

Analyse de l'effet du paysage sur les insectes ravageurs de la luzerne porte-graine

Clément CHEVALEYRE et Corentin BARBU

Coordinateur FNAMS : Benjamin COUSSY

Février 2026

1. Introduction

L'UMR agronomie travaille depuis plus près de 20 ans sur l'impact du paysage sur les bioagresseurs des cultures (Delaune et al. 2021¹, Rusch et al., 2013²). Dans ce rapport, nous appliquons les méthodes appliquées par Delaune et al. (2021) aux données rassemblées sur les insectes ravageurs de la luzerne porte-graine par la FNAMS ces dernières années. Nous étendons ces méthodes pour quantifier les corrélations avec les périmètres et non seulement avec les surfaces des différentes occupations du sol. Nous cherchons ainsi à estimer l'impact de la composition et de la configuration des paysages sur la présence des différents ravageurs.

2. Matériel et méthode

Pour quantifier les corrélations entre éléments paysagers et ravageurs de la luzerne porte-graine, nous utilisons un modèle linéaire de type LASSO avec validation croisée et bootstrap. Nous modélisons les observations d'abondance des ravageurs à l'échelle de la commune ou de la parcelle, en fonction des données paysagères et de variables de contrôle. Les variables de contrôle sont ajoutées aux modèles pour prendre en compte des paramètres qui pourraient avoir un rôle sur la distribution des bioagresseurs tout en étant corrélées avec les variables paysagères.

En plus des variables de contrôles explicites telles que les variables météorologiques nous prenons en compte des effets des territoires ou des années. Ces effets peuvent servir de variable approchant la distribution d'autres variables (proxy), telles que les pratiques agricoles plus ou moins présentes suivant les territoires ou le climat, qui pourraient être corrélées avec les variables paysagères. La prise en compte de ces variables de contrôle peut éviter d'attribuer aux paysages des effets liés à d'autres facteurs corrélés aux paysages mais peut inversement expliquer une part de la variance effectivement induite par les différences de paysages. L'introduction ou pas des différentes variables de contrôle représente donc des hypothèses alternatives sur la part des paysages dans l'explication de la variance. La présentation des résultats des différents modèles alternatifs permet d'éprouver la robustesse des estimations aux hypothèses de structure de la variabilité.

¹ Delaune, T., Ouattara, M.S., Ballot, R., Sausse, C., Felix, I., Maupas, F., Chen, M., Morison, M., Makowski, D. and Barbu, C. (2021), Landscape drivers of pests and pathogens abundance in arable crops. *Ecography*, 44: 1429-1442. <https://doi.org/10.1111/ecog.05433>

² Rusch, A., Valantin-Morison, M., Sarthou, J. P., & Roger-Estrade, J. (2013). Effect of crop management and landscape context on insect pest populations and crop damage. *Agriculture, Ecosystems & Environment*, 166, 118-125.

2.1. Les jeux de données et variables explicatives

Jeux de données sur le suivi des principaux ravageurs de la luzerne

Les données portent sur sept insectes ravageurs : l'apion (*Holotrichapion pisi*), les punaises mirides (*Lygus* sp. et *Adelphocoris lineolatus*), le tychius (*Tychius aureolus*), la tordeuse (*Cydia medicaginis*), le sitone (*Sitona* sp.), le négril (*Colaspidema barbarum*) et le phytonome (*Hypera postica*). Pour chaque point d'observation, le suivi des ravageurs est réalisé sur une année donnée. Selon les jeux de données, la localisation des observations varie :

- Période 2006–2019 : données localisées à l'échelle communale.
- Périodes 2019–2022 et 2021–2025 : données localisées à l'échelle parcellaire.

L'agglomération des bases permet d'obtenir un jeu de données initial de 1 555 parcelles, après filtrage des doublons (en privilégiant systématiquement la localisation parcellaire).

Surface en luzerne porte-graine

Données disponibles par département, couvrant la période 2012–2025 (source : SEMAE).

Données fournies par l'INRAE

- **Données météorologiques** : issues de la base SAFRAN (Météo France), avec une maille de 8 km.
- **Données paysagères** : issues de la BD TOPO® (IGN, édition 2017).
- **Données du registre parcellaire graphique (RPG)** : déclarations PAC des agriculteurs (2015–2024).
- **Délimitation des régions agricoles** : issues de Agreste³

Il existe deux grandes catégories de variables explicatives :

- **Variables paysagères.** Ces variables décrivent la structure du paysage autour des parcelles. À l'exception des surfaces en luzerne porte-graine à l'échelle départementale, nous utilisons différents rayons autour de la position géographique de la parcelle ou du centroïde de la commune pour calculer :
 - **Depuis le RPG (Registre Parcellaire Graphique) (à partir de 2015) :**
 - Surfaces de **luzerne (tous types)** pour l'année en cours et l'année précédente (soit 2 variables).
 - Surfaces de **prairies** pour l'année en cours.
 - **Depuis la BDTPO (version 2017),** les surfaces et périmètres de :
 - **Haies**
 - **Forêts**, pour les haies et les forêts, les distinctions faites par l'IGN étant parfois douteuses, nous avons réassigné les surfaces étant à la fois longues et petites à des haies.
 - **Bois traités** : surfaces en culture pérenne (vigne, arboriculture, peupleraies).
 - **Landes** : terrains avec un taux de recouvrement de végétation naturelle ou semi-naturelle (ligneuse ou non) supérieur à 10 %, et un taux de recouvrement relatif des ligneux bas (arbustes et sous-arbrisseaux) supérieur ou égal à 25 %. Le taux de recouvrement des arbres ne doit pas dépasser 10 %. Les landes ne présentent pas de traces d'intervention humaine régulière de type agricole (clôtures, parcellaire visible). On y trouve des zones de parcours extensifs, des friches, des terrains vacants, des garrigues et des maquis non boisés.

³ <https://agreste.agriculture.gouv.fr/agreste-web/methodon/Z.1!/searchurl/listeTypeMethodon/>

- **Surface en luzerne porte-graine départementale**

Les surfaces en luzerne porte-graine ne sont intégrées dans le modèle que pour l'année en cours. En effet, à une échelle aussi large, cette variable reflète davantage un contexte environnemental global, qui peut raisonnablement être considéré comme stable d'une année sur l'autre. L'inclusion des surfaces pour les deux années (N et N-1) introduirait une forte colinéarité entre ces variables. Dans ce cas, le modèle LASSO, incapable de différencier leur poids explicatif respectif, sélectionnerait arbitrairement l'une des deux variables (généralement la première rencontrée dans le jeu de données), ce qui biaiserait l'interprétation des résultats.

Les effets paysagers (à l'exception de la luzerne porte-graine) sont inclus dans les modèles pour différents buffers simultanément, le LASSO permettant de ne sélectionner que les distances les plus pertinentes pour expliquer les abondances de bio-agresseurs :

- **À l'échelle communale** : rayons de 5 000 m et 10 000 m.
- **À l'échelle parcellaire** : rayons supplémentaires de 200 m et 1 000 m.

– **Variables environnementales non paysagères.** Ces variables captent la variabilité liée au contexte local (météo, climat, sol, etc.) à différentes échelles spatiales et temporelles.

- **Ancienne région administrative** : identifie les parcelles situées dans une même région (effet spatial à gros grain).
- **Régions agricoles** : identifie les parcelles situées dans une même région agricole (effet spatial à grain plus fin).
- **Ancienne région administrative – année = effet météo indirect** : identifie les parcelles situées dans une même région et une même année.
- **Météo = effet météo direct** : pour chaque variable météorologique (TM = température moyenne, TN = température minimale, TX = température maximale, RR = cumul des précipitations et RG = rayonnement), nous utilisons les données des 12 mois précédant le mois de la date moyenne des observations des ravageurs (calculée précédemment).

Pour trouver la date la plus représentative d'abondances élevées, la date moyenne des observations est pondérée par l'abondance à chaque date :

$$dateMoyObs = \frac{\sum_{i=1}^n jour_i * w_i}{\sum_{i=1}^n w_i}$$

Où :

- $jour_i$ = valeur numérique de la date du jour i
- w_i = abondance du ravageur associé au jour i
- Les valeurs sont moyennées par mois (ou sommées pour les précipitations), sauf pour les périodes automnales et hivernales, qui sont agrégées :
 - **ONDJ** : octobre, novembre, décembre de l'année $n-1$ et janvier de l'année n .
 - **FMA** : février, mars et avril de l'année n . Ces variables apportent généralement des informations de premier ordre dans les modèles, expliquant une part importante de la variance.

2.2. Procédure d'analyse

Pour quantifier les ravageurs, nous utilisons le **logarithme de l'observation maximale plus 1**, pour un point d'observation donné. Ce choix se justifie par l'objectif agronomique final : minimiser les pics de présence des ravageurs, qui sont directement liés aux risques de dégâts sur les cultures et aux pertes de rendement. Le passage au logarithme permet de réduire l'influence des valeurs extrêmes sur la régression. L'augmentation des effectifs d'un point permet d'inclure les observations nulles et donc de conserver l'information biologique associée aux absences.

Nous **testons des modèles** incluant différents groupes de variables, chacune correspondant à des **hypothèses sous-jacentes distinctes en fonction des variables de contrôle intégrées**. L'objectif est de vérifier que les résultats de ces modèles convergent vers les mêmes conclusions, renforçant ainsi la robustesse de nos analyses. Toutes les combinaisons d'hypothèses ne pouvant pas être explorées, nous définissons initialement un modèle « central » (*init*) et nous étudions ensuite les hypothèses introduisant des variations par rapport à ce modèle central.

En parallèle, une analyse similaire est menée sur le **logarithme de la médiane des observations**, qui reflète davantage le niveau de base des populations de ravageurs dans le temps. Les résultats de ces modèles ne sont présentés qu'en annexe 5.2 mais les divergences notables d'avec le modèle central sont mentionnées dans le texte principal.

A. Préparation des données pour le modèle central

Les variables explicatives sont **normalisées**, par bioagresseur, sur l'ensemble des observations les différentes années (2006-2025) afin de permettre la comparaison directe des coefficients entre les variables en terme d'importance relative pour expliquer la variance des abondances de ravageurs.

Le modèle inclut l'**ensemble des variables paysagères** (luzerne porte-graine spécifique, luzerne tous types de conduite pour les années N et N-1, prairies, landes, forêts, haies et bois traités), ainsi que des variables de contrôle : les **effets territoire** (régions-administratives et régions-agricoles), ajoutés pour éviter que des différences liées à des pratiques agricoles ou des contextes pédoclimatiques locaux ne soient confondues avec les effets paysagers, et les **effets interannuels** (région-année), permettant de prendre en compte indirectement les variations météorologiques.

Les analyses sont d'abord menées à l'**échelle communale**. Cette approche permet de conserver les points d'observations pour lesquelles nous n'avons pas de localisation exacte. En raison de la disponibilité limitée des données cultures issues du RPG (2015–2024), nous avons exclu les parcelles des périodes 2006–2014 et 2025. Ainsi, le jeu de données final est restreint à environ **500–600 points d'observation selon les ravageurs**.

B. Définition des variantes du modèle central

À partir du modèle central, nous **retirons ou substituons** certaines variables pour tester différentes hypothèses ou prise en compte de données d'observations différentes. Un **modèle spécifique à l'échelle parcellaire** (*parcelle_init*) est notamment développé pour analyser les **effets du paysage à des distances plus fines** (200 m et 1 000 m), en reprenant les variables du modèle de référence. L'ensemble de ces modèles est détaillé dans le Tableau 1.

