> (variable offset) (Rapport des nombres moyens de spécimens de l'espèce par piège)	Présence/Absence <i>P. coracinus</i> (Odds ratio)
TE vs CB T2 T4 T8 (non scarifiée)	-	3.427	3.534	8.1503
T2 vs T4 T8	1.402	0.223	0.207	9.544
Scari x Période linéaire	-	-	-	-
Période linéaire (sans scarification)	0.586	0.905	0.997	-
Période linéaire (Avec scarification)	0.380	0.472	0.8668	- -

FIGURES

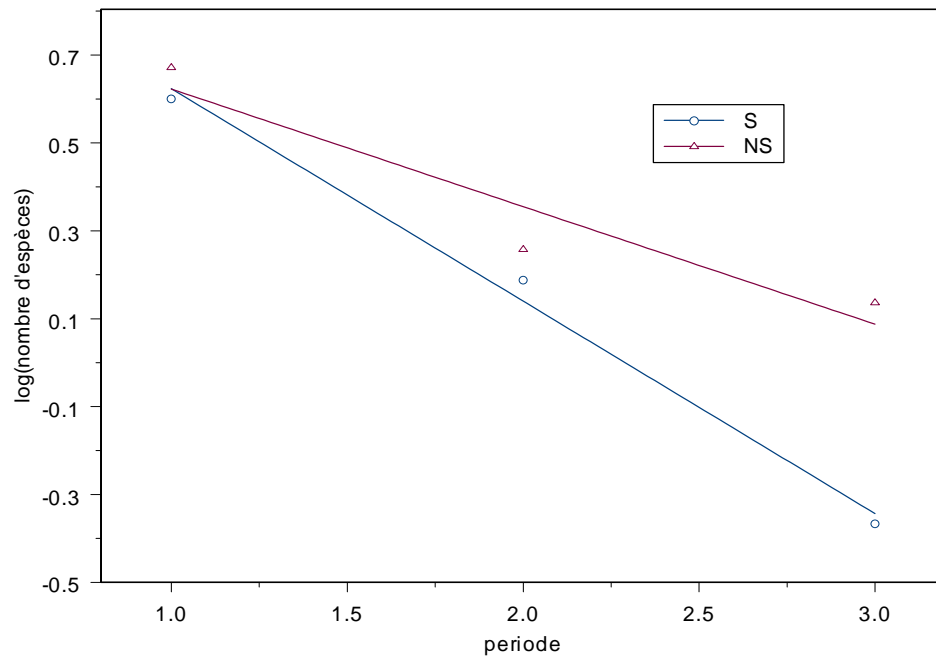


Figure 1 : logarithme du nombre d'espèces selon la scarification et la période (modèle d'analyse de la biodiversité des espèces)

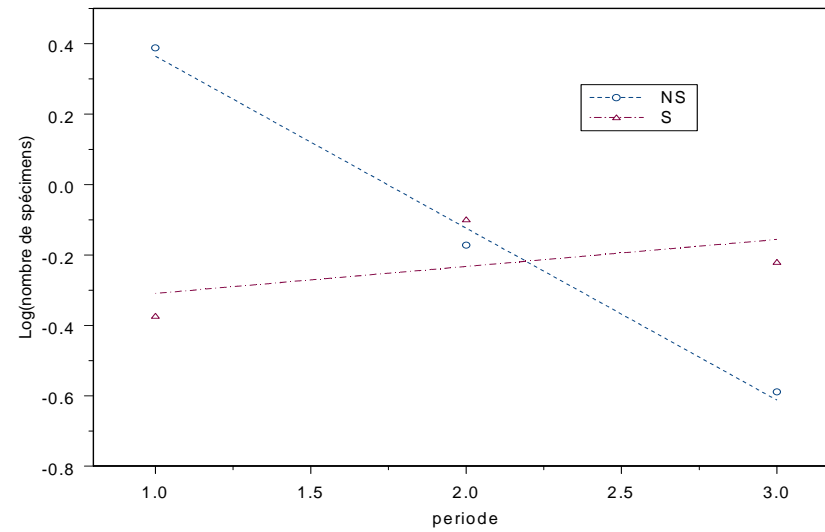


Figure 2 : Logarithme moyen du nombre de spécimens de l'espèce *P. coracinus* selon la scarification et la période (modèle d'analyse de l'abondance de l'espèce *P. coracinus* avec effet aléatoire de piège)

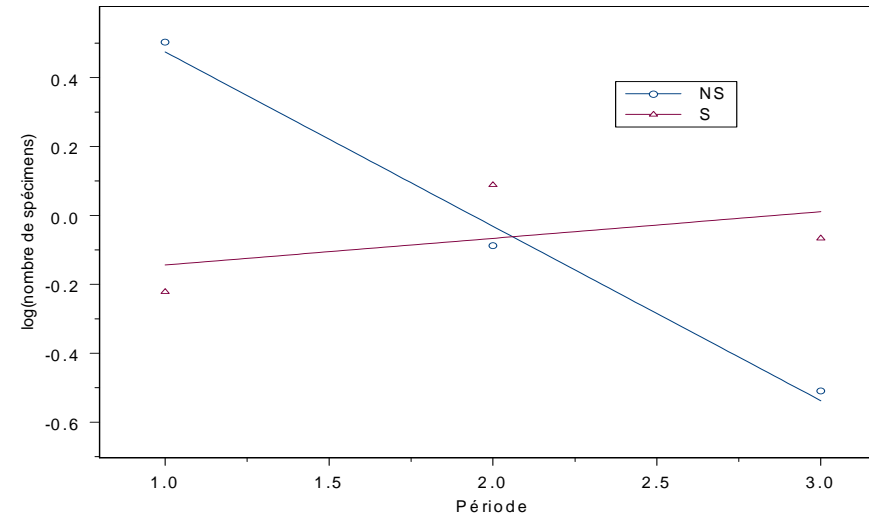


Figure 3 : Logarithme moyen du nombre de spécimens de l'espèce *P. coracinus* selon la période et la scarification (modèle d'analyse de l'abondance de l'espèce *P. coracinus* avec variable *offset*)

PROGRAMME SAS :

```

*****;
* PROGRAMME SAS POUR LE MODÈLE D'ANALYSE DE L'ESPÈCE P. CORACINUS ;
* AVEC EFFET ALÉATOIRE DE PIÈGE ;
*****;

*****;
* EXPLORATION DES DONNÉES ;
*****;

options nodate pageno=1 linesize=65 pagesize=45;

*importation du fichier d'origine de SAS;

PROC IMPORT OUT = sasuser.carabides
  DATAFILE = "P:\essai\CarabidaeSSAM2001_2003_09_15.xls"
  DBMS = EXCEL2000 REPLACE;
  RANGE="DATA$";
  GETNAMES = YES;
RUN;

*on ne conserve que la famille des carabidés et l'espèce "Pterostichus ; *coracinus (Newman)";
* conversion de trap1 de alphanumérique en numérique;
* on opte pour semaine 1, 2, ...13 plutôt que des dates;
* les parcelles témoins sont non scarifiées donc le X est changé en N ; *(non scarifié);

data sasuser.caral;
set sasuser.carabides;
if Famille='Carabidae' and Identification='Pterostichus coracinus (Newman)';
Trap1 = substr(Trap,5,3);
Trap2 = input(Trap1,8.0);
if Period='29-V - 5-VI' or Period='30-V - 6-VI' then Semaine=1;
if Period='5-VI - 12-VI' or Period='6-VI - 13-VI' then Semaine=2;
if Period='12-VI - 19 VI' or Period='13-VI - 20 VI' then Semaine=3;
if Period='19-VI - 26-VI' or Period='20-VI - 27-VI' then Semaine=4;
if Period='26-VI - 3-VII' or Period='27-VI - 4-VII' then Semaine=5;

```

```

if Period='3-VII - 10-VII' or Period='4-VII - 11-VII' then Semaine=6;
if Period='10-VII - 17-VII' or Period='11-VII - 18-VII' then Semaine=7;
if Period='17-VII - 24-VII' or Period='18-VII - 25-VII' then Semaine=8;
if Period='24-VII - 31-VII' or Period='25-VII - 1-VIII' then Semaine=9;
if Period='31-VII - 7-VIII' or Period='1-VIII - 8-VIII' then Semaine=10;
if Period='7-VIII - 14-VIII' or Period='8-VIII - 15-VIII' then Semaine=11;
if Period='14-VIII - 21-VIII' or Period='15-VIII - 22-VIII' then Semaine=12;
if Period='21-VIII - 28-VIII' or Period='22-VIII - 29-VIII' then Semaine=13;
if Scarification='N' or Scarification='X' then Scarification='N';
else Scarification='Y';
  keep Block Trap2 Treatment1 Scarification Semaine Speciesno;
run;

*****;
* Manipulation des données afin de les apprêter à l'analyse ;
*****;

*on considère le bloc C où se trouve les pièges 1 à 30;
data blocC;
set sasuser.caral;
if block='C';
run;

*création de piegel contenant les pièges du bloc C;
data piegel;
do trap1=1 to 30;
output;
end;
keep trap1;
run;

*on crée un fichier contenant les 13 semaines;
proc sort data=sasuser.caral;
by Semaine;
run;

data semain;
set sasuser.caral;
by Semaine;