[Tableau 1 : Structures des modèles testés et variables explicatives associées](#)

Nom du modèle	Description du modèle et hypothèse sous-jacentes
<i>init</i>	Inclut l'ensemble des variables paysagères et les variables de contrôle environnementales (effet territoire et année).

<i>area_only</i>	Dans ce modèle on teste uniquement les effets du paysage liées aux surfaces (retrait des effets périmètres du paysage). Cela dans l'hypothèse que la surface des composantes du territoire aient un impact plus fort sur le ravageur que les interfaces.
<i>perimeter_only</i>	Dans ce modèle on teste uniquement les effets du paysage liés aux périmètres (retrait des effets surfaces du paysage). Cela dans l'hypothèse que les interfaces de bordure aient un impact plus important que la surface des composantes du paysage sur le ravageur.
<i>sans_regAgri</i>	Dans ce modèle on retire le contrôle de la région-agricole dans l'hypothèse ou celle-ci capterait des potentiels effets liés au paysage. Certaines corrélations avec le paysage peuvent apparaître ou se renforcer.
<i>sans_regionAn</i>	Dans ce modèle on retire les effets d'année (notamment liée à la météo) pour voir si la présence du ravageur est impacté par un effet année ou de météo. Si tel est le cas, la capacité prédictive (R^2) du modèle pourrait plus ou moins se dégrader.
<i>meteo</i>	Dans ce modèle les effets années sont substitués par de l'information météo. Cela permet de voir si les résultats sont robustes à la prise en compte explicite de la météo.
<i>paysage_seul</i>	Dans ce modèle, on conserve uniquement les variables du paysage (retrait de l'ensemble des effets territoires, météo ou année) afin d'identifier ce que pourrait être l'impact du paysage dans l'hypothèse forte ou l'abondance du ravageur serait exclusivement liée à celui-ci.
<i>parcelle_init</i>	Correspond au modèle central (<i>init</i>), mais appliqué au jeu de données restreint ayant la localisation précise des parcelles. Cela permet d'également prendre en compte les effets du paysage à courte distance : 200 m et 1 000 m.

C. Procédure LASSO et validation croisée

Ce type de régression linéaire minimise la somme des carrés des écarts en lui ajoutant un terme de pénalisation proportionnel d'une part à la valeur des coefficients estimés et d'autre part à un paramètre de pénalisation. Cette pénalisation permet de sélectionner les variables utilisées et d'ajuster les valeurs des coefficients des variables conservées. Les variables non-sélectionnées ont un coefficient estimé exactement à 0, les valeurs sélectionnées ont un coefficient strictement négatif ou positif. Lorsque le coefficient est positif, les variables sont positivement corrélées à l'abondance des ravageurs. Lorsque le coefficient est négatif, les variables sont négativement corrélées à l'abondance de ravageur.

Pour estimer la valeur de pénalisation (= la force de sélection de variables) pertinente, une validation croisée par dixième est utilisée : pour chaque valeur du paramètre de pénalisation lambda, le modèle est ajusté 10 fois en excluant à chaque fois un dixième des données. Ce dixième est ensuite utilisé pour calculer la somme des carrés des écarts associée à une valeur donnée du paramètre de pénalisation. Ici, nous choisissons conformément à l'habitude dans la littérature une valeur « conservatrice » de lambda au sens statistique du terme. Nous retenons la pénalisation la plus forte pour laquelle la somme des carrés des écarts ne dépasse pas la valeur minimale obtenue pour les différentes valeurs testées plus un écart-type de la somme des carrés des écarts, tel qu'estimé théoriquement. Pour le package R glmnet que nous utilisons cela revient à utiliser cv.glmnet avec l'option lambda égal à « lambda.1se ». Cette méthode limite le risque de sur-ajustement, ce qui est crucial pour éviter des interprétations erronées.

D. Robustesse des estimées LASSO

Étant donné que le **LASSO intègre une composante aléatoire dans la sélection des variables, il est essentiel d'ajuster un même modèle plusieurs fois**. Pour cela nous réalisons un bootstrap de **1000 répétitions**. Cela permet de visualiser l'ensemble des paramètres pouvant être sélectionnés de manière aléatoire, ainsi que la variabilité de l'estimation des coefficients.

Pour chaque variable, la **valeur médiane du coefficient** sur l'ensemble des 1000 répétitions bootstrap est représentée, accompagnée de son **intervalle de confiance à 90 %** (laissant 5% de chaque côté) ; si l'intervalle de confiance ne coupe pas zéro, la p-value de ce coefficient est alors inférieure à 5% (la probabilité d'obtenir un coefficient aussi différent de 0 et du même signe par chance est inférieure à 5%). Le taux de sélection de la variable est égal à 1- (p-value) si le coefficient est de même signe à chaque fois que la variable est sélectionnée.

A titre d'exemple, la figure 1 présente les résultats obtenus pour les 1000 répétitions du modèle central (*init*) appliqué à la tordeuse de la luzerne. La luzerne porte-graine (Luzerne PG) et les forêts jusqu'à 10km (Forêts - 10km) sont sélectionnées par le modèle pour plus de 50 % des répétitions (médiane = point bleu détaché de 0) et elles sont aussi significativement différentes de 0 par bootstrap avec un seuil de 5% (intervalle détaché de 0). Les autres variables, quant à elles, sont sélectionnées par le LASSO dans moins de 50 % des répétitions et sont donc *a fortiori* non significatives à 5%.

L'utilité du bootstrap sur 1000 répétitions est particulièrement évidente ici, car il permet de mettre en lumière des intervalles de confiance parfois très larges, reflétant une incertitude importante dans l'estimation des effets. Cette approche renforce la fiabilité des conclusions en identifiant les variables les plus stables et significatives.

Une limitation de cette représentation réside dans le fait que, pour chaque composante, la distance du buffer pour laquelle la corrélation est examinée est analysée indépendamment des autres. Ainsi, si deux distances sont sélectionnées par le modèle et présentent le même signe, les effets globaux de l'élément paysager peuvent être fortement sous-estimés. À l'inverse, si deux distances sont sélectionnées avec des signes opposés, les effets de l'élément paysager dans son ensemble pourraient être surévalués (puisqu'ils peuvent s'annuler).

Sélection des variables paysagères par bootstrap LASSO - Tordeuse

Variables classées par fréquence de sélection (1000 répétitions)
Seuil de sélection : 5%

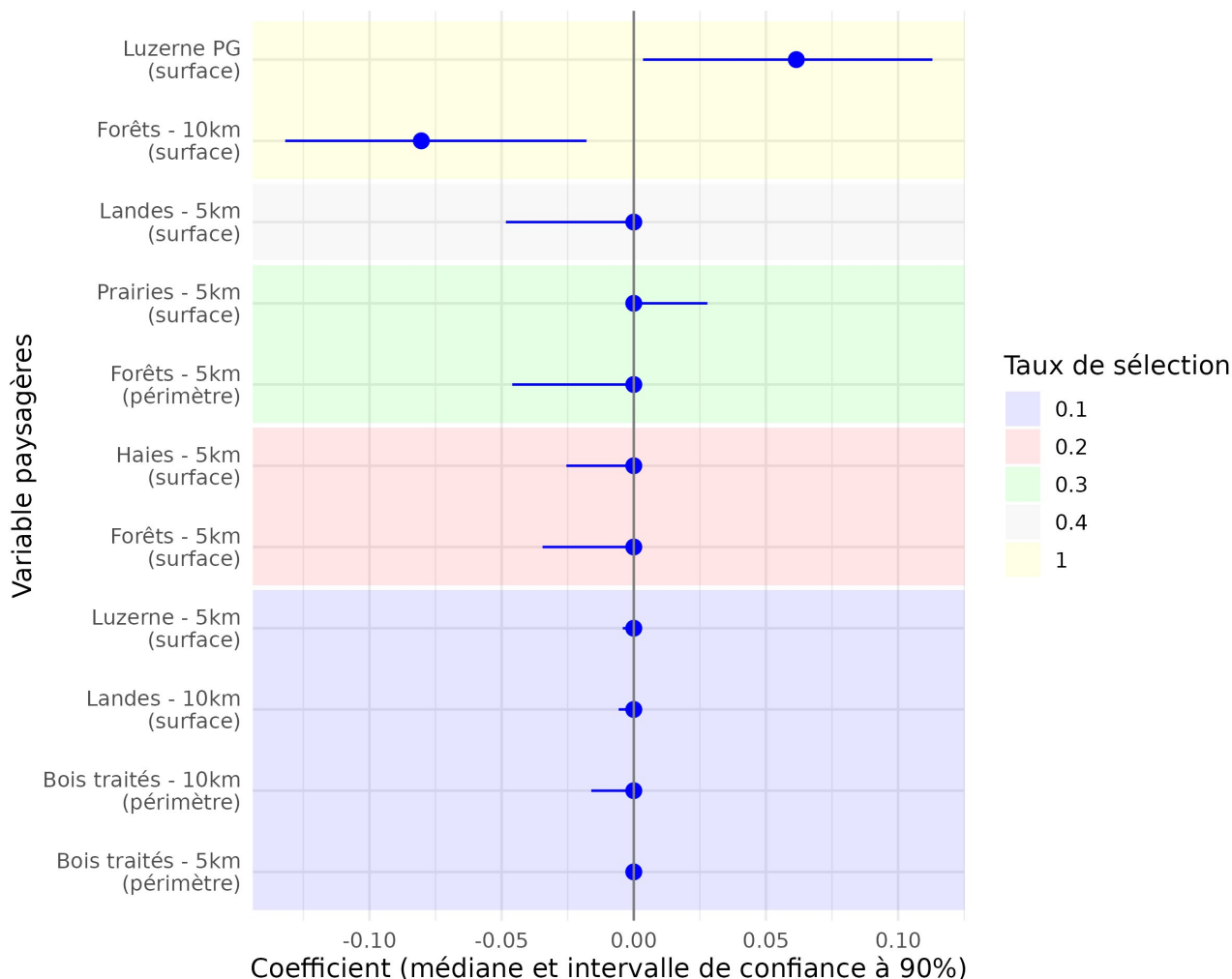


Figure 1. Sélection et stabilité des variables paysagères explicatives du logarithme des observations maximales de la Tordeuse pour le modèle *init* sur 1000 répétitions de bootstrap. Les variables sont classées par fréquence de sélection et seules les variables sélectionnées dans au moins 5% des répétitions sont représentées par un souci de lisibilité. Le rond bleu correspond au coefficient médian et le trait représente l'intervalle de confiance à 90% (quantile 5% et 95%). Les distances indiquées (en kilomètres) correspondent au rayon du buffer autour de chaque centroïde de la commune des parcelles pour le calcul des métriques paysagères.

E. Représentation conjointe de l'ensemble des modèles alternatifs pour un ravageur

Pour étudier l'effet global de chaque composante paysagère, nous avons **agrégé les coefficients des variables apparentées** en les sommant, **indépendamment de leur nature** (périmètre, surface) ou du rayon considéré. La sommation est permise par la normalisation préalable des données explicatives qui met toutes les données dans la même échelle en termes de variabilité observée dans les données.

Cette somme est calculée **individuellement pour chacune des 1000 répétitions**, ce qui permet d'analyser ensuite la **distribution du coefficient agrégé**. Cette méthode offre une vision globale de

l'effet d'une composante paysagère, tout en **captant les tendances dominantes** et en intégrant d'éventuelles **corrélations opposées** au sein d'une même catégorie.

Les résultats pour l'ensemble des modèles alternatifs sont présentés sur une **figure unique par ravageur** pour les différents modèles testés (figures 2 à 8).

Chaque graphique de synthèse comprend :

1. **La distribution des coefficients sommés** pour chaque composante du paysage, selon les différentes variantes de modèle.
2. **Le R² médian** reflétant la part de variance expliquée par l'ensemble du modèle. Lors des comparaisons entre variantes du modèle, il ne faut cependant pas oublier que certaines variantes sont ajustés sur des observations différentes (échelle, années sélectionnées) ce qui change la capacité à expliquer les données.
3. **La distance moyenne d'effet** de chaque composante paysagère. Les différentes distances sélectionnées sont pondérées par la probabilité de sélection. Deux valeurs sont indiquées, la première faisant la moyenne pour l'ensemble des modèles à l'échelle communale et la deuxième pour le modèle à l'échelle parcellaire.

Dans les explications, nous ne commentons pas les résultats de tous les modèles mais nous nous concentrons sur ceux qui semblent avoir une signification statistique.

Les corrélations estimées pour les variables météorologiques ne sont généralement pas significatives. Nous ne les présentons pas ici d'une part car elles sont peut souvent significatives et d'autres part car la robustesse des estimations n'a pas été vérifiée. Si certaines corrélations sont significatives, nous les mentionnons cependant à la fin des résultats pour chaque bioagresseur.

3. Résultats et discussion par bioagresseur

3.1. Apion (figure 2)

Absence de lien avec les surfaces de luzernes

Les modèles ajustés pour l'apion à l'échelle communale ne révèlent aucune relation entre l'abondance de ce ravageur et les surfaces de luzerne, qu'il s'agisse de la luzerne porte-graine ou de la luzerne en général. Ces résultats indiquent que la dynamique de l'apion n'est pas directement liée à la disponibilité de la culture dans le paysage, du moins aux échelles spatiales considérées dans cette étude. Cela pourrait s'expliquer soit parce que la luzerne cultivée représente une part assez faible des plantes exploitées par l'apion soit parce que le développement de l'apion avec la luzerne se fait en parallèle avec celui d'auxiliaires qui le régulent. Le modèle à l'échelle parcellaire a cependant une tendance nettement positive bien que non significative par bootstrap à 95% de la luzerne l'année N, à assez courte distance (2171m). Bien que l'on ne puisse pas être sûr qu'il y ait un effet, si il y en a un, il est probablement positif et l'absence d'effet des luzernes dans les autres modèles pourrait être liée à l'imprécision de la localisation.

Effet négatif des prairies

Une tendance à la corrélation négative est observée entre l'abondance de l'apion et la présence de prairies. La médiane pour tous les modèles est inférieure à zéro bien que pour la majorité des modèles, l'intervalle de confiance du bootstrap à 95% inclut ou est très proche de zéro. Les trois modèles pour lesquels l'effet des prairies est clairement négatif sont des modèles qui retirent tout ou partie des variables de contrôle territoriales. Les territoires où il y a plus de prairies ont donc moins d'apions. L'effet protecteur des prairies contre les apions pourrait en partie résulter d'une confusion d'effet avec d'autres caractéristiques des territoires prairiaux non pris en compte dans le modèle (météo, élevage, ...).

Tendance positive des landes et ligneux

Pour les autres composantes paysagères (landes, forêts, haies, bois traités), la tendance semble plutôt corrélée positivement avec l'abondance du ravageur, bien que cela reste incertain. Lorsqu'on cherche à expliquer le logarithme de la médiane des observations (Annexe 5.2.1), le modèle central devient plus significativement différent de zéro pour les landes. Même si les résultats restent imprécis, l'effet des landes apparaît plus associé aux périmètres qu'aux surfaces, ce qui suggère que des landes fragmentées auraient plus d'impact que des landes d'un seul tenant de surface égale. Un effet positif des bois traités et des forêts à courte distance dans le modèle à l'échelle parcellaire est possible mais très peu certain.

Influence potentielle des variables météorologiques

La lecture préliminaire des coefficients des variables météorologiques (modèle météo), non présentée ici, révèle une corrélation positive et significative entre les précipitations du mois de juillet et l'abondance de l'apion. La robustesse de cet effet à des modifications du modèle n'a pas été évaluée. Si cet effet était confirmé il pourrait expliquer en partie les pics d'abondance maximale observés en juillet.

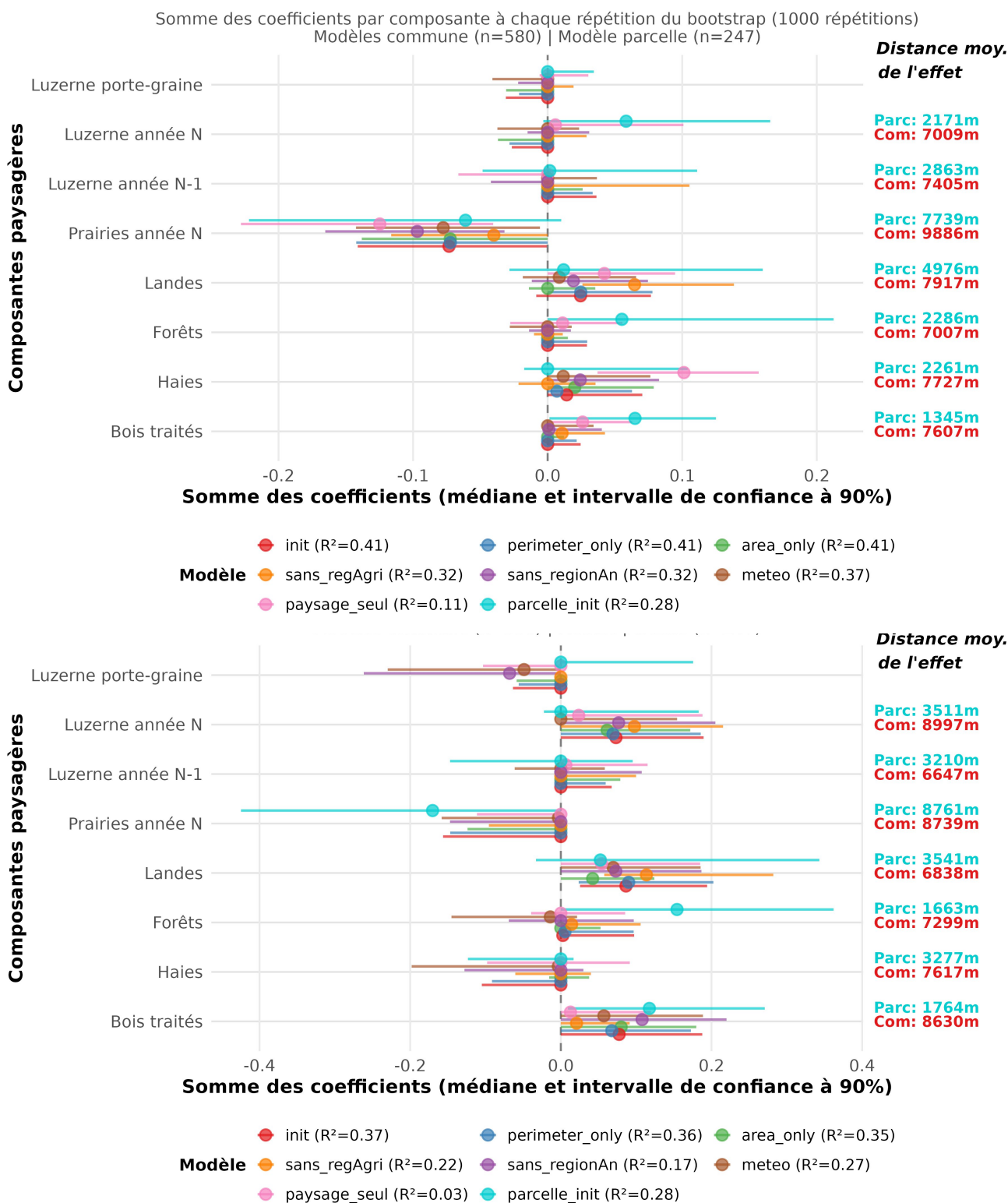


Figure 2. Contribution relative des composantes paysagères dans l'explication du logarithme des observations maximales (haut) ou médiane (bas) d'abondance de l'Apion ($\log(\text{abondance} + 1)$) pour différentes variantes de modèles LASSO (1000 répétitions bootstrap)

Pour chaque composante paysagère, cette figure présente la distribution des sommes des coefficients de toutes les variables associées, calculées individuellement à chaque répétition. Le R² médian est indiqué pour chaque variante du modèle afin d'évaluer sa capacité explicative. Les distances moyennes (en mètres), affichées à droite des graphiques, précisent l'échelle spatiale à laquelle l'effet de chaque composante est le plus prononcé.

3.2. Punaises mirides (figure 3)



Punaise miride adulte

Absence de lien clair avec les surfaces en luzerne

Aucune corrélation significative n'est observée entre les surfaces en luzerne et les abondances maximales de punaise miride, bien qu'une tendance positive apparaisse avec la luzerne porte-graine. Cette tendance est renforcée pour les abondances médianes : les corrélations avec les abondances de luzerne dans le paysage l'année précédente (N-1) à 6 km deviennent très significatives et robustes aux spécifications de modèle (Annexe 5.2.2), ce qui suggère un effet d'inoculum paysager, éventuellement limité par des régulations naturelles pouvant s'exercer plus facilement à courtes distances (0,5 km, modèle parcellaire). Ces régulations pourraient limiter les très fortes abondances ce qui expliquerait le manque de corrélation avec les maximum d'abondance. Les corrélations globalement faibles et incohérentes (médiane versus max) entre abondances de punaises mirides et surfaces de luzerne pourraient s'expliquer par le caractère peu spécialisé de la punaise miride, capable d'exploiter la luzerne comme ressource parmi d'autres, plus abondantes dans le paysage.

Rôle limité des prairies et des landes

Les prairies et les landes ne présentent pas de lien clair avec l'abondance maximale des punaises mirides, leurs coefficients restant proches de zéro dans la majorité des variantes de modèles. Bien qu'une légère tendance négative soit observable pour les prairies, ces milieux ne semblent pas jouer un rôle déterminant dans la dynamique des populations des punaises mirides.

Effet négatif des forêts

Les forêts montrent une tendance à la corrélation négative avec le ravageur, tendance renforcée pour les abondances médianes et notamment pour les données à la parcelle. Cet effet négatif, plutôt associé aux surfaces, suggère un effet de tampon des forêts, avec une dilution paysagère des surfaces de plantes hôtes.

Tendances positives des haies et bois traités

Les haies présentent une tendance à des corrélations positives, bien que ces résultats restent peu robustes et incertains. Cette tendance est légèrement plus marquée pour les vergers, vignes et peupleraies (bois traités) ce qui pourrait refléter une certaine disponibilité de plantes hôtes dans ces espaces.

Influence potentielle des variables météorologiques

La lecture préliminaire des variables météorologiques suggère un effet potentiel des précipitations : une corrélation négative en mai et positive en juin et juillet, la robustesse de ces corrélations aux spécifications de modèle n'a cependant pas été évaluée.