```

```
if first.Semaine;
keep Semaine;
run;

* on parcourt chacun des pièges pendant les 13 semaines;
proc sql;
create table struct1 as
select a.trap1, b.Semaine
from piegel as a, semaine as b
order by trap1, Semaine;
quit;

* Le fichier complet C contient les dénombrements de tous les pièges ; *pendant les 13 semaines;
* ainsi que les types de traitement et de scarification;
proc sql;
create table completC as
select a.*, b.*
from struct1 as a left join BlocC as b
on a.trap1 = b.Trap2 and
a.Semaine = b.Semaine
order by trap1, Semaine ;
quit;

* correspondances entre les pièges, les traitements et scarification;
* les dénombrements manquants sont remplacés par des 0 et on ne garde ; *pas trap2, trap1 suffit;
data completC;
set completC;
if Block='' then Block='C';
if ((trap1=1 or trap1=2 or trap1=3) and Treatment1='') then Treatment1='CP' ;
if ((trap1=4 or trap1=5 or trap1=6) and Treatment1='') then Treatment1='CP' ;
if ((trap1=7 or trap1=8 or trap1=9) and Treatment1='') then Treatment1='C6' ;
if ((trap1=10 or trap1=11 or trap1=12) and Treatment1='') then Treatment1='C6' ;
if ((trap1=13 or trap1=14 or trap1=15) and Treatment1='') then Treatment1='C2' ;
if ((trap1=16 or trap1=17 or trap1=18) and Treatment1='') then Treatment1='C2' ;
if ((trap1=19 or trap1=20 or trap1=21) and Treatment1='') then Treatment1='C4' ;
if ((trap1=22 or trap1=23 or trap1=24) and Treatment1='') then Treatment1='C4' ;
if ((trap1=25 or trap1=26 or trap1=27 or trap1=28 or trap1=29 or trap1=30) and Treatment1='') then
Treatment1='F' ;
```

```

if ((trap1=1 or trap1=2 or trap1=3) and Scarification='') then Scarification='Y' ;
if ((trap1=4 or trap1=5 or trap1=6) and Scarification='') then Scarification='N' ;
if ((trap1=7 or trap1=8 or trap1=9) and Scarification='') then Scarification='Y' ;
if ((trap1=10 or trap1=11 or trap1=12) and Scarification='') then Scarification='N' ;
if ((trap1=13 or trap1=14 or trap1=15) and Scarification='') then Scarification='Y' ;
if ((trap1=16 or trap1=17 or trap1=18) and Scarification='') then Scarification='N' ;
if ((trap1=19 or trap1=20 or trap1=21) and Scarification='') then Scarification='Y' ;
if ((trap1=22 or trap1=23 or trap1=24) and Scarification='') then Scarification='N' ;
if ((trap1=25 or trap1=26 or trap1=27 or trap1=28 or trap1=29 or trap1=30) and Scarification='') then
Scarification='N' ;
if speciesno='.' then speciesno=0;
drop Trap2;
run;

*****;

* On fait exactement les mêmes manipulations pour      ;
* les trois autres blocs à savoir les bloc D, A et B ;
*****;

*bloc D contenant les pièges de 31 à 60;
data blocD;
set sasuser.caral;
if block='D';
run;
data piege2;
do trap1=31 to 60;
output;
end;
keep trap1;
run;

proc sql;
create table struct2 as
select a.trap1, b.Semaine
from piege2 as a, semain as b
order by trap1, Semaine;
quit;

```

```
proc sql;
create table completD as
select a.*, b.*
from struct2 as a left join BlocD as b
on    a.trap1 = b.Trap2 and
      a.Semaine = b.Semaine
order by trap1, Semaine ;
quit;

data completD;
set completD;
if Block='' then Block='D';
if ((trap1=31 or trap1=32 or trap1=33) and Treatment1='') then Treatment1='CP' ;
if ((trap1=34 or trap1=35 or trap1=36) and Treatment1='') then Treatment1='CP' ;
if ((trap1=37 or trap1=38 or trap1=39) and Treatment1='') then Treatment1='C6' ;
if ((trap1=40 or trap1=41 or trap1=42) and Treatment1='') then Treatment1='C6' ;
if ((trap1=43 or trap1=44 or trap1=45) and Treatment1='') then Treatment1='C2' ;
if ((trap1=46 or trap1=47 or trap1=48) and Treatment1='') then Treatment1='C2' ;
if ((trap1=49 or trap1=50 or trap1=51) and Treatment1='') then Treatment1='C4' ;
if ((trap1=52 or trap1=53 or trap1=54) and Treatment1='') then Treatment1='C4' ;
if ((trap1=55 or trap1=56 or trap1=57 or trap1=58 or trap1=59 or trap1=60) and Treatment1='') then
Treatment1='F' ;

if ((trap1=31 or trap1=32 or trap1=33) and Scarification='') then Scarification='Y' ;
if ((trap1=34 or trap1=35 or trap1=36) and Scarification='') then Scarification='N' ;
if ((trap1=37 or trap1=38 or trap1=39) and Scarification='') then Scarification='Y' ;
if ((trap1=40 or trap1=41 or trap1=42) and Scarification='') then Scarification='N' ;
if ((trap1=43 or trap1=44 or trap1=45) and Scarification='') then Scarification='Y' ;
if ((trap1=46 or trap1=47 or trap1=48) and Scarification='') then Scarification='N' ;
if ((trap1=49 or trap1=50 or trap1=51) and Scarification='') then Scarification='Y' ;
if ((trap1=52 or trap1=53 or trap1=54) and Scarification='') then Scarification='N' ;
if ((trap1=55 or trap1=56 or trap1=57 or trap1=58 or trap1=59 or trap1=60) and Scarification='') then
Scarification='N' ;

if speciesno='.' then speciesno=0;
drop Trap2;
run;
```

```
*bloc A contient les pièges 61 à 90;
data blocA;
set sasuser.caral;
if block='A';
run;

data piege3;
do trap1=61 to 90;
output;
end;
keep trap1;
run;

proc sql;
create table struct3 as
select a.trap1, b.Semaine
from piege3 as a, semain as b
order by trap1, Semaine;
quit;

proc sql;
create table completA as
select a.*, b.*
from struct3 as a left join BlocA as b
on a.trap1 = b.Trap2 and
a.Semaine = b.Semaine
order by trap1, Semaine ;
quit;

data completA;
set completA;
if Block=' ' then Block='A';
if ((trap1=61 or trap1=62 or trap1=63) and Treatment1='') then Treatment1='CP' ;
if ((trap1=64 or trap1=65 or trap1=66) and Treatment1='') then Treatment1='CP' ;
if ((trap1=67 or trap1=68 or trap1=69) and Treatment1='') then Treatment1='C6' ;
if ((trap1=70 or trap1=71 or trap1=72) and Treatment1='') then Treatment1='C6' ;
if ((trap1=73 or trap1=74 or trap1=75) and Treatment1='') then Treatment1='C2' ;
if ((trap1=76 or trap1=77 or trap1=78) and Treatment1='') then Treatment1='C2' ;
```

```
if ((trap1=79 or trap1=80 or trap1=81) and Treatment1='') then Treatment1='C4' ;
if ((trap1=82 or trap1=83 or trap1=84) and Treatment1='') then Treatment1='C4' ;
if ((trap1=85 or trap1=86 or trap1=87 or trap1=88 or trap1=89 or trap1=90) and Treatment1='') then
Treatment1='F' ;

if ((trap1=61 or trap1=62 or trap1=63) and Scarification='') then Scarification='Y' ;
if ((trap1=64 or trap1=65 or trap1=66) and Scarification='') then Scarification='N' ;
if ((trap1=67 or trap1=68 or trap1=69) and Scarification='') then Scarification='Y' ;
if ((trap1=70 or trap1=71 or trap1=72) and Scarification='') then Scarification='N' ;
if ((trap1=73 or trap1=74 or trap1=75) and Scarification='') then Scarification='Y' ;
if ((trap1=76 or trap1=77 or trap1=78) and Scarification='') then Scarification='N' ;
if ((trap1=79 or trap1=80 or trap1=81) and Scarification='') then Scarification='Y' ;
if ((trap1=82 or trap1=83 or trap1=84) and Scarification='') then Scarification='N' ;
if ((trap1=85 or trap1=86 or trap1=87 or trap1=88 or trap1=89 or trap1=90) and Scarification='') then
Scarification='N' ;

if speciesno='.' then speciesno=0;
drop Trap2;
run;

*bloc B contient les pièges 91 à 120;
data blocB;
set sasuser.caral;
if block='B';
run;

data piege4;
do trap1=91 to 120;
output;
end;
keep trap1;
run;

proc sql;
create table struct4 as
select a.trap1, b.Semaine
from piege4 as a, semain as b
order by trap1, Semaine; quit;
```

```
proc sql;
create table completB as
select a.*, b.*
from struct4 as a left join BlocB as b
on    a.trap1 = b.Trap2 and
      a.Semaine = b.Semaine
order by trap1, Semaine ;
quit;

data completB;
set completB;
if Block='' then Block='B';
if ((trap1=91 or trap1=92 or trap1=93) and Treatment1='') then Treatment1='CP' ;
if ((trap1=94 or trap1=95 or trap1=96) and Treatment1='') then Treatment1='CP' ;
if ((trap1=97 or trap1=98 or trap1=99) and Treatment1='') then Treatment1='C6' ;
if ((trap1=100 or trap1=101 or trap1=102) and Treatment1='') then Treatment1='C6' ;
if ((trap1=103 or trap1=104 or trap1=105) and Treatment1='') then Treatment1='C2' ;
if ((trap1=106 or trap1=107 or trap1=108) and Treatment1='') then Treatment1='C2' ;
if ((trap1=109 or trap1=110 or trap1=111) and Treatment1='') then Treatment1='C4' ;
if ((trap1=112 or trap1=113 or trap1=114) and Treatment1='') then Treatment1='C4' ;
if ((trap1=115 or trap1=116 or trap1=117 or trap1=118 or trap1=119 or trap1=120) and Treatment1='') then
Treatment1='F' ;

if ((trap1=91 or trap1=92 or trap1=93) and Scarification='') then Scarification='Y' ;
if ((trap1=94 or trap1=95 or trap1=96) and Scarification='') then Scarification='N' ;
if ((trap1=97 or trap1=98 or trap1=99) and Scarification='') then Scarification='Y' ;
if ((trap1=100 or trap1=101 or trap1=102) and Scarification='') then Scarification='N' ;
if ((trap1=103 or trap1=104 or trap1=105) and Scarification='') then Scarification='Y' ;
if ((trap1=106 or trap1=107 or trap1=108) and Scarification='') then Scarification='N' ;
if ((trap1=109 or trap1=110 or trap1=111) and Scarification='') then Scarification='Y' ;
if ((trap1=112 or trap1=113 or trap1=114) and Scarification='') then Scarification='N' ;
if ((trap1=115 or trap1=116 or trap1=117 or trap1=118 or trap1=119 or trap1=120) and Scarification='') then
Scarification='N' ;

if speciesno='.' then speciesno=0;
drop Trap2; run;
```

```
* fusion des données des 4 blocs;
* on ne garde que les 12 premières semaines;
data sasuser.donnees;
set completC completD completA completB;
if semaine=13 then delete;
run;

proc transpose data=sasuser.donnees out=sasuser.donnees1;
by trap1 Block Treatment1 Scarification;
id Semaine;
var speciesno;
run;

* agrégation des dénombrements en faisant la somme par 4 semaines;
data sasuser.donnees1;
set sasuser.donnees1;
  semaine1234=_1+_2+_3+_4;
  semaine5678=_5+_6+_7+_8;
  semaine9101112=_9+_10+_11+_12;
output;
run;

*on transforme de telle sorte que la période soit univariée;
data sasuser.donnees2;
set sasuser.donnees1;
periode=1; compte=semaine1234; output;
periode=2; compte=semaine5678; output;
periode=3; compte=semaine9101112; output;
rename Treatment1=Treat Scarification=Scari ;
keep trap1 block Treatment1 Scarification periode compte;
run;

* nombre de fréquences nulles;
proc sql;
select count(compte) as c
from sasuser.donnees2
where compte=0; quit;
```

```

*****;
*MODÈLE AVEC EFFET ALÉATOIRE DE PIÈGE EN UTILISANT LA MACRO ;
* GLIMMIX DE SAS: ABONDANCE DE L'ESPÈCE P. CORACINUS      ;
*****;
options nomprint nomlogic nosymbolgen;

%glimmix(data=sasuser.donnees2,
  procopt= ratio covtest,
  stmts=%str(
    class Block Treat Scari periode trap1;
    model compte=Treat|Scari|periode;
    random trap1(Block Treat Scari) Block*Treat*Scari*periode/solution;
    repeated periode / subject=trap1(Block Treat Scari) type=AR(1);
    lsmeans Treat periode Scari*periode / pdiff cl;
estimate 'temoin' intercept 3
          Treat 0 0 0 0 3
          Scari 3 0
          periode 1 1 1
          Treat*Scari 0 0
                    0 0
                    0 0
                    0 0
                    3
          Treat*periode 0 0 0
                        0 0 0
                        0 0 0
                        0 0 0
                        1 1 1
          Scari*periode 1 1 1
                       0 0 0
          Treat*Scari*periode 0 0 0 0 0 0
                              0 0 0 0 0 0
                              0 0 0 0 0 0
                              0 0 0 0 0 0
                              1 1 1/divisor=3;
estimate 'T8 T4 T2 CB@NS' intercept 12
          Treat 3 3 3 3 0
                    Scari 12 0