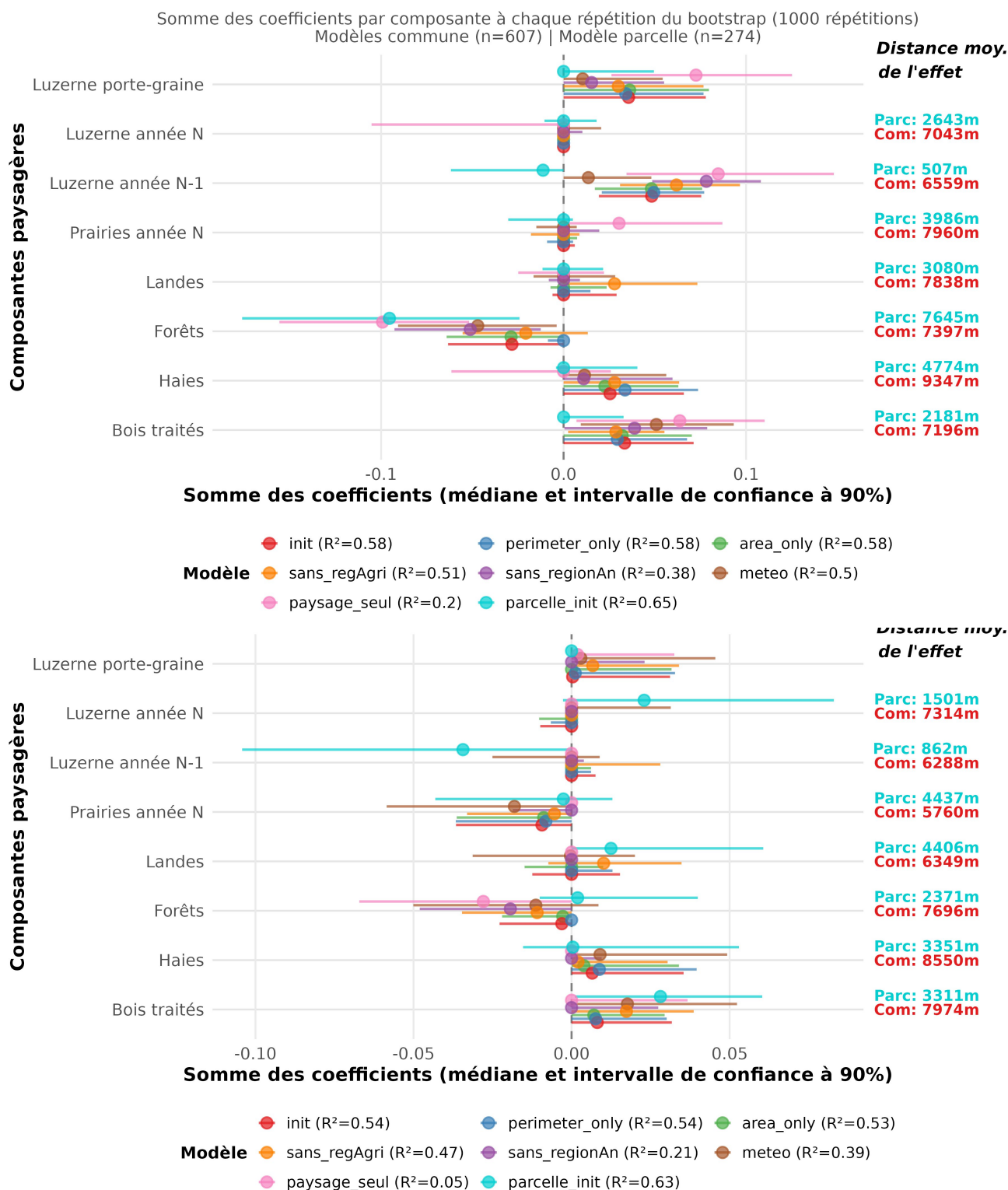


Figure 3. Contribution relative des composantes paysagères dans l'explication du logarithme des observations maximales (haut) et médianes (bas) d'abondance des punaises mirides
Voir détails sur la légende de la figure 1

3.3. Sitone (figure 4)

Effet négatif et robuste des forêts

La corrélation négative entre abondances maximales de sitones et forêts (sans que l'on puisse indiquer si c'est plus le périmètre ou la surface qui joue) est assez forte en plus d'être significative et cohérente. Elle n'est plus significative mais reste négative qu'à l'échelle parcellaire, ce qui pourrait s'expliquer par le nombre nettement plus faible d'observations à cette échelle. Plusieurs mécanismes pourraient expliquer cet effet, comme un rôle de barrière physique ou une association avec des contextes paysagers moins favorables au ravageur (réduction des surfaces cultivées hôtes ou augmentation de la régulation biologique). Cependant, sans données spécifiques sur les auxiliaires ou les cultures hôtes alternatives, ces hypothèses restent à confirmer. Contrairement aux forêts, les haies n'affichent qu'une corrélation négative faible et non significative avec les abondances de sitone.

Absence de lien clair avec les surfaces de luzerne

Les modèles à l'échelle communale ne révèlent aucun lien entre les surfaces de luzerne et l'abondance du sitone. Un élargissement et une précision géographique améliorée du jeu de données pourrait faire apparaître des tendances plus nettes (tendance positive pour le modèle parcellaire) mais probablement assez faible.

Tendance protectrice des prairies

Les prairies montrent une tendance à la corrélation négative avec l'abondance du sitone, bien que cet effet ne soit pas significatif dans la majorité des modèles. Un effet significatif est observé dans le modèle excluant les régions agricoles, ce qui pourrait indiquer une confusion entre l'effet des prairies et celui d'autres variables environnementales.

Faibles tendance positive des landes et bois traités

Les landes et les bois traités ne présentent aucun lien avec les pics maximaux d'abondance de la sitone. En revanche, pour les valeurs médianes (Annexe 5.2.3), une tendance à la corrélation positive est observée. À l'échelle parcellaire, cette tendance s'inverse, avec une corrélation négative plus marquée et à plus petite échelle. A ce stade les tailles d'effets et significativités sont trop faibles pour y attacher de l'importance.

Perspectives pour une meilleure compréhension

Globalement, ces résultats suggèrent que la dynamique du sitone dépend moins de la présence de luzerne que de la structure générale du paysage. Les éléments ligneux, en particulier les forêts, semblent jouer un rôle clé dans la réduction des abondances. Étant donné le caractère assez généraliste du sitone, des analyses complémentaires intégrant la distribution des légumineuses cultivées dans le paysage pourrait permettre d'identifier plus précisément les facteurs favorisant sa présence.

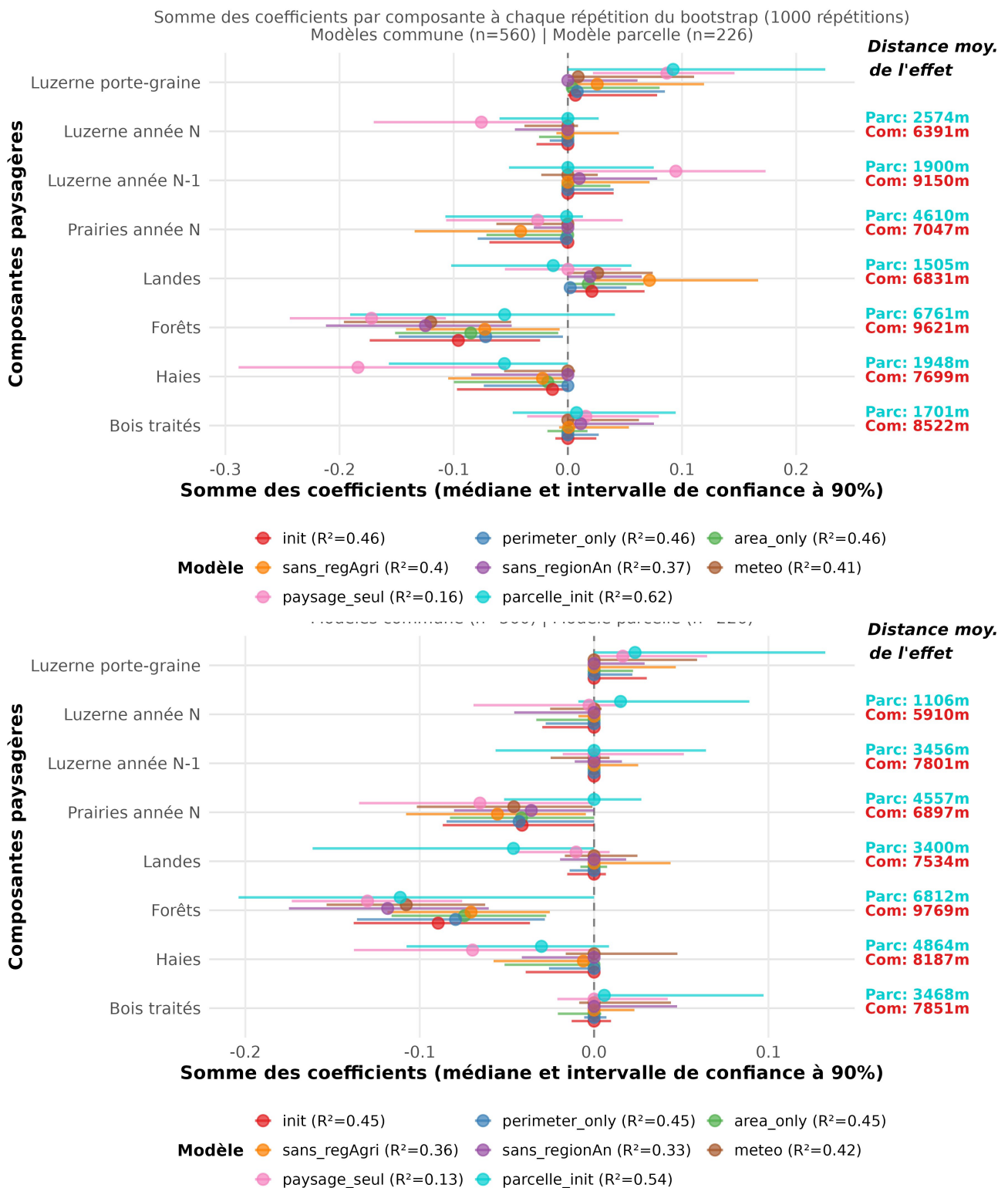


Figure 4. Contribution relative des composantes paysagères dans l'explication du logarithme des observations maximales (haut) et médianes (bas) d'abondance de sitones
Voir détails sur la légende de la figure 1

3.4. Tordeuse (figure 5)



Tordeuse adulte

Les données disponibles pour la tordeuse sont plus limitées que pour les autres ravageurs étudiés, car l'échantillonnage se fait de manière différente (via un piège à phéromone), ce qui peut restreindre la capacité des modèles à détecter certains effets paysagers. Malgré cela, plusieurs résultats apparaissent robustes et cohérents avec la biologie de ce ravageur.

Effet positif marqué des surfaces de luzerne porte-graine

Une corrélation positive marquée est observée entre l'abondance de la tordeuse et les surfaces de luzerne porte-graine à l'échelle départementale. Cet effet est stable dans les différentes variantes de modèles et est significatif pour la majorité des modèles d'abondance maximale (mais pas pour l'abondance médiane, annexe 5.2.4). Il n'y a en revanche aucun lien avec les surfaces de luzerne issues du RPG : ce résultat est en accord avec le cycle biologique de la tordeuse, dont les larves se développent spécifiquement dans les gousses de la luzerne (donc spécifique à la culture porte-graine). Cela traduit un effet de la disponibilité de la ressource reproductive qui pourrait être exploré à une échelle plus fine que départementale si les surfaces de luzernes porte-graines à l'échelle parcellaire étaient rendues disponibles.

Effet négatif significatif des forêts et des haies

Les forêts et les haies sont associées à une corrélation négative significative avec l'abondance de la tordeuse. Cet effet est observé à grande échelle (8,5 km pour les modèles communaux) mais n'est pas significatif pour le modèle parcellaire. Cet effet est clairement lié aux surfaces de forêts plus qu'aux périmètres, suggérant un effet de masse dans le paysage. La tordeuse étant un papillon, une explication plausible serait que ces éléments paysagers abritent des prédateurs naturels (oiseaux, chauves-souris), limitant ainsi sa présence, mais les forêts pourraient aussi simplement avoir un effet tampon avec de grandes surfaces impropres à la reproduction car exemptes de luzerne.