```

```

      periode 4 4 4
      Treat*Scari 3 0
                    3 0
                    3 0
                    3 0
                    0
      Treat*periode 1 1 1
                    1 1 1
                    1 1 1
                    0 0 0
      Scari*periode 4 4 4
                    0 0 0
      Treat*Scari*periode 1 1 1 0 0 0
                          1 1 1 0 0 0
                          1 1 1 0 0 0
                          1 1 1 0 0 0
                          1 1 1 0 0 0
                          0 0 0/divisor=12;

estimate 'T2' intercept 6
treat 0 0 6 0 0
scari 3 3
treat*scari 0 0
            0 0
            3 3
            0 0
            0
periode 2 2 2
treat*periode 0 0 0
              0 0 0
              2 2 2
              0 0 0
              0 0 0
scari*periode 1 1 1
              1 1 1

```

```

treat*scari*periode 0 0 0 0 0 0
                   0 0 0 0 0 0
                   1 1 1 1 1 1
                   0 0 0 0 0 0
                   0 0 0/divisor=6;
estimate 'T8 et T4' intercept 12
          Treat 6 6 0 0 0
          Scari 6 6
          periode 4 4 4
          Treat*Scari 3 3
                    3 3
                    0 0
                    0 0
                    0
          Treat*periode 2 2 2
                      2 2 2
                      0 0 0
                      0 0 0
                      0 0 0
          Scari*periode 2 2 2
                      2 2 2
          Treat*Scari*periode 1 1 1 1 1 1
                              1 1 1 1 1 1
                              0 0 0 0 0 0
                              0 0 0 0 0 0
                              0 0 0 /divisor=12;
estimate 'PeriodeLinSca' intercept 8
          Treat 2 2 2 2 0
          Scari 0 8
          periode 4 0 4
          Treat*Scari 0 2
                    0 2
                    0 2
                    0 2
                    0

```

```

Treat*periode 1 0 1
                1 0 1
                1 0 1
                1 0 1
                0 0 0
Scari*periode  0 0 0
                4 0 4
Treat*Scari*periode 0 0 0 1 0 1
                    0 0 0 1 0 1
                    0 0 0 1 0 1
                    0 0 0 1 0 1
                    0 0 0 / divisor=8 cl;
estimate 'PeriodeLinNS' intercept 8
        Treat 2 2 2 2 0
            Scari 8 0
            periode 4 0 4
            Treat*Scari 2 0
                    2 0
                    2 0
                    2 0
                    0
            Treat*periode 1 0 1
                            1 0 1
                            1 0 1
                            1 0 1
                            0 0 0
            Scari*periode 4 0 4
                            0 0 0
            Treat*Scari*periode 1 0 1 0 0 0
                                1 0 1 0 0 0
                                1 0 1 0 0 0
                                1 0 1 0 0 0
                                0 0 0 / divisor=8 cl;
estimate 'Scari@S*Periode@1' intercept 4
        treat 1 1 1 1 0
        scari 0 4

```

```

      treat*scari 0 1
                0 1
                0 1
                0 1
                0
      periode 4 0 0
      treat*periode 1 0 0
                  1 0 0
                  1 0 0
                  1 0 0
                  0 0 0
      scari*periode 0 0 0
                  4 0 0
      treat*scari*periode 0 0 0 1 0 0
                        0 0 0 1 0 0
                        0 0 0 1 0 0
                        0 0 0 1 0 0
                        0 0 0/divisor=4;
estimate 'Scari@S*Periode@2' intercept 4
      treat 1 1 1 1 0
      scari 0 4
      treat*scari 0 1
                0 1
                0 1
                0 1
                0
      periode 0 4 0
      treat*periode 0 1 0
                  0 1 0
                  0 1 0
                  0 1 0
                  0 0 0
      scari*periode 0 0 0
                  0 4 0

```

```

treat*scari*periode 0 0 0 0 1 0
                    0 0 0 0 1 0
                    0 0 0 0 1 0
                    0 0 0 0 1 0
                    0 0 0/divisor=4;
estimate 'Scari@S*Periode@3' intercept 4
    treat 1 1 1 1 0
    scari 0 4
    treat*scari 0 1
                0 1
                0 1
                0 1
                0
    periode 0 0 4
    treat*periode 0 0 1
                  0 0 1
                  0 0 1
                  0 0 1
                  0 0 0
    scari*periode 0 0 0
                  0 0 4

treat*scari*periode 0 0 0 0 0 1
                    0 0 0 0 0 1
                    0 0 0 0 0 1
                    0 0 0 0 0 1
                    0 0 0/divisor=4;
estimate 'Scari@NPeriode@1' intercept 4
treat 1 1 1 1 0
scari 4 0
treat*scari 1 0
            1 0
            1 0
            1 0
            0
periode 4 0 0

```

```

treat*periode 1 0 0
                1 0 0
                1 0 0
                1 0 0
                0 0 0
scari*periode  4 0 0
                0 0 0
treat*scari*periode 1 0 0 0 0 0
                    1 0 0 0 0 0
                    1 0 0 0 0 0
                    1 0 0 0 0 0
                    0 0 0/divisor=4;
estimate 'Scri@NPeriode@2' intercept 4
treat 1 1 1 1 0
scari 4 0
treat*scari 1 0
             1 0
             1 0
             1 0
             0
periode 0 4 0
treat*periode 0 1 0
              0 1 0
              0 1 0
              0 1 0
              0 0 0
scari*periode 0 4 0
              0 0 0
treat*scari*periode 0 1 0 0 0 0
                    0 1 0 0 0 0
                    0 1 0 0 0 0
                    0 1 0 0 0 0
                    0 0 0/divisor=4;
estimate 'Scri@NPeriode@3' intercept 4
treat 1 1 1 1 0
scari 4 0