Les haies et les bois traités ont aussi une corrélation négative avec les abondances mais de manière moins importante et moins significative. Cependant, à l'échelle parcellaire, une légère tendance à la corrélation positive avec les bois traités à faibles distances est observée, tendance un peu plus nette pour les abondances médianes que les abondances maximales.

Tendance positive des prairies

Les prairies montrent une tendance à la corrélation positive avec l'abondance de la tordeuse, surtout pour les abondances médianes (Annexe 5.2.4). Cela pourrait être lié à la présence de légumineuses capables de monter en graines dans ces milieux, maintenant une population de base dans les environs, bien que cette hypothèse ne puisse être confirmée avec les données actuelles.

Variables météorologiques et effets régionaux peu informatifs

La capacité explicative (R^2) des modèles n'est pas dégradée lorsque les variables météorologiques ou les effets régionaux agricoles sont retirés. Cela indique que, pour la tordeuse, les variations interannuelles ou climatiques apportent peu d'information supplémentaire par rapport aux variables paysagères et au contexte environnemental.

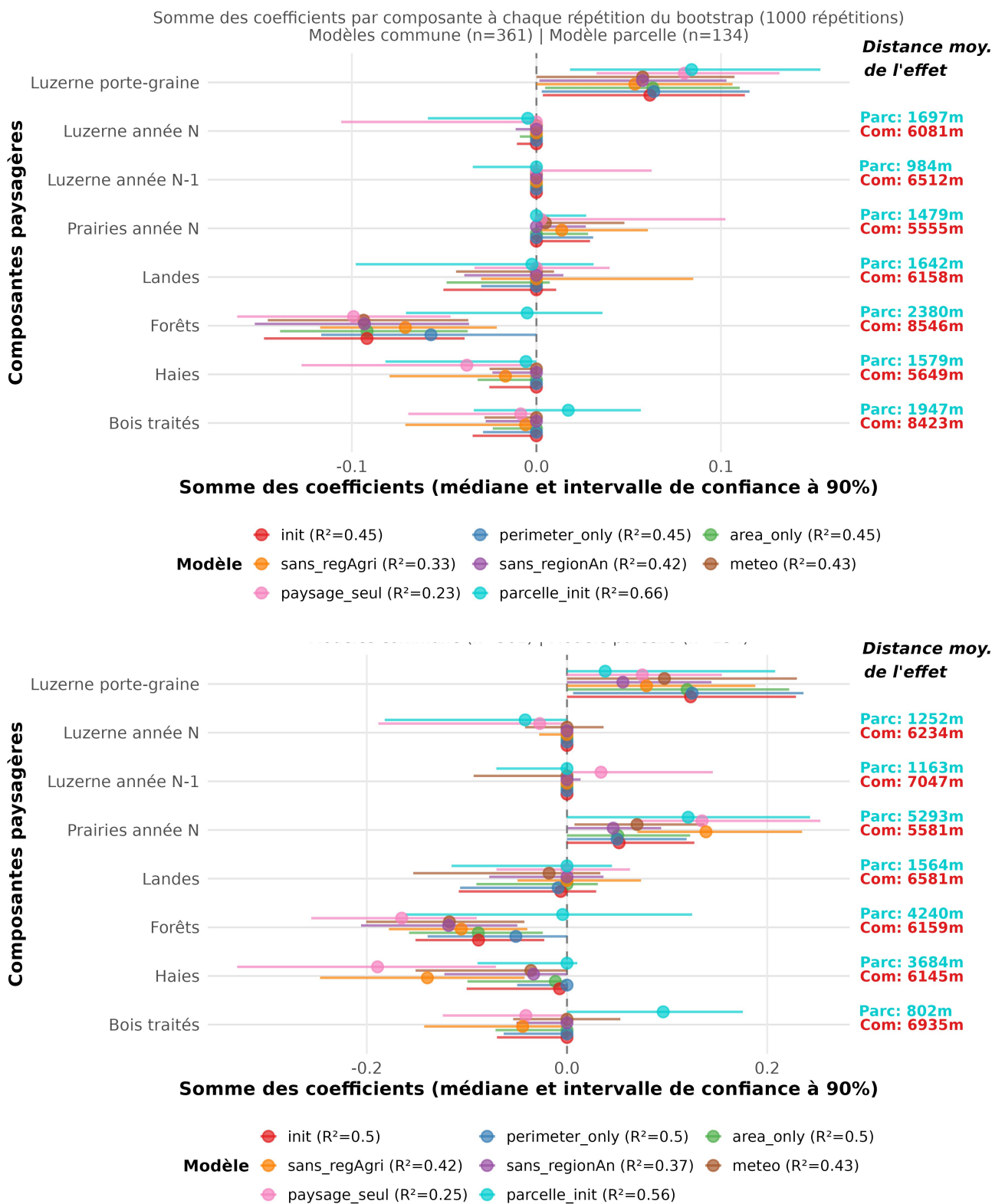


Figure 5. Contribution relative des composantes paysagères dans l'explication du logarithme des observations maximales (haut) et médianes (bas) d'abondance de tordeuse
Voir détails sur la légende de la figure 1

3.5. Tychius (figure 6)



Tychius adulte

On note pour le R^2 une valeur explicative importante du modèle central (0.71), même le modèle paysage seul a un R^2 relativement élevé (0.39) ce qui suggère que les effets paysage pourraient être importants.

Corrélation positive forte avec les surfaces de luzerne porte-graine

Une corrélation positive marquée est observée entre l'abondance de Tychius et les surfaces de luzerne porte-graine à large échelle. Cet effet, fort, significatif et stable entre les différentes variantes de modèles, est cohérent avec le cycle biologique de ce ravageur, dont le développement larvaire s'effectue dans les gousses de luzerne. On note cependant que l'effet de la luzerne porte-graine à l'échelle départementale est moins net pour le modèle parcellaire, cela pourrait s'expliquer si il y avait des variations des surfaces de luzerne porte-graine dans certaines régions échantillonnées entre 2015 et 2023. Dans ce cas, la sélection des données avec information parcellaire qui se focalise sur les années 2019-2023 pourrait favoriser la sélection d'un effet « région » par rapport à un effet « abondance départementale de la luzerne ». A l'inverse, la prise en compte de l'ensemble de l'échantillon permettrait de corrélérer l'augmentation de la luzerne porte-graine au sein d'une même région avec les abondances de tychius.

Aucun lien robuste n'est détecté avec les autres types de conduites de luzerne.

Effet négatif peu robuste des prairies et des forêts

Les prairies présentent une corrélation négative à 5km pour le modèle à l'échelle de la parcelle. En revanche, cet effet n'est sensible à l'échelle communale que pour le modèle sans aucune covariable. Les forêts ont tendance à être négativement corrélées à l'abondance de tychius, mais cet effet reste non significatif et incertain.

Absence de lien avec les landes et les haies, mais tendance positive pour les bois traités

Aucun lien n'est mis en évidence pour les landes et les haies, dont les coefficients restent nuls ou proches de zéro dans la majorité des modèles. En revanche, les bois traités semblent positivement corrélés à l'abondance de tychius, bien que cet effet reste non significatif, sans que l'on puisse avancer d'hypothèse quant à ce résultat.

Faible impact des variations interannuelles et climatiques

Comme pour la tordeuse, le retrait des effets région-année (sans_regionaAn) n'entraîne qu'une faible dégradation de la capacité explicative des modèles ($R^2 = 0.66$). Cela indique que les variations interannuelles apportent peu d'information supplémentaire par rapport aux variables paysagères et au contexte environnemental.

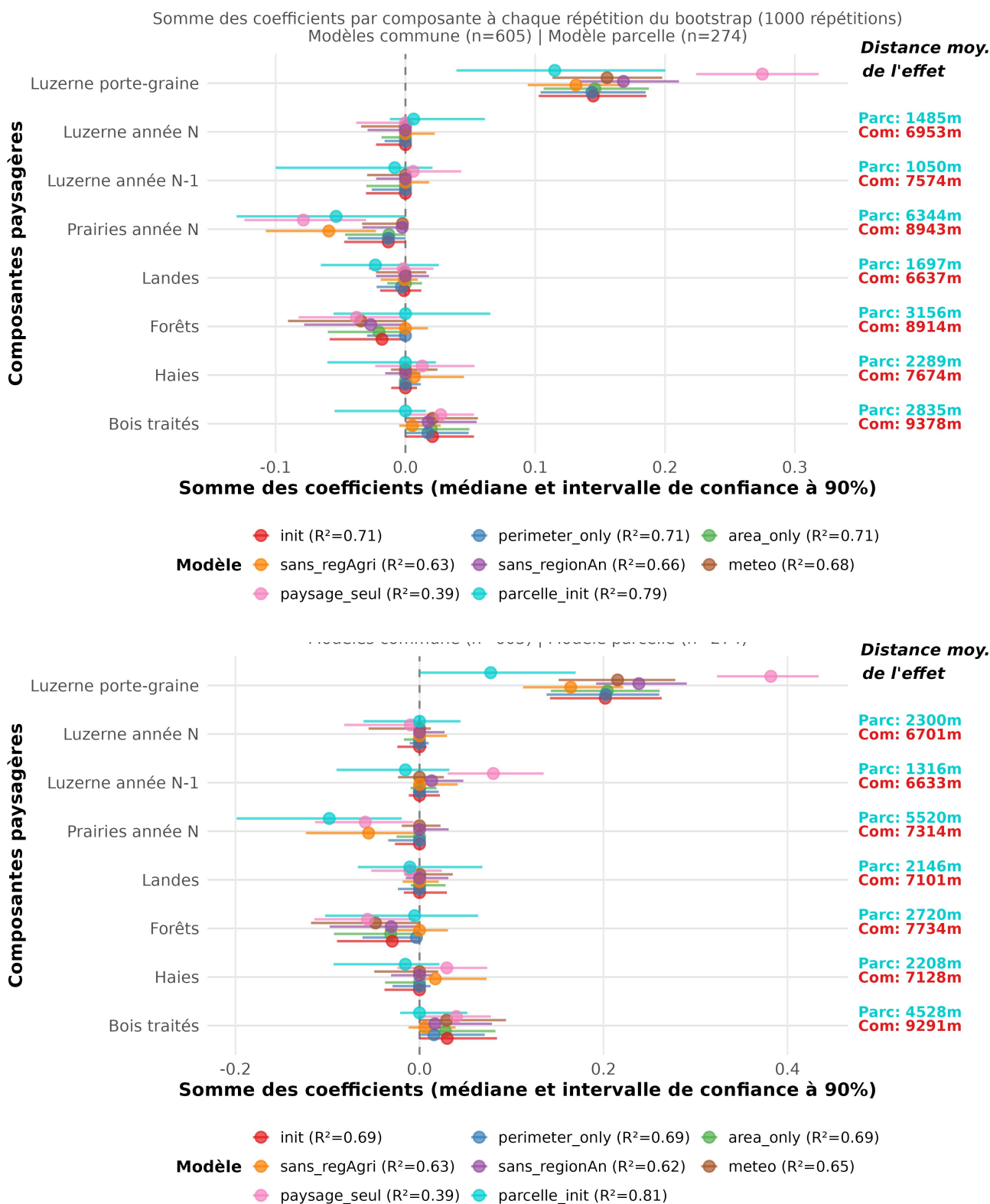


Figure 6. Contribution relative des composantes paysagères dans l'explication du logarithme des observations maximales (haut) et médianes (bas) d'abondance de tychius
Voir détails sur la légende de la figure 1

3.6. Larves de négril (figure 7)



Larve de négril

Le négril présente une répartition géographique principalement limitée au sud de la France, ce qui se traduit par de nombreuses observations nulles dans les régions plus au nord. Pour limiter un biais lié à cette absence structurelle, les analyses ont été restreintes aux régions administratives où le ravageur est effectivement présent.

Seule la modélisation du maximum observé permet d'obtenir une capacité explicative quelconque des modèles (Figure 7 et annexe 5.2.6 - figure 1). Cela peut s'expliquer par la composition du jeu de données qui contient presque exclusivement des observations médianes à 0 (Annexe 5.2.6 - figure 2). L'utilisation du maximum permet d'avoir un jeu de données avec plus d'observations positives même si elles restent très minoritaires (7%).

Tendances limitées pour les composantes paysagères

Aucune composante paysagère ne montre de liens significatifs et robustes avec l'abondance du négril. Cependant, certaines tendances se dégagent :

- Les surfaces de luzerne semblent positivement corrélées avec la présence du ravageur. Étant donné que le négril est essentiellement inféodé à la luzerne, mais peut aussi se développer sur le trèfle, la présence de ces cultures dans le paysage pourrait favoriser son abondance.
- Les prairies présentent une tendance à la corrélation négative, suggérant un effet protecteur potentiel. Elles pourraient, par exemple, maintenir une population d'auxiliaires sans que l'effet soit très fort ni très significatif à ce jour.
- Aucune tendance claire ne se dégage pour les landes, les forêts et les bois traités.
- Les haies sont l'élément paysager montrant la corrélation positive la plus marquée, avec des médianes de coefficients différentes de zéro. Cet effet devient même significatif et très fort pour le modèle à l'échelle parcellaire, modèle qui a aussi un pouvoir explicatif élevé : le coefficient médian est très élevé, le plus grand de toute l'étude et deux fois celui reliant luzerne porte-graine et tychius qui est le deuxième. A l'échelle communale cependant, seul le modèle n'incluant que les variables paysagères devient significatif mais avec une dégradation importante du R^2 rapport aux autres modèles. Cette contradiction est d'autant plus surprenante que les distances de corrélation sélectionnées sont assez élevées (6 à 9km). Malgré cette contradiction, les valeurs élevées trouvées pour le modèle à l'échelle parcellaire mériteraient une étude plus poussée, idéalement en prenant en compte des données bien géoréférencées et plus diversifiées spatialement. Cela permettrait aussi de mieux évaluer les effets de la luzerne qui pourraient encore être importants (barres d'erreurs très longues à droite).

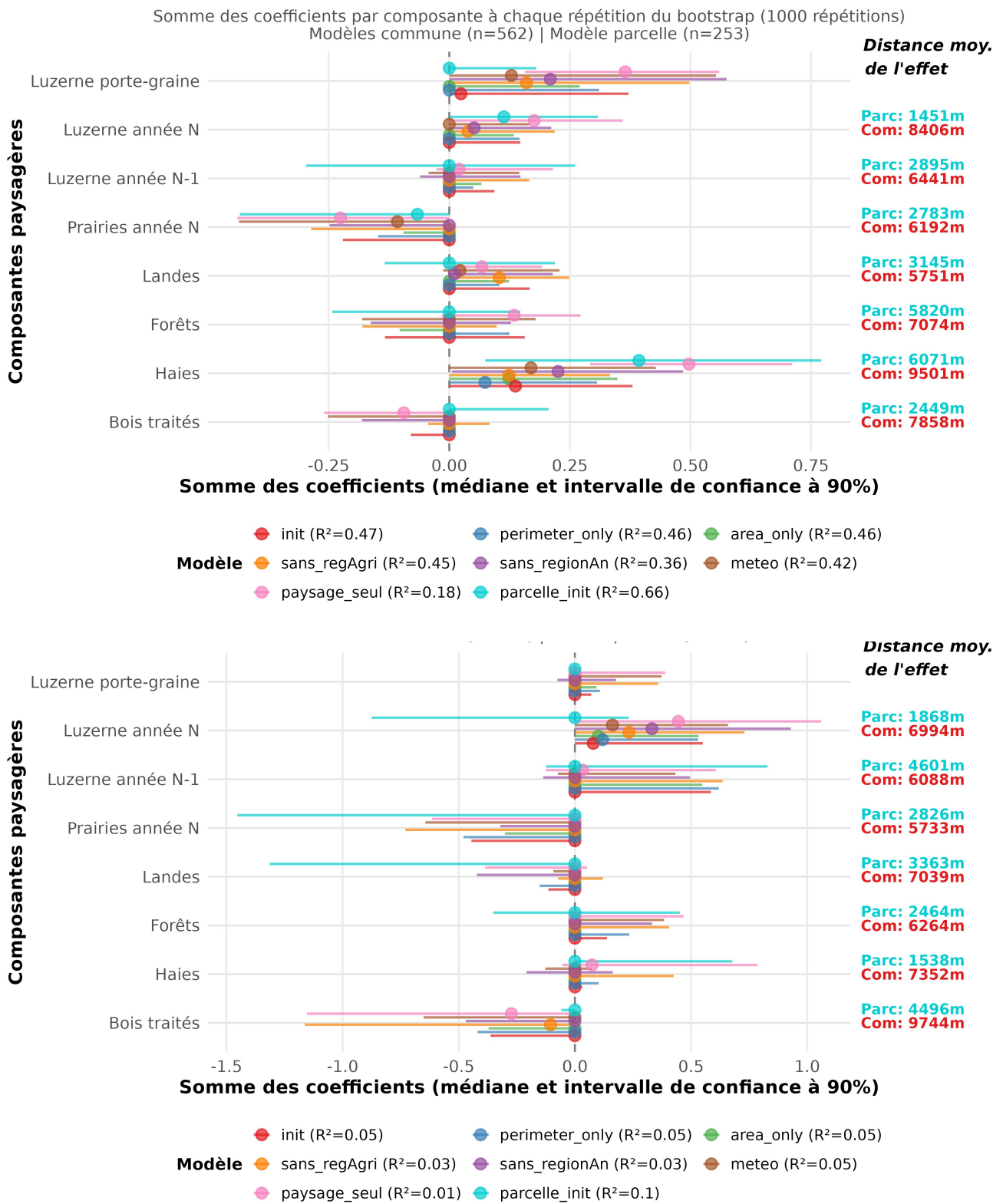


Figure 7. Contribution relative des composantes paysagères dans l'explication du logarithme des observations maximales (haut) et médianes (bas) d'abondance de Négril
Voir détails sur la légende de la figure 1

3.7. Larves de phytonome (figure 8)



Larve de phytonome

Absence de lien significatif à l'échelle parcellaire

Le modèle à l'échelle parcellaire ne révèle aucun lien clair avec les composantes paysagères, avec des médianes des coefficients presque systématiquement nulles ou proches de zéro et des incertitudes faibles autour de zéro.

Corrélation négative faible avec les prairies à l'échelle communale

À l'échelle communale, nous observons un lien potentiel avec les surfaces en luzerne (années N et N-1), bien que cette corrélation négative reste non significative. La corrélation négative avec les prairies apparaît significative pour l'ensemble des modèles des maximums d'abondances mais la tendance opposée est observée pour les médianes (Annexe 5.2.7). Cette apparente contradiction s'explique par le fait que les modèles n'expliquent pas les mêmes aspects de la dynamique du phytonome : les prairies pourraient maintenir une population de base du ravageur (corrélation positive avec le log de la médiane), tout en favorisant la présence d'auxiliaires qui limitent les pics d'abondance (corrélation négative avec le log des observations maximales). L'absence d'effet à l'échelle parcellaire pourrait simplement être liée au plus faible nombre d'observations, surtout avec les tailles d'effets estimées qui sont assez faibles et donc facilement « pénalisées » par le LASSO.

Absence d'effet des autres composantes paysagères

Pour les autres éléments du paysage (landes, forêts, haies et bois traités), aucun lien notable n'est observé avec les abondances maximales du ravageur. On note cependant une tendance à la corrélation positive avec les haies, non significative à 5% mais robuste entre modèle.

Faible capacité prédictive des modèles

Les capacités prédictives des modèles pour les larves de phytonome restent bien inférieures à celles obtenues pour les autres ravageurs. Ces résultats suggèrent qu'il manque des éléments clés pour caractériser la présence et l'abondance de ce ravageur dans son environnement. Il est possible que les pratiques culturales à la parcelle, bords de parcelles ou autres éléments non pris en compte ici aient un impact déterminant sur la présence de larves de phytonome.

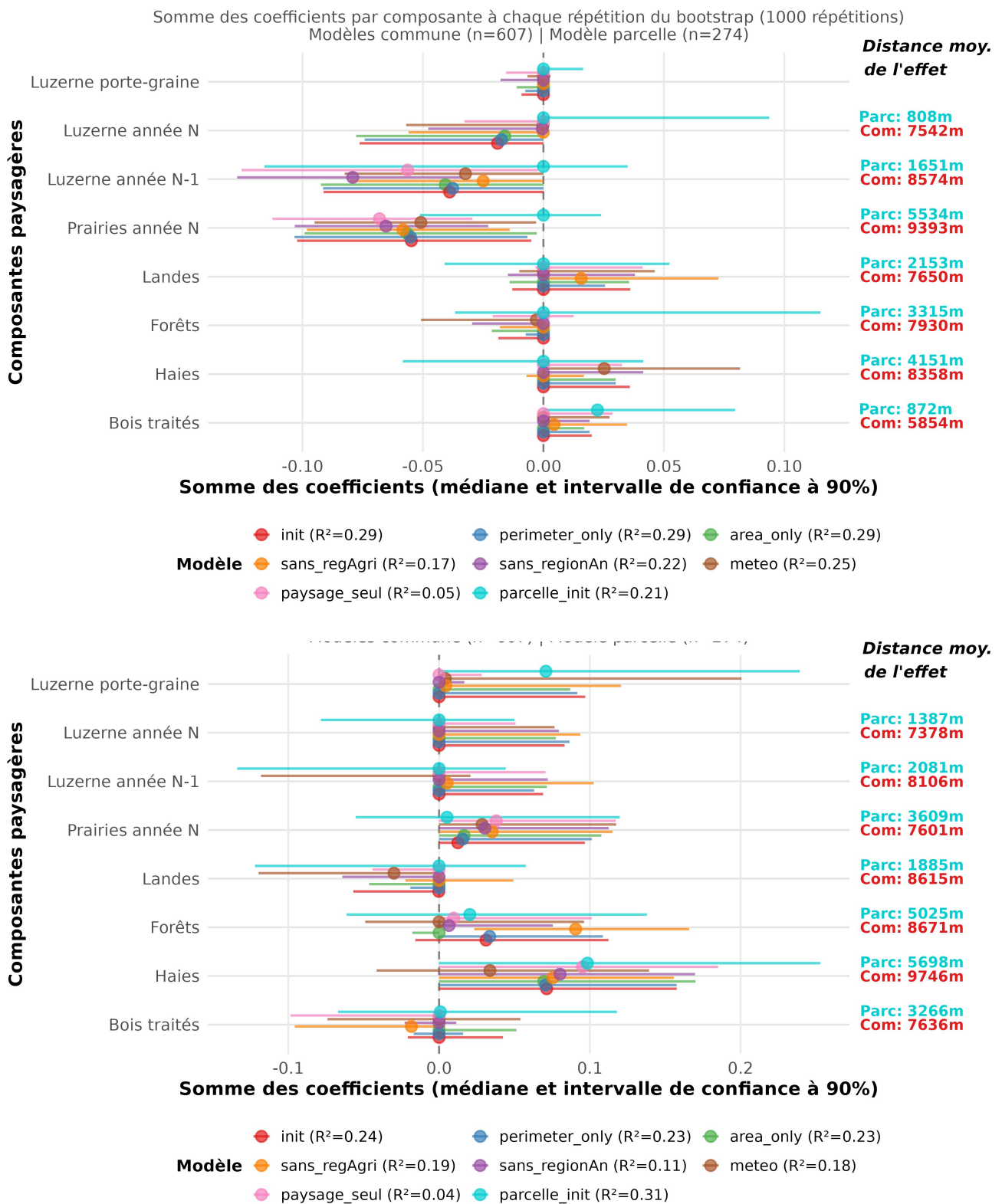


Figure 8. Contribution relative des composantes paysagères dans l'explication du logarithme des observations maximales (haut) et médianes (bas) d'abondance de phytonome
Voir détails sur la légende de la figure 1