```

```

treat*scari 1 0
            1 0
            1 0
            1 0
            0
periode 0 0 4

treat*periode 0 0 1
              0 0 1
              0 0 1
              0 0 1
              0 0 1
              0 0 0
scari*periode 0 0 4
              0 0 0
treat*scari*periode 0 0 1 0 0 0
                   0 0 1 0 0 0
                   0 0 1 0 0 0
                   0 0 1 0 0 0
                   0 0 1 0 0 0
                   0 0 0/divisor=4;
estimate 'ScariN*PLin' periode -4 0 4
        Treat*periode -1 0 1
                       -1 0 1
                       -1 0 1
                       -1 0 1
                       0 0 0
        Scari*periode -4 0 4
                       0 0 0
        Treat*Scari*periode -1 0 1 0 0 0
                              -1 0 1 0 0 0
                              -1 0 1 0 0 0
                              -1 0 1 0 0 0
                              0 0 0 / divisor=4;
estimate 'ScariY*PLin' periode -4 0 4
        Treat*periode -1 0 1
                       -1 0 1
                       -1 0 1
                       -1 0 1
                       0 0 0

```

```

Scari*periode 0 0 0
               -4 0 4
Treat*Scari*periode 0 0 0 -1 0 1
                   0 0 0 -1 0 1
                   0 0 0 -1 0 1
                   0 0 0 -1 0 1
                   0 0 0 / divisor=4;
estimate 'TE vs éclaircies' Treat -3 -3 -3 -3 12
                          Treat*Scari -3 0
                                      -3 0
                                      -3 0
                                      -3 0
                                      12
                          Treat*periode -1 -1 -1
                                          -1 -1 -1
                                          -1 -1 -1
                                          -1 -1 -1
                                          4 4 4
Treat*Scari*periode -1 -1 -1 0 0 0
                   -1 -1 -1 0 0 0
                   -1 -1 -1 0 0 0
                   -1 -1 -1 0 0 0
                   4 4 4 / divisor=12;
estimate 'T2 vs (T8 et T4)' Treat -6 -6 12 0 0

                          Treat*Scari -3 -3
                                      -3 -3
                                      6 6
                                      0 0
                                      0
                          Treat*periode -2 -2 -2
                                          -2 -2 -2
                                          4 4 4
                                          0 0 0
                                          0 0 0

```

```

Treat*Scari*periode -1 -1 -1 -1 -1 -1
                    -1 -1 -1 -1 -1 -1
                    2 2 2 2 2 2
                    0 0 0 0 0 0
                    0 0 0/ divisor=12;
contrast 'Scari*PLin' Scari*periode -4 0 4
                                   4 0 -4
Treat*Scari*periode -1 0 1 1 0 -1
                    -1 0 1 1 0 -1
                    -1 0 1 1 0 -1
                    -1 0 1 1 0 -1
                    0 0 0;
Contrast 'Scari*PQua' Scari*periode -4 8 -4
                                   4 -8 4
Treat*Scari*periode -1 2 -1 1 -2 1
                    -1 2 -1 1 -2 1
                    -1 2 -1 1 -2 1
                    -1 2 -1 1 -2 1
                    0 0 0;

contrast 'TE vs éclaircies'
Treat -3 -3 -3 -3 12
Treat*Scari -3 0
            -3 0
            -3 0
            -3 0
            12
Treat*periode -1 -1 -1
              -1 -1 -1
              -1 -1 -1
              -1 -1 -1
              4 4 4

Treat*Scari*periode -1 -1 -1 0 0 0
                    -1 -1 -1 0 0 0
                    -1 -1 -1 0 0 0
                    -1 -1 -1 0 0 0
                    4 4 4 ;
contrast 'CB vs T2 T4 T8' Treat -1 -1 -1 3 0 ;

```

```

        contrast 'T2 vs T4 T8' Treat -1 -1 2 0 0 ;
        contrast 'T4 vs T8' Treat -1 1 0 0 0 ;

    ),
    error=poisson,
    link=log
);

*Les estimations sont sauvegardées dans oumara.lsm pour ensuite faire des graphiques avec SPLUS;
data oumara.lsm;
set _est;
run;

*****;
* PROGRAMME POUR L'ANALYSE DE L'ABONDANCE DE L'ESPÈCE P. CORACINUS ;
*
* MODÈLE AVEC VARIABLE OFFSET ;
*****;

*****;
* QUELQUES MANIPULATIONS DES DONNÉES ;
*****;
* Pour chaque grande parcelle, on somme le nombre d'insectes sur tous les pièges des sous-parcelles
scarifiées d'une part;
* et sur tous les pièges des sous parcelles non scarifiées d'autre part;
proc sql;
create table sasuser.toutes as
select Block, Treat, Scari, Periode, sum(compte) as compt
from sasuser.donnees2
group by Block, Treat, Scari, Periode;
quit;

*attribution des valeurs de la variable offset;
data sasuser.donneesfinales;
set sasuser.toutes;
if Treat='CP' or Treat='C2' or Treat='C6' or Treat='C4' then nb_piege=3;
else nb_piege=6;
run;

```

```

*on crée une variable qui est le logarithme de la variable offset;
*elle est l'argument de l'instruction offset dans GLIMMIX;
data sasuser.donneesfinales1;
set sasuser.donneesfinales;
logoffset=log(nb_piege);
run;

*****;
*   MODÈLE STATISTIQUE FINAL AVEC VARIABLE OFFSET POUR L'ANALYSE DE   *
L'ESPÈCE P. CORACINUS           ;
*****;

options nomprint nomlogic nosymbolgen;

%glimmix(data=sasuser.donneesfinales1,
         procopt= ratio covtest,
         stmts=%str(
           class Block Treat Scari periode;
           model compt=Treat|Scari|periode/solution;
           random Block*Treat/solution;
           repeated periode / subject=Block*Treat*Scari ;
           lsmeans Treat|Scari|periode / pdiff cl;
estimate 'temoin' intercept 3
          Treat 0 0 0 0 3
          Scari 3 0
          periode 1 1 1

          Treat*Scari 0 0
                      0 0
                      0 0
                      0 0
                      3
          Treat*periode 0 0 0
                       0 0 0
                       0 0 0
                       0 0 0
                       1 1 1

```

```

Scari*periode 1 1 1
              0 0 0
Treat*Scari*periode 0 0 0 0 0 0
                   0 0 0 0 0 0
                   0 0 0 0 0 0
                   0 0 0 0 0 0
                   1 1 1/divisor=3 e;

estimate 'T8 T4 T2 CB@NS' intercept 12
        Treat 3 3 3 3 0
           Scari 12 0
           periode 4 4 4
           Treat*Scari 3 0
                    3 0
                    3 0
                    3 0
                    0

Treat*periode 1 1 1
              1 1 1
              1 1 1
              1 1 1
              0 0 0
Scari*periode 4 4 4
              0 0 0
Treat*Scari*periode 1 1 1 0 0 0
                   1 1 1 0 0 0
                   1 1 1 0 0 0
                   1 1 1 0 0 0
                   1 1 1 0 0 0
                   0 0 0/divisor=12 e;

estimate 'T2' intercept 6
treat 0 6 0 0 0
scari 3 3
treat*scari 0 0
           3 3
           0 0
           0 0
           0
periode 2 2 2

```

```

treat*periode 0 0 0
                2 2 2
                0 0 0
                0 0 0
                0 0 0
scari*periode 1 1 1
                1 1 1

treat*scari*periode 0 0 0 0 0 0
                    1 1 1 1 1 1
                    0 0 0 0 0 0
                    0 0 0 0 0 0
                    0 0 0/divisor=6 e;
estimate 'T8 et T4' intercept 12
          Treat 6 6 0 0 0
          Scari 6 6
          periode 4 4 4
          Treat*Scari 3 3
                    3 3
                    0 0
                    0 0
                    0
          Treat*periode 2 2 2
                        2 2 2
                        0 0 0
                        0 0 0
                        0 0 0
          Scari*periode 2 2 2
                        2 2 2
          Treat*Scari*periode 1 1 1 1 1 1
                              1 1 1 1 1 1
                              0 0 0 0 0 0
                              0 0 0 0 0 0
                              0 0 0 / divisor=12 e ;
estimate 'PeriodeLinSca' intercept 8
          Treat 2 2 2 2 0

```

```

      Scari 0 8
      periode 4 0 4
      Treat*Scari 0 2
                  0 2
                  0 2
                  0 2
                  0
      Treat*periode 1 0 1
                  1 0 1
                  1 0 1
                  1 0 1
                  0 0 0
      Scari*periode 0 0 0
                  4 0 4
      Treat*Scari*periode 0 0 0 1 0 1
                        0 0 0 1 0 1
                        0 0 0 1 0 1
                        0 0 0 1 0 1
                        0 0 0 / divisor=8 e cl;
estimate 'PeriodeLinNS' intercept 8
      Treat 2 2 2 2 0
      Scari 8 0
      periode 4 0 4
      Treat*Scari 2 0
                  2 0
                  2 0
                  2 0
                  0
      Treat*periode 1 0 1
                  1 0 1
                  1 0 1
                  1 0 1
                  0 0 0
      Scari*periode 4 0 4
                  0 0 0

```

```

Treat*Scari*periode 1 0 1 0 0 0
                    1 0 1 0 0 0
                    1 0 1 0 0 0
                    1 0 1 0 0 0
                    0 0 0 / divisor=8 e cl;
estimate 'Scari@S*Periode@1' intercept 4
        treat 1 1 1 1 0
        scari 0 4
        treat*scari 0 1
                   0 1
                   0 1
                   0 1
                   0
        periode 4 0 0
        treat*periode 1 0 0
                    1 0 0
                    1 0 0
                    1 0 0
                    0 0 0
        scari*periode 0 0 0
                    4 0 0
        treat*scari*periode 0 0 0 1 0 0
                           0 0 0 1 0 0
                           0 0 0 1 0 0
                           0 0 0 1 0 0
                           0 0 0/divisor=4 e;
estimate 'Scari@S*Periode@2' intercept 4
        treat 1 1 1 1 0
        scari 0 4
        treat*scari 0 1
                   0 1
                   0 1
                   0 1
                   0
        periode 0 4 0

```

```

treat*periode 0 1 0
                0 1 0
                0 1 0
                0 1 0
                0 0 0
scari*periode 0 0 0
                0 4 0
treat*scari*periode 0 0 0 0 1 0
                    0 0 0 0 1 0
                    0 0 0 0 1 0
                    0 0 0 0 1 0
                    0 0 0/divisor=4 e;
estimate 'Scari@S*Periode@3' intercept 4
        treat 1 1 1 1 0
        scari 0 4
        treat*scari 0 1
                    0 1
                    0 1
                    0 1
                    0
        periode 0 0 4
treat*periode 0 0 1
                0 0 1
                0 0 1
                0 0 1
                0 0 0
scari*periode 0 0 0
                0 0 4
treat*scari*periode 0 0 0 0 0 1
                    0 0 0 0 0 1
                    0 0 0 0 0 1
                    0 0 0 0 0 1
                    0 0 0/divisor=4 e;
estimate 'Scari@NPeriode@1' intercept 4
treat 1 1 1 1 0
scari 4 0

```

```
treat*scari 1 0
            1 0
            1 0
            1 0
            0
periode 4 0 0
treat*periode 1 0 0
              1 0 0
              1 0 0
              1 0 0
              0 0 0
scari*periode 4 0 0
              0 0 0
treat*scari*periode 1 0 0 0 0 0
                   1 0 0 0 0 0
                   1 0 0 0 0 0
                   1 0 0 0 0 0
                   0 0 0/divisor=4 e;
estimate 'Scri@NPeriode@2' intercept 4
treat 1 1 1 1 0
scari 4 0
treat*scari 1 0
            1 0
            1 0
            1 0
            0
periode 0 4 0
treat*periode 0 1 0
              0 1 0
              0 1 0
              0 1 0
              0 0 0
scari*periode 0 4 0
              0 0 0
```

```

treat*scari*periode 0 1 0 0 0 0
                    0 1 0 0 0 0
                    0 1 0 0 0 0
                    0 1 0 0 0 0
                    0 0 0/divisor=4 e ;
estimate 'Scari@NPeriode@3' intercept 4
treat 1 1 1 1 0
scari 4 0
treat*scari 1 0
           1 0
           1 0
           1 0
           0
periode 0 0 4
treat*periode 0 0 1
              0 0 1
              0 0 1
              0 0 1
              0 0 1
              0 0 0
scari*periode 0 0 4
              0 0 0
treat*scari*periode 0 0 1 0 0 0
                   0 0 1 0 0 0
                   0 0 1 0 0 0
                   0 0 1 0 0 0
                   0 0 1 0 0 0
                   0 0 0/divisor=4 e ;
estimate 'ScariN*PLin'
          periode -5 0 5
            Treat*periode -1 0 1
              -1 0 1
              -1 0 1
              -1 0 1
              -1 0 1
            Scari*periode -5 0 5
              0 0 0

```

```

Treat*Scari*periode -1 0 1 0 0 0
                    -1 0 1 0 0 0
                    -1 0 1 0 0 0
                    -1 0 1 0 0 0
                    -1 0 1 / divisor=5 e;
estimate 'ScariY*PLin' periode -4 0 4
                    treat*periode -1 0 1
                                    -1 0 1
                                    -1 0 1
                                    -1 0 1
                                    0 0 0
                    Scari*periode 0 0 0
                                    -4 0 4
                    Treat*Scari*periode 0 0 0 -1 0 1
                                            0 0 0 -1 0 1
                                            0 0 0 -1 0 1
                                            0 0 0 -1 0 1
                                            0 0 0 -1 0 1
                                            0 0 0 / divisor=4 e;

estimate 'TE vs éclaircies' Treat -3 -3 -3 -3 12
                    Treat*Scari -3 0
                                    -3 0
                                    -3 0
                                    -3 0
                                    12
                    Treat*periode -1 -1 -1
                                    -1 -1 -1
                                    -1 -1 -1
                                    -1 -1 -1
                                    4 4 4
                    Treat*Scari*periode -1 -1 -1 0 0 0
                                            -1 -1 -1 0 0 0
                                            -1 -1 -1 0 0 0
                                            -1 -1 -1 0 0 0
                                            4 4 4 /divisor=12 e;

```

```

estimate 'T2 vs (T8 et T4)' Treat -6 -6 12 0 0
      Treat*Scari -3 -3
                -3 -3
                6 6
                0 0
                0
      Treat*periode -2 -2 -2
                -2 -2 -2
                4 4 4
                0 0 0
                0 0 0
      Treat*Scari*periode -1 -1 -1 -1 -1 -1
                        -1 -1 -1 -1 -1 -1
                        2 2 2 2 2 2
                        0 0 0 0 0 0
                        0 0 0/ divisor=12 e;

contrast 'TE vs éclaircies' Treat -3 -3 -3 -3 12
      Treat*Scari -3 0
                -3 0
                -3 0
                -3 0
                12
      Treat*periode -1 -1 -1
                -1 -1 -1
                -1 -1 -1
                -1 -1 -1
                4 4 4

      Treat*Scari*periode -1 -1 -1 0 0 0
                        -1 -1 -1 0 0 0
                        -1 -1 -1 0 0 0
                        -1 -1 -1 0 0 0
                        4 4 4 ;

contrast 'CB vs T2 T4 T8' Treat -1 -1 -1 3 0 ;
contrast 'T2 vs T4 T8' Treat -1 -1 2 0 0 ;
contrast 'T4 vs T8' Treat -1 1 0 0 0 ;
contrast '1.5: Scari*PLin'

```

```
Scari*periode -4 0 4
               4 0 -4

Treat*Scari*periode -1 0 1 1 0 -1
                    -1 0 1 1 0 -1
                    -1 0 1 1 0 -1
                    -1 0 1 1 0 -1
                    0 0 0;
Contrast '1.6:Scari*PQua' Scari*periode -4 8 -4
                                       4 -8 4
Treat*Scari*periode -1 2 -1 1 -2 1
                    -1 2 -1 1 -2 1
                    -1 2 -1 1 -2 1
                    -1 2 -1 1 -2 1
                    0 0 0;

),
error=poisson,
link=log,
offset=logoffset

);

*sauvegarde des estimations afin de faire des graphiques dans SPLUS;

data oumarou.lsm;
set _est;
run;
```

```

*****;
* MODÈLE D'ANALYSE DE LA PROBABILITÉ DE PRÉSENCE/ABSENCE DE L'ESPÈCE
*                               P. CORACINUS
*****;

options nodate pageno=1 linesize=65 pagesize=45

*****;
* Création d'une variable de Bernoulli: si un piège situé dans une
* sous-parcelle à une période donnée capture au moins un spécimen de
variable prend la valeur 1 sinon ; * l'espèce P.coracinus alors la
* elle prend la valeur 0
*****;
proc sort data=sasuser.donnees2;
by block treat scari periode;
run;

data sasuser.presabs;
set sasuser.donnees2;
if compte=0 then bino=0;
else bino=1;
by block treat scari periode;
keep trap1 block treat scari periode compte bino ;
run;
* Résultats: Des probabilités certaines causent un problème de convergence pour GLIMMIX;

*****;
* Alternative: Enlever pour tous les blocs, les données correspondant
sous-parcelles non scarifiés des ; * aux parcelles témoins et les
période ;
* parcelles avec 8 trouées mais seulement pour la première
data sasuser.presabs1;
set sasuser.presabs;
if (block='A' or block='B' or block='C' or block='D' ) and (Treat='TE' and Scari='N' and periode=1) then
delete;
if (block='A' or block='B' or block='C' or block='D' ) and (treat='T8' and Scari='N' and periode=1) then
delete; run;

```

```

*****;
*           MACRO GLIMMIX DE SAS           ;
*****;
options nomprint nomlogic nosymbolgen;

%glimmix(data=sasuser.presabs1,
  procopt= ratio covtest,
  stmts=%str(
    class Block Treat Scari periode trap1;
    model bino=Treat|Scari|periode;
    random Block Block*Treat Block*Treat*Scari Block*Treat*Scari*periode;
    repeated periode / subject=trap1(Block Treat Scari) RCORR;
    lsmeans treat;
contrast 'TE vs coupes@NS'
  treat -2 -2 -2 -2 8
    treat*scari -2 0
      -2 0
      -2 0
      -2 0
      8
    treat*periode 0 -1 -1
      0 -1 -1
      0 -1 -1
      0 -1 -1
      4 4
    treat*scari*periode 0 -1 -1 0 0 0
      0 -1 -1 0 0 0
      0 -1 -1 0 0 0
      -1 -1 0 0 0
      4 4;
contrast 'CB vs eclaircis@P=2,3'
  treat -12 4 4 4 0
  treat*scari -6 -6
    2 2
    2 2
    2 2
    0