3.8. Poids des différents types de variables

Pour représenter le poids des différents types de variables dans la modélisation, nous représentons l'écart-type des prédictions pour l'ensemble du jeu de donnée en utilisant seulement un type de variables (régions agricoles agrégées, paysage ...). Plus l'écart-type généré par un groupe de variables est grand plus ce groupe de variables a d'impact dans le modèle ajusté. Nous indiquons aussi l'écart-type dans les données observées (Data) et l'écart-type généré par le modèle complet (Complet) pour donner l'ordre de grandeur de la variabilité à expliquer (écart-type de « Data ») et de la variabilité expliquée par l'ensemble du modèle (écart-type du modèle « complet »). Nous montrons ces valeurs premièrement pour les modèles « Init » puis « météo » sur les logs de maximums d'abondance, le premier utilisant les coefficients « région-année » comme proxy de la météo une année donnée et l'autre utilisant les variables météo (Figure 9).

On retrouve le très faible poids des variables paysage (Landscape) pour les punaises, surtout comparé aux variables de « région-année » qui représente la variabilité annuelle par région. Au contraire, le tychius et la tordeuse ont un poids du paysage assez conséquent, plus important que celui de la variabilité inter-annuelle (Region-année ou Meteo).

Les écart-types générés par les variables météo sont généralement plus faibles que ceux générés par les régions-années, ce qui n'est pas très surprenant puisque les variables région-année, en tant qu'effets fixes, peuvent accommoder de la variabilité non décrite par les variables météo prises en compte.

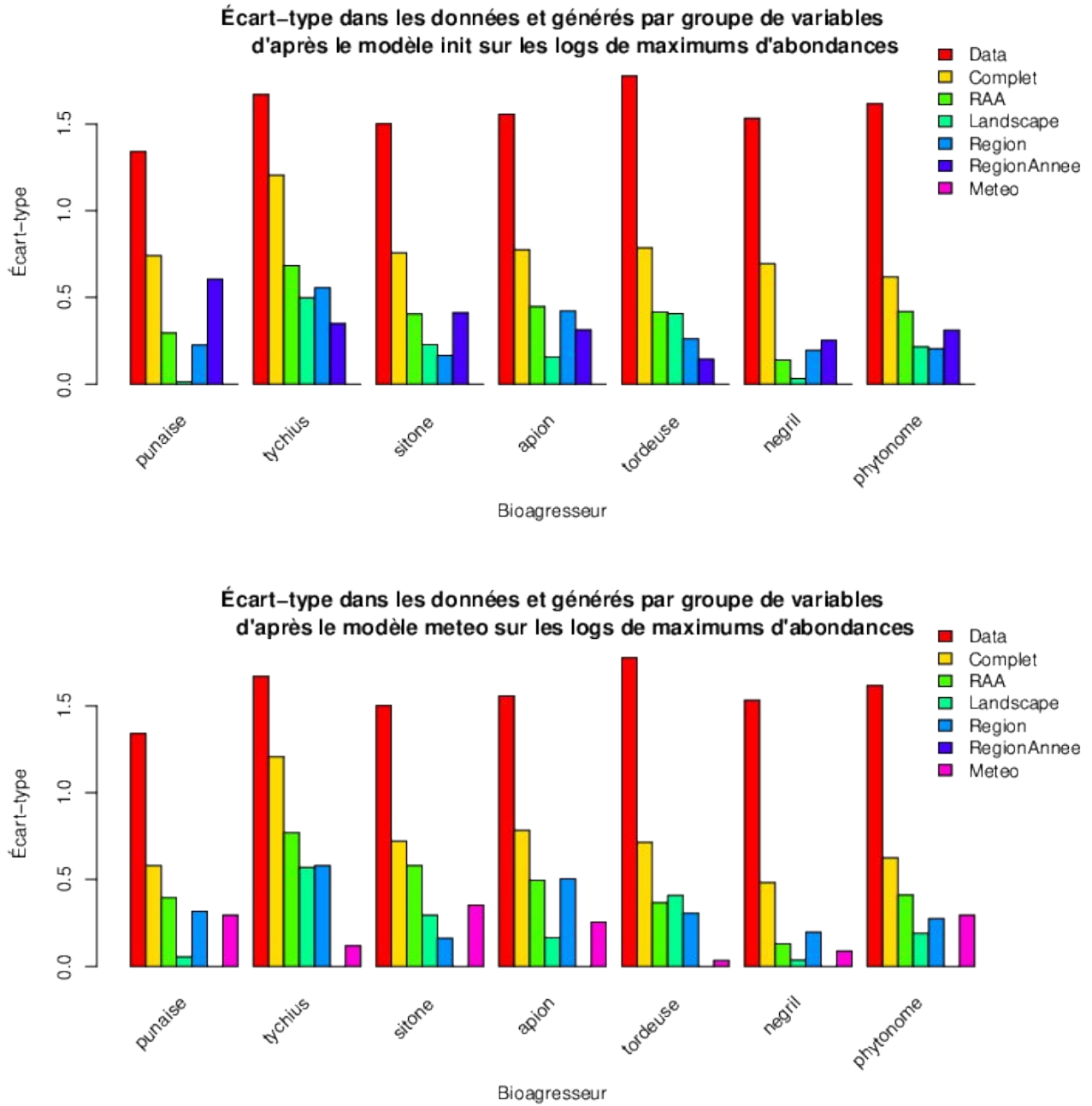


Figure 9 Importance relative des groupes de variables dans les modèles « init » (haut) ou « meteo » (bas) du logs de maximums d'abondance

L'importance est quantifiée par l'écart-type généré par les facteurs de ces groupes de variables sur l'ensemble du jeu de données, l'écart-type observé dans le jeu de données (Data) et par le modèle complet (Complet) sont aussi indiqué à titre de comparaison.

4. Conclusion

Le travail mené met en évidence que la structure et la composition du paysage peuvent influencer l'abondance de plusieurs insectes ravageurs de la luzerne porte-graine, mais ces effets dépendent fortement de la biologie des espèces et de l'échelle spatiale considérée.

Tableau 2 : Synthèse des résultats d'analyses des corrélations des composantes du paysage dans l'explication du logarithme de l'abondance maximale ($\log(\text{abondance_max}+1)$) des différents ravageurs de la luzerne porte-graine à l'échelle communale.

Une corrélation positive favorise le ravageur et une corrélation négative correspond à un effet protecteur de la composante.

Composante	Ravageur						
	Apion	Punaise miride	Sitone	Tordeuse	Tychius	Négril	Phytonome
Luzerne porte-graine	0	(+)	(+)	+	++*	(+)	0
Luzerne (tout type) année N	0	0	(-)	0	0	(+)	(-)
Luzerne (tout type) année N-1	0	0	0	0	0	(+)	-
Prairie année N	-	-	-	(+)	-	(-)	-*
Landes	+	0	0	(-)	0	(+)	(+)
Forêts	0	-	-*	-*	-	0	(-)
Haies	+	+	(-)	(-)	0	+	(+)
Bois traité	(+)	+	0	(-)	+	0	(+)

0 : aucune corrélation (médiane à 0 et l'intervalle de confiance sont autant positif que négatif) ;

+ : corrélation positive (l'abondance du ravageur est favorisée) ;

- : corrélation négative (l'abondance du ravageur est défavorisée) ;

(+/-) : Faible tendance non significative (les coefficients médians ne sont pas différents de 0 mais les intervalles de confiance sont d'un signe unique) ;

+/- : Corrélation (les coefficients médians sont différents de 0 et inférieur à 0.1) ;

++/- : Corrélation forte (les coefficients médians sont différents de 0 et supérieur à 0.1)

* : Corrélation significative au seuil 5% (les intervalles de confiance sont différents de 0)

Tableau 3 : Synthèse des résultats d'analyses des corrélations des composantes du paysage dans l'explication du logarithme de l'abondance médiane ($\log(\text{abondance_med}+1)$) des différents ravageurs de la luzerne porte-graines à l'échelle communale.

Une corrélation positive favorise le ravageur et une corrélation négative correspond à un effet protecteur de la culture par la composante.

Composante	Ravageur						
	Apion	Punaise miride	Sitone	Tordeuse	Tychius	Négril	Phytonome
Luzerne porte-graine	0	+	(+)	+	++*	(+)	(+)
Luzerne (tout type) année N	+	0	0	0	0	+	(+)
Luzerne (tout type) année N-1	(+)	+	0	0	0	(+)	(+)
Prairie année N	(-)	0	(-)	+	(-)	(-)	+
Landes	+	0	+	0	0	0	(-)
Forêts	0	-	-*	-*	-	(+)	(+)
Haies	0	+	-	-	0	0	+
Bois traité	+	+	0	(-)	+	(-)	0

0 : aucune corrélation (médiane à 0 et l'intervalle de confiance sont autant positif que négatif) ;

- + : **corrélation positive** (l'abondance du ravageur est favorisée) ;
- : **corrélation négative** (l'abondance du ravageur est défavorisée) ;
- (+/-) : **Faible tendance non significative** (les coefficients médians ne sont pas différents de 0 mais les intervalles de confiances sont d'un signe unique) ;
- +/- : **Corrélation** (les coefficients médians sont différents de 0 et inférieur à 0.1) ;
- ++/- - : **Corrélation forte** (les coefficients médians sont différents de 0 et supérieur à 0.1)
- *: **Corrélation significative au seuil 5%** (les intervalles de confiance sont différents de 0)

Un premier enseignement concerne le rôle de la luzerne dans le paysage. La **luzerne porte-graine** présente toujours une corrélation positive ou nulle avec les abondances des ravageurs, que l'on considère les abondances médianes (le niveau de population général sur la parcelle) ou maximales (le pic de population sur la parcelle). Pour les ravageurs dont le cycle biologique est étroitement lié à la montée en graine de la culture (**tychius** et **tordeuse**), la surface de **luzerne porte-graine** à l'échelle départementale apparaît comme un facteur explicatif significatif et plus marqué des abondances. Ces résultats traduisent un effet de contexte régional ou d'**inoculum paysager**, cohérent avec la spécialisation de ces espèces. La luzerne de tous types n'a que des tendances à la corrélations positive ou nulle avec les abondances médianes mais des tendances négatives peu ou pas significatives peuvent apparaître pour les abondances maximales.

Les **prairies** semblent jouer un rôle globalement **protecteur** face aux abondances maximales (Tableau 2), bien que la corrélation ne soit significative que pour un seul ravageur, le phytonome. La corrélation peut cependant s'inverser pour la médiane (Tableau 3, phytonome). Les prairies pourraient d'une part, servir de réservoir pour les ravageurs, et d'autre part, favoriser la présence d'auxiliaires, ce qui contribuerait à limiter les pics d'abondance des ravageurs.