```

```
treat*periode 0 -6 -6
              0 2 2
              0 2 2
              0 2 2
              0 0
```

```
treat*scari*periode 0 -3 -3 0 -3 -3
                   0 1 1 0 1 1
                   0 1 1 0 1 1
                   1 1 0 1 1
                   0 0;
```

```
contrast 'CB vs eclaircis@S'
```

```
treat -9 3 3 3 0
treat*scari 0 -9
           0 3
           0 3
           0 3
           0
```

```
treat*periode -3 -3 -3
              1 1 1
              1 1 1
              1 1 1
              0 0
```

```
treat*scari*periode 0 0 0 -3 -3 -3
                   0 0 0 1 1 1
                   0 0 0 1 1 1
                   0 0 1 1 1
                   0 0;
```

```
contrast 'T2 vs T4 T8@P=2,3' treat 0 -8 4 4 0
```

```
treat*scari 0 0
           -4 -4
           2 2
           2 2
           0
```

```

treat*periode 0 0 0
               0 -4 -4
               0 2 2
               0 2 2
               0 0
treat*scari*periode 0 0 0 0 0 0
                   0 -2 -2 0 -2 -2
                   0 1 1 0 1 1
                   1 1 0 1 1
                   0 0;
contrast 'T2 vs T4 T8@S' treat 0 -4 2 2 0
treat*scari 0 0
            0 -4
            0 2
            0 2
            0
treat*periode 0 0 0
              0 -2 -2
              0 1 1
              0 1 1
              0 0
treat*scari*periode 0 0 0 0 0 0
                   0 0 0 0 -2 -2
                   0 0 0 0 1 1
                   0 0 0 1 1
                   0 0;
contrast 'T4 vs T8@P=2,3' treat 0 0 -4 4 0
treat*scari 0 0
            0 0
            -2 -2
            2 2
            0
treat*periode 0 0 0
              0 0 0
              0 -2 -2
              0 2 2
              0 0

```

```

treat*scari*periode 0 0 0 0 0 0
                    0 0 0 0 0 0
                    0 -1 -1 0 -1 -1
                    1 1 0 1 1
                    0 0;
contrast 'T4 vs T8@S' treat 0 0 -3 3 0
treat*scari 0 0
            0 0
            0 -3
            0 3
            0
treat*periode 0 0 0
              0 0 0
              -1 -1 -1
              1 1 1
              0 0
treat*scari*periode 0 0 0 0 0 0
                   0 0 0 0 0 0
                   0 0 0 -1 -1 -1
                   0 0 1 1 1
                   0 0;
estimate 'TE@NS&Periode=2,3' intercept 8
treat 0 0 0 0 8
scari 8 0
treat*scari 0 0
            0 0
            0 0
            0 0
            8
periode 0 4 4
treat*periode 0 0 0
              0 0 0
              0 0 0
              0 0 0
              4 4
scari*periode 0 4 4
              0 0 0

```

```

treat*scari*periode 0 0 0 0 0 0
                    0 0 0 0 0 0
                    0 0 0 0 0 0
                    0 0 0 0 0
                    4 4/divisor=8;
estimate 'CB T2 T4 T8@NS' intercept 8
treat 2 2 2 2 0
scari 8 0
treat*scari 2 0
            2 0
            2 0
            2 0
            0
periode 0 4 4

treat*periode 0 1 1
              0 1 1
              0 1 1
              0 1 1
              0 0
scari*periode 0 4 4
              0 0 0
treat*scari*periode 0 1 1 0 0 0
                   0 1 1 0 0 0
                   0 1 1 0 0 0
                   1 1 0 0 0
                   0 0 /divisor=8;
estimate 'T2' intercept 4
treat 0 4 0 0 0
scari 2 2
treat*scari 0 0
            2 2
            0 0
            0 0
            0
periode 0 2 2

```

```

treat*periode 0 0 0
                0 2 2
                0 0 0
                0 0 0
                0 0
scari*periode 0 1 1
                0 1 1
treat*scari*periode 0 0 0 0 0 0
                    0 1 1 0 1 1
                    0 0 0 0 0 0
                    0 0 0 0 0
                    0 0 /divisor=4;
estimate 'T4, T8' intercept 8
treat 0 0 4 4 0
scari 4 4
treat*scari 0 0
              0 0
              2 2
              2 2
              0
periode 0 4 4
treat*periode 0 0 0
              0 0 0
              0 2 2
              0 2 2
              0 0
scari*periode 0 2 2
              0 2 2
treat*scari*periode 0 0 0 0 0 0
                    0 0 0 0 0 0
                    0 1 1 0 1 1
                    1 1 0 1 1
                    0 0 /divisor=8;
estimate 'TE vs coupes@NS' treat -2 -2 -2 -2 8

```

```

treat*scari -2 0
            -2 0
            -2 0
            -2 0
            8
treat*periode 0 -1 -1
              0 -1 -1
              0 -1 -1
              0 -1 -1
              4 4
treat*scari*periode 0 -1 -1 0 0 0
                   0 -1 -1 0 0 0
                   0 -1 -1 0 0 0
                   -1 -1 0 0 0
                   4 4/divisor=8;
estimate '1.T2 vs (T4 ou T8)' treat 0 -8 4 4 0
treat*scari 0 0
           -4 -4
           2 2
           2 2
           0
treat*periode 0 0 0
              0 -4 -4
              0 2 2
              0 2 2
              0 0
treat*scari*periode 0 0 0 0 0 0
                   0 -2 -2 0 -2 -2
                   0 1 1 0 1 1
                   1 1 0 1 1
                   0 0/divisor=8;
estimate '2.T2 vs (T4 ou T8)' treat 0 -4 2 2 0
treat*scari 0 0
           0 -4
           0 2
           0 2
           0

```

```

treat*periode 0 0 0
               0 -2 -2
               0 1 1
               0 1 1
               0 0
treat*scari*periode 0 0 0 0 0 0
                   0 0 0 0 -2 -2
                   0 0 0 0 1 1
                   0 0 0 1 1
                   0 0/divisor=4;
               ),
error=binomial,
link=logit

);

*****;
* MODÈLE FINAL D'ANALYSE DE LA BIODIVERSITÉ DES ESPÈCES: par mois et ;
* avec variable offset ;
*****;

*****;
* MANIPULATION DES DONNÉES ;
*****;

* fusion des données des 4 blocs;
* on garde l'ensemble des 13 semaines et on crée une variable de classification numérique pour
Identification;
data sasuser.carab2;
set sasuser.completC1 sasuser.completD1 sasuser.completA1 sasuser.completB1;
if Identification='Agonum cupripenne (Say)' then Identi=1;
if Identification='Agonum fidele Casey' then Identi=2;
if Identification='Agonum gratiosum (Mannerheim)' then Identi=3;
if Identification='Agonum melanarium Dejean' then Identi=4;
if Identification='Agonum retractum LeConte' then Identi=5;

```

```
if Identification='Bembidion quadrimaculatum oppositum Say' then Identi=6;
if Identification='Bembidion versicolor (LeConte)' then Identi=7;
if Identification='Bembidion wingatei Bland' then Identi=8;
if Identification='Bradycellus nigrinus (Dejean)' then Identi=9;
if Identification='Calathus ingratus Dejean' then Identi=10;
if Identification='Calosoma calidum (Fabricius)' then Identi=11;
if Identification='Chlaenius sericeus sericeus (Forster)' then Identi=12;
if Identification='Clivina fossor (Linné)' then Identi=13;
if Identification='Cymindis cribricollis Dejean' then Identi=14;
if Identification='Disamara arenaria (LeConte)' then Identi=15;
if Identification='Harpalus compar LeConte' then Identi=16;
if Identification='Harpalus rufipes (DeGeer)' then Identi=17;
if Identification='Harpalus solitarius Dejean' then Identi=18;
if Identification='Harpalus somnulentus Dejean' then Identi=19;
if Identification='Leiodes punctostriata Kirby' then Identi=20;
if Identification='Loricera pilicornis pilicornis (Fabricius)' then Identi=21;
if Identification='Notiophilus aeneus (Herbst)' then Identi=22;
if Identification='Olisthopus parmatus (Say)' then Identi=23;
if Identification='Platynus decentis (Say)' then Identi=24;
if Identification='Pseudamara arenaria (LeConte)' then Identi=25;
if Identification='Pterostichus adoxus (Say)' then Identi=26;
if Identification='Pterostichus adstrictus Eschscholtz' then Identi=27;
if Identification='Pterostichus coracinus (Newman)' then Identi=28;
if Identification='Pterostichus melanarius (Illiger)' then Identi=29;
if Identification='Pterostichus pensylvanicus LeConte' then Identi=30;
if Identification='Pterostichus punctatissimus (Randall)' then Identi=31;
if Identification='Pterostichus tenuis (Casey)' then Identi=32;
if Identification='Pterostichus tristis (Dejean)' then Identi=33;
if Identification='Sphaeroderus canad. canadensis Chaudoir' then Identi=34;
if Identification='Sphaeroderus nitidicollis brevoorti LeConte' then Identi=35;
if Identification='Sphaeroderus stenostomus (Weber)' then Identi=36;
if Identification='Syntomus americanus (Dejean)' then Identi=37;
if Identification='Synuchus impunctatus (Say)' then Identi=38;
if Identification='Trechus apicalis Motschulsky' then Identi=39;
if Identification='Trechus rubens (Fabricius)' then Identi=40;

run;
```

```

*décompte du nombre d'espèce différentes;
proc sql;
select Identification, count(Identification) as c
from sasuser.carab2
group by Identification;
quit;

*Résultats: il y a en tout 40 espèces différentes, autrement dit, durant les 13 semaines;
* 40 espèces différentes ont été capturées. Chaque espèce est subdivisée en specimen;

*tri du fichier carab2 pour le calcul du nombre d'espèces par piège et par semaine;
*Le fichier résultats est carab3;
proc sort data=sasuser.carab2 out=sasuser.carab3;
by Block Treatment1 Scarification trap1 semaine identification;
run;

* nombre d'espèces par piège par semaine;
proc means data=sasuser.carab3 n noprint;
by Block Treatment1 Scarification trap1 semaine;
var Identi;
output out=sasuser.carab4 n=nbsp;
run;

*nombre de zéros de la variable nombre d'espèces par piège par semaine;
proc sql;
select count(nbsp) as oo
from sasuser.carab4
where nbsp=0;
quit;

*****;
*Résultats: la variable réponse #nombre d'espèces# contient 741 zéros ;
* sur 1560 soit un pourcentage de 47.55% de zéros. Ce pourcentage est ;
* très élevé. Conséquence: la macro GLIMMIX ne converge pas ;
*Alternative: Faire des regroupements i.e. nbre d'espèces par semaine ;
* pour l'ensemble des pièges d'une sous-parcelle, mais en ne comptant ;
* pas la même espèce plus d'une fois ;
*****;

```

```
*tri du fichier carab2 pour le calcul du nombre d'espèces par semaine ; * et par sous-parcelle. Le fichier
résultat est carab3 ;

proc sort data=sasuser.carab2 out=sasuser.carab3;
by Block Treatment1 Scarification semaine identi;
run;

*décompte des espèces sans redondance par sous-parcelle et par semaine;
proc means data=sasuser.carab3 n noprint;
by Block Treatment1 Scarification semaine;
var Identi;
output out=sasuser.carab5 n=nbespe;
run;

*dans sasuser.carab5, seules les semaines pour lesquelles;
*on trouve au moins un insecte sont considérées;
*On souhaite plutôt considérer toutes les semaines, y compris celles où;
*aucun insecte n'est capturé pour une sous-parcelle donnée;
*Le programme ci-dessous nous permet de faire cela;

*on fait un tri par semaine et ...;
proc sort data=sasuser.carab1;
by Semaine;
run;

*...on considère les 13 semaines...;
data semain;
set sasuser.carab1;
by Semaine;
if first.Semaine;
keep Semaine;
run;

*... et on fait de même pour scarification...;

proc sort data=sasuser.carab1; by Scarification; run;
```

```
data scarifi; set sasuser.carab1; by scarification; if first.scarification;
keep scarification; run;

*... pour block...;
proc sort data=sasuser.carab1; by Block; run;

data blok; set sasuser.carab1; by block; if first.block;
keep block; run;

*...et pour traitement.;
proc sort data=sasuser.carab1; by Treatment1; run;

data treatm; set sasuser.carab1; by treatment1; if first.treatment1;
keep treatment1; run;

*On fait une jonction des 4 fichiers ainsi créés...;
proc sql;
create table btss as
select a.block, b.treatment1, c.scarification, d.Semaine
from blok as a, treatm as b, scarifi as c, semain as d
order by block, treatment1, scarification, Semaine;
quit;

*...puis on fait une jonction avec le fichier sasuser.carab5;
proc sql;
create table sasuser.carab00 as
select a.*, b.*
from btss as a left join sasuser.carab5 as b
on a.block = b.block and
a.treatment1=b.treatment1 and
a.scarification=b.scarification and
a.semaine=b.semaine
order by block, treatment1, scarification, Semaine ; quit;
*on transforme les valeurs manquantes en 0;
data sasuser.carab000; set sasuser.carab00;
if nbespe='.' then nbespe=0; run;
```

```
*le fichier sasuser.varoffset contient les nombres d'espèces non redondantes par;  
*sous=parcelle et par semaine;  
*évidemment une valeur manquante n'est pas comptabilisée;  
proc sql;  
create table sasuser.varoffset as  
select Block, Treatment1, Scarification, semaine, count(distinct identification) as nbespoffset  
from sasuser.carab3  
where identification NE '.'  
group by Block, Treatment1, Scarification, semaine;  
quit;  
  
*le programme ci-dessous a pour but de vérifier que::  
*le nbre total d'espèces est bien supérieur au nbre d'espèces distinctes dans une sous-parcelle par semaine  
;  
data combine;  
merge sasuser.carab000 sasuser.varoffset;  
by Block Treatment1 Scarification Semaine;  
keep block treatment1 scarification semaine nbespe nbespoffset;  
run;  
  
*le nombre d'espèces distinctes par sous-parcelle par semaine est nul s'il est manquant;  
*on crée une variable offset correspondant au nombre de pièges par sous-parcelles;  
data sasuser.combine; set combine;  
if nbespoffset='.' then nbespoffset=0;  
if Treatment1='CB' or Treatment1='T8' or Treatment1='T2' or Treatment1='T4' then nb_piege=3;  
else nb_piege=6;  
run;  
  
* nbre de zéros dans la variable de dénombrement qui est le nbre d'espèces par sous-parcelle par semaine;  
proc sql;  
select count(nbespoffset) as o  
from sasuser.combine  
where nbespoffset=0;  
quit;  
  
*Résultat: sur les 468 observations 131 zéros sont dénombrés soit une pourcentage de 25.2;  
*Il reste à ajuster un modèle linéaire généralisé avec une erreur de Poisson à la variable nbespoffset;  
*Pour cela on utilise la MACRO GLIMMIX de SAS;
```

* Pour compter les espèces par sous-parcelle par groupe de 4 semaines;

```
data sasuser.carab6;  
set sasuser.carab2;  
if semaine=13 then delete;  
mois=int(semaine/4)+1;  
if mod(semaine,4)=0 then mois=mois-1;  
run;
```

```
proc sort data=sasuser.carab6 out=sasuser.carab7;  
by Block Treatment1 Scarification trap1 mois Identification;  
run;
```

```
proc sort data=sasuser.carab7;  
by block treatment1 scarification trap1 mois identification;  
run;
```

```
proc print data=sasuser.carab7 noobs; run;
```

```
proc sql;  
create table sasuser.carab8 as  
select Block, Treatment1, Scarification, mois, count(distinct identification) as nbespoffset  
from sasuser.carab7  
where Identification NE '.'  
group by Block, Treatment1, Scarification, mois;  
quit;
```

*Résultat: Sur les 108 observations il y a un seul 0. Les fréquences ; *nulles ne devraient plus être un souci;

*on crée la variable offset, qui est le nombre de pièges dans une sous-parcelle;

```
data sasuser.carab9; set sasuser.carab8;  
*if block='D' and Treatment1='T2' and Scarification='Y' and mois=2 then nbespoffset=0;  
if Treatment1='CB' or Treatment1='T8' or Treatment1='T2' or Treatment1='T4' then nb_piege=3;  
else nb_piege=6;  
rename Treatment1=Treat Scarification=Scari mois=periode;  
run;
```

```

*logoffset est le logarithme de la variable offset;
data sasuser.carabfinal;
set sasuser.carab9;
logoffset=log(nb_piege);
run;

```

*Il reste à ajuster un GLMM à la variable nbespoffset. Il y a des mesures répétées par mois et aussi une variable offset;

```

*****;
*          MODÉLISATION AVEC LA MACRO GLIMMIX          ;
*****;

%glimmix(data=sasuser.carabfinal,
  procopt= ratio covtest,
  stmts=%str(
    class Block Treat Scari periode;
    model nbespoffset=Treat|Scari|periode;
    random Block Block*Treat Block*Treat*Scari;
    repeated periode / subject=Block*Treat*Scari type=VC RCORR;
lsmeans Treat|Scari|periode / pdiff cl;
estimate 'temoin' intercept 3
          Treat 0 0 0 0 3
          Scari 3 0
          periode 1 1 1
          Treat*Scari 0 0
                    0 0
                    0 0
                    0 0
                    3
          Treat*periode 0 0 0
                       0 0 0
                       0 0 0
                       0 0 0
                       1 1 1

```

```

Scari*periode 1 1 1
              0 0 0
Treat*Scari*periode 0 0 0 0 0 0
                   0 0 0 0 0 0
                   0 0 0 0 0 0
                   0 0 0 0 0 0
                   1 1 1/divisor=3;

estimate 'CB T2 T4 T8@NS' intercept 12
      Treat 3 3 3 3 0
           Scari 12 0
           periode 4 4 4
           Treat*Scari 3 0
                   3 0
                   3 0
                   3 0
                   0

Treat*periode 1 1 1
              1 1 1
              1 1 1
              1 1 1
              0 0 0
Scari*periode 4 4 4
              0 0 0
Treat*Scari*periode 1 1 1 0 0 0
                   1 1 1 0 0 0
                   1 1 1 0 0 0
                   1 1 1 0 0 0
                   0 0 0/divisor=12;

estimate 'T8 et T4' intercept 12
      Treat 0 0 6 6 0
           Scari 6 6
           periode 4 4 4
           Treat*Scari 0 0
                   0 0
                   3 3
                   3 3
                   0

```

```

Treat*periode 0 0 0
               0 0 0
               2 2 2
               2 2 2
               0 0 0
Scari*periode 2 2 2
               2 2 2
Treat*Scari*periode 0 0 0 0 0 0
                   0 0 0 0 0 0
                   1 1 1 1 1 1
                   1 1 1 1 1 1
                   0 0 0 / divisor=12;
estimate 'T2' intercept 6
         treat 0 6 0 0 0
         scari 3 3
         treat*scari 0 0
                   3 3
                   0 0
                   0 0
                   0
periode 2 2 2
treat*periode 0 0 0
               2 2 2
               0 0 0
               0 0 0
               0 0 0
scari*periode 1 1 1
               1 1 1
treat*scari*periode 0 0 0 0 0 0
                   1 1 1 1 1 1
                   0 0 0 0 0 0
                   0 0 0 0 0 0
                   0 0 0/divisor=6;
estimate 'TE vs CB T2 T4 T8@NS' Treat -3 -3 -3 -3 12

```

```

Treat*Scari -3 0
             -3 0
             -3 0
             -3 0
             12
Treat*periode -1 -1 -1
              -1 -1 -1
              -1 -1 -1
              -1 -1 -1
              4 4 4

Treat*Scari*periode -1 -1 -1 0 0 0
                   -1 -1 -1 0 0 0
                   -1 -1 -1 0 0 0
                   -1 -1 -1 0 0 0
                   4 4 4 / divisor=12;
estimate 'T2 vs (T8 et T4)' Treat 0 -12 6 6 0
              Treat*Scari 0 0
                          -6 -6
                          3 3
                          3 3
                          0
Treat*periode 0 0 0
              -4 -4 -4
              2 2 2
              2 2 2
              0 0 0
Treat*Scari*periode 0 0 0 0 0 0
                   -2 -2 -2 -2 -2 -2
                   1 1 1 1 1 1
                   1 1 1 1 1 1
                   0 0 0 / divisor=12;
estimate 'Periode lineaire@NS'
              periode -4 0 4

```

```

treat*periode -1 0 1
              -1 0 1
              -1 0 1
              -1 0 1
              0 0 0
Scari*periode -4 0 4
              0 0 0
Treat*Scari*periode -1 0 1 0 0 0
                   -1 0 1 0 0 0
                   -1 0 1 0 0 0
                   -1 0 1 0 0 0
                   0 0 0/divisor=4;
estimate 'Periode lineaire@S'
        periode -4 0 4
        treat*periode -1 0 1
                   -1 0 1
                   -1 0 1
                   -1 0 1
                   0 0 0
        Scari*periode 0 0 0
                   -4 0 4
        Treat*Scari*periode 0 0 0 -1 0 1
                           0 0 0 -1 0 1
                           0 0 0 -1 0 1
                           0 0 0 -1 0 1
                           0 0 0 -1 0 1
                           0 0 0/divisor=4;
estimate 'Scari@S*Periode@1' intercept 4
        treat 1 1 1 1 0
        scari 0 4
        treat*scari 0 1
                   0 1
                   0 1
                   0 1
                   0
        periode 4 0 0

```

```

treat*periode 1 0 0
                1 0 0
                1 0 0
                1 0 0
                0 0 0
scari*periode  0 0 0
                4 0 0
treat*scari*periode 0 0 0 1 0 0
                    0 0 0 1 0 0
                    0 0 0 1 0 0
                    0 0 0 1 0 0
                    0 0 0/divisor=4;
estimate 'Scari@S*Periode@2' intercept 4
        treat 1 1 1 1 0
        scari  0 4
        treat*scari 0 1
                    0 1
                    0 1
                    0 1
                    0
        periode 0 4 0
treat*periode 0 1 0
                0 1 0
                0 1 0
                0 1 0
                0 0 0
scari*periode 0 0 0
                0 4 0
treat*scari*periode 0 0 0 0 1 0
                    0 0 0 0 1 0
                    0 0 0 0 1 0
                    0 0 0 0 1 0
                    0 0 0/divisor=4;
estimate 'Scari@S*Periode@3' intercept 4
        treat 1 1 1 1 0
        scari  0 4

```

```

      treat*scari 0 1
                0 1
                0 1
                0 1
                0
      periode 0 0 4
      treat*periode 0 0 1
                  0 0 1
                  0 0 1
                  0 0 1
                  0 0 0
      scari*periode 0 0 0
                  0 0 4
      treat*scari*periode 0 0 0 0 0 1
                        0 0 0 0 0 1
                        0 0 0 0 0 1
                        0 0 0 0 0 1
                        0 0 0 0 0 1
                        0 0 0/divisor=4;
estimate 'Scri@NPeriode@1' intercept 4
treat 1 1 1 1 0
scari 4 0
treat*scari 1 0
          1 0
          1 0
          1 0
          0
periode 4 0 0
treat*periode 1 0 0
             1 0 0
             1 0 0
             1 0 0
             0 0 0
scari*periode 4 0 0
             0 0 0

```

```
treat*scari*periode 1 0 0 0 0 0
                    1 0 0 0 0 0
                    1 0 0 0 0 0
                    1 0 0 0 0 0
                    0 0 0/divisor=4;
estimate 'Scri@NPeriode@2' intercept 4
treat 1 1 1 1 0
scari 4 0
treat*scari 1 0
           1 0
           1 0
           1 0
           0
periode 0 4 0
treat*periode 0 1 0
              0 1 0
              0 1 0
              0 1 0
              0 1 0
              0 0 0
scari*periode 0 4 0
              0 0 0
treat*scari*periode 0 1 0 0 0 0
                   0 1 0 0 0 0
                   0 1 0 0 0 0
                   0 1 0 0 0 0
                   0 1 0 0 0 0
                   0 0 0/divisor=4;
estimate 'Scri@NPeriode@3' intercept 4
treat 1 1 1 1 0
scari 4 0
treat*scari 1 0
           1 0
           1 0
           1 0
           0
periode 0 0 4
```

```

treat*periode 0 0 1
               0 0 1
               0 0 1
               0 0 1
               0 0 0
scari*periode 0 0 4
               0 0 0

treat*scari*periode 0 0 1 0 0 0
                    0 0 1 0 0 0
                    0 0 1 0 0 0
                    0 0 1 0 0 0
                    0 0 1 0 0 0
                    0 0 0/divisor=4;
estimate 'Periode 1 2 3@NS' intercept 12
treat 3 3 3 3 0
scari 12 0
treat*scari 3 0
            3 0
            3 0
            3 0
            0
periode 4 4 4
treat*periode 1 1 1
              1 1 1
              1 1 1
              1 1 1
              0 0 0
scari*periode 4 4 4
              0 0 0

treat*scari*periode 1 1 1 0 0 0
                    1 1 1 0 0 0
                    1 1 1 0 0 0
                    1 1 1 0 0 0
                    0 0 0/divisor=12;

estimate 'Periode 1 2 3@S' intercept 12
treat 3 3 3 3 0
scari 0 12

```

```

treat*scari 0 3
            0 3
            0 3
            0 3
            0
periode 4 4 4
treat*periode 1 1 1
              1 1 1
              1 1 1
              1 1 1
              0 0 0
scari*periode 0 0 0
              4 4 4
treat*scari*periode 0 0 0 1 1 1
                    0 0 0 1 1 1
                    0 0 0 1 1 1
                    0 0 0 1 1 1
                    0 0 0/divisor=12;
contrast 'TE vs CB T2 T4 T8@NS' Treat -3 -3 -3 -3 12
                                         Treat*Scari -3 0
                                                         -3 0
                                                         -3 0
                                                         -3 0
                                                         12

Treat*periode -1 -1 -1
              -1 -1 -1
              -1 -1 -1
              -1 -1 -1
              4 4 4
Treat*Scari*periode -1 -1 -1 0 0 0
                    -1 -1 -1 0 0 0
                    -1 -1 -1 0 0 0
                    -1 -1 -1 0 0 0
                    4 4 4 ;

```

```

contrast 'CB vs T2 T4 T8' Treat 3 -1 -1 -1 0 ;
contrast 'T2 vs T4, T8' Treat 0 2 -1 -1 0 ;
contrast 'T4 vs T8' Treat 0 0 -1 1 0 ;
contrast 'Scari*Periode Lineaire' Scari*periode -4 0 4
                                                4 0 -4

Treat*Scari*periode -1 0 1 1 0 -1
                    -1 0 1 1 0 -1
                    -1 0 1 1 0 -1
                    -1 0 1 1 0 -1
                    0 0 0;

```

```

Contrast 'Scari*PQua' Scari*periode -4 8 -4
                                      4 -8 4
Treat*Scari*periode -1 2 -1 1 -2 1
                    -1 2 -1 1 -2 1
                    -1 2 -1 1 -2 1
                    -1 2 -1 1 -2 1
                    0 0 0;

```

```

),
error=poisson,
link=log,
offset=logoffset

```

```
);
```

*on sauvegarde les estimations des moyennes dans le fichier oumar.lsm, certaines serviront à faire des graphiques;

* avec SPLUS, pour l'étude de l'effet linéaire deu temps selon la scarification;

```
data oumar.lsm;
```

```
set _est;
```

```
run;
```