Pour les **landes** et les **bois traités**, les corrélations peuvent être dans un sens ou dans l'autre suivant les ravageurs et elles restent généralement faibles.

Les éléments ligneux ne constituent pas un ensemble homogène du point de vue de leurs effets. Les **forêts** et les **haies** présentent parfois des effets de signes opposés selon les ravageurs, ce qui exclut toute interprétation globale du rôle des « ligneux » dans le paysage. Les **forêts** sont cependant majoritairement associées à des **effets protecteurs** plus ou moins visibles (punaise miride, sitone, tordeuse, tychius et phytonome), ce qui pourrait refléter des mécanismes de barrière naturelle ou de régulation biologique. Les **haies** ont des corrélations contrastées suivant les ravageurs. Elles peuvent renforcer l'effet négatifs des forêts sur certains ravageurs (sitone et tordeuse) mais aussi à contrario, suggérer un effet inverse à la forêt bien que toujours non significatif.

La comparaison entre les modèles à l'**échelle communale** et à l'**échelle parcellaire** montre que certains effets peuvent s'exprimer à des distances courtes, mais ces résultats reposent sur des jeux de données plus restreints et doivent être interprétés avec prudence.

Perspectives

Plusieurs éléments permettraient de renforcer les analyses. La **localisation précise des parcelles** de luzerne porte-graine constitue le principal levier d'amélioration, indispensable pour analyser correctement les effets à courte distance et les processus opérant à l'échelle parcellaire. Les données de contexte (territoires, météorologie,...) sont actuellement considérées comme des variables de contrôle pour l'effet des bioagresseurs. Elles ont parfois un rôle important comme le montre la baisse du R² pour les modèles « paysage seul ». Le poids de la variabilité annuelle (représentée par les

variables météo ou par les facteurs « région-année » peut aussi être très variable suivant les ravageurs. Une analyse plus détaillée pourrait être intéressante pour comprendre les déterminants non paysagers de l'abondance des ravageurs (variables météo spécifiques ayant un impact ou régions agricoles agricoles particulièrement touchées par exemple mais aussi variabilité de ces effets suivant les modèles). Un **renforcement des données contextuelles** (mode de conduite et traitement des cultures, cultures hôtes ou ressources alternatives des ravageurs) permettrait d'améliorer la capacité explicative des modèles et de mieux caractériser les facteurs de premier ordre qui structurent les abondances observées et conditionnent la détection d'éventuels effets paysagers de second ordre. De plus, on pourrait ajouter d'autres **variantes de modèles**, comme : 1) Un modèle **sans les données du RPG**, pour presque doubler la taille du jeu de données à l'échelle communale, tout en conservant les autres composantes du paysage (forêts, haies, landes, bois traités et luzerne porte-graine) ; 2) Un modèle **avec les données du RPG d'avant 2015, en assimilant le groupe "culture fourragère" à la luzerne** (éventuellement avec un coefficient correspondant à la proportion de luzerne dans ce groupe), afin d'augmenter également la taille du jeu de données.

Dans l'ensemble, les données qui nous ont été fournies contiennent beaucoup d'information et le temps imparti a permis de faire ressortir des éléments très significatifs. La concentration des surfaces de luzerne grainière est clairement un facteur de risque même si l'échelle spatiale et l'amplitude de l'effet ne peuvent être affirmés avec certitude à ce stade. Les prairies et les autres surfaces de luzerne ont un effet plus contrasté et plus faible. Les haies mériteraient un examen plus détaillé avec plus de données bien géolocalisées car les corrélations observées sur le négril sont extrêmement marquées. Enfin, examiner les facteurs de risque météorologiques et éventuellement les dynamiques temporelles pourrait être très intéressant mais n'était pas possible dans le cadre du budget prévu pour ce contrat.

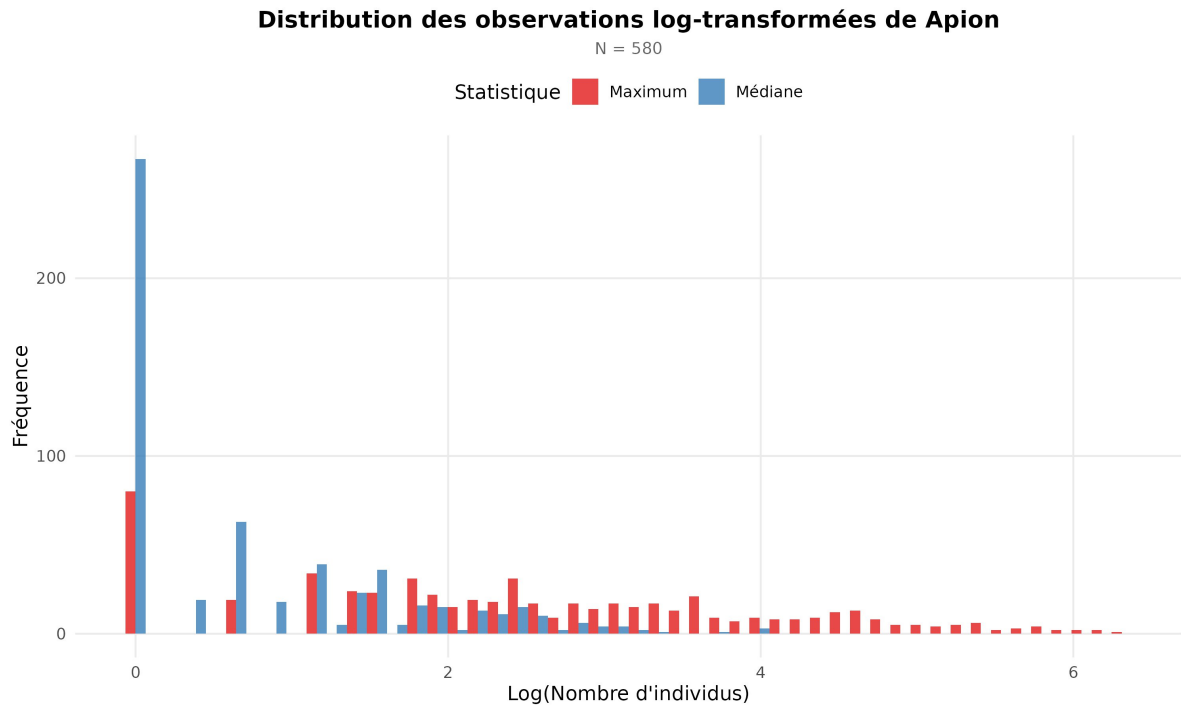
5. Annexes

5.1. Annexe 1 : Variables utilisées dans les modèles avec leurs correspondances dans le code de calcul

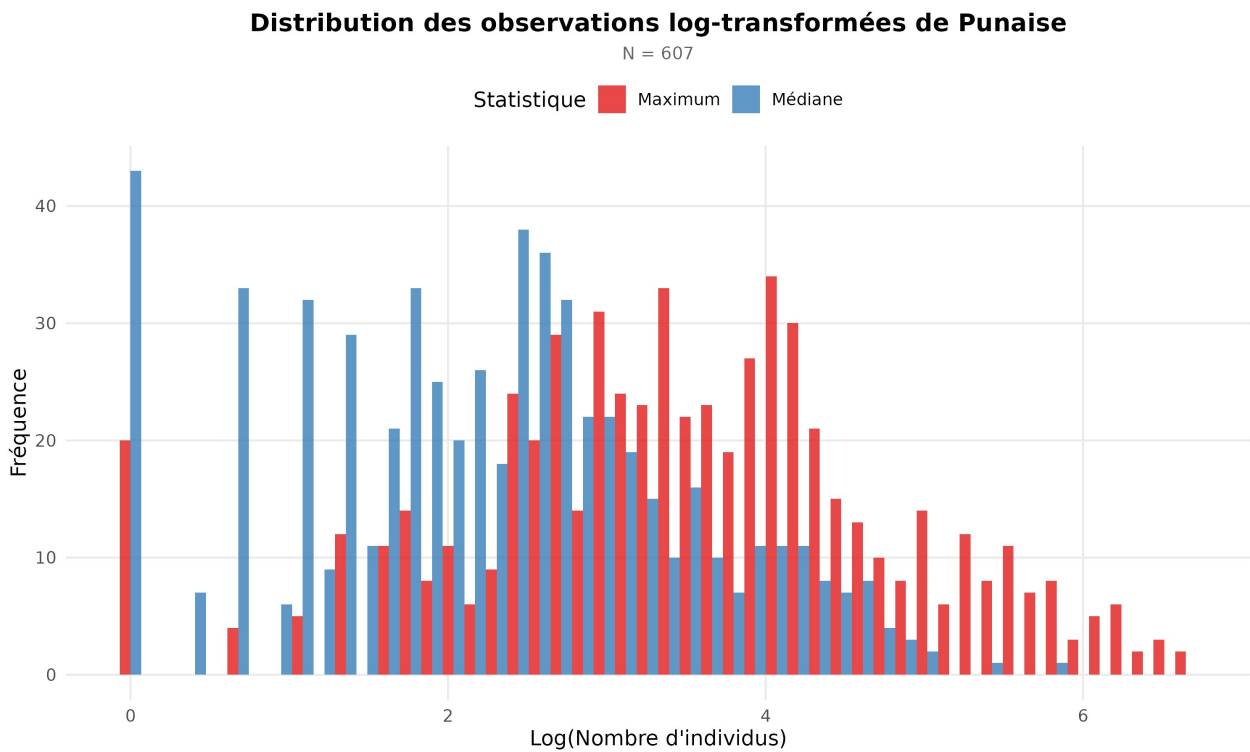
Variables paysagères				
Source	Composante paysagère	Composante paysagère (code)	Nom des variables (code) XX = taille de l'effet	Taille des effets
BDTOPO 2017	Haies	BDTOPO2017GetHaie	BDTOPO2017GetHaie_wXX_area BDTOPO2017GetHaie_wXX_perimeter	200m; 1000m; 5000m; 10000m
	Forêts	BDTOPO2017GetForet	BDTOPO2017GetForet_wXX_area BDTOPO2017GetForet_wXX_perimeter	
	Bois traités	BDTOPO2017GetBoisTraitesIGN	BDTOPO2017GetBoisTraitesIGN_wXX_area BDTOPO2017GetBoisTraitesIGN_wXX_perimeter	
	Landes	BDTOPO2017GetLandelIGN	BDTOPO2017GetLandelIGN_wXX_area	
			BDTOPO2017GetLandelIGN_wXX_perimeter	
	RPG	Luzerne année N	LUZ_all.*area\$	
Luzerne année N-1		LUZ_all.*area_Prec\$	LUZ_all_wXX_area_Prec	
Prairies année N		PRAIRIE_all	PRAIRIE_all_wXX_area	
SEMAE	Luzerne porte-graine	LuzPG_surface\$	LuzPG_surface	Département
Variables environnementales non paysagères				
Source	Nom des types de variables	Correspondance des types variables (code)	Nom des variables (code)	Abréviations
INSEE	Anciennes régions administratives	region	region_YY	YY = code région ou région agricole YY = année
	Anciennes régions administratives - année	regionAnnee	regionAnnee_YY-YY	
Agrete	Régions agricole	regAgri	ra_YY	
SAFRAN	Température moyenne	CLIM	CLIM_TM_MM	MM = mois ou groupe de mois
	Température minimale		CLIM_TN_MM	
	Température maximale		CLIM_TX_MM	
	Cumul de précipitation		CLIM_RR_MM	
	Rayonnement global		CLIM_RG_MM	

5.2. Distributions des abondances par bioagresseur

5.2.1. Apion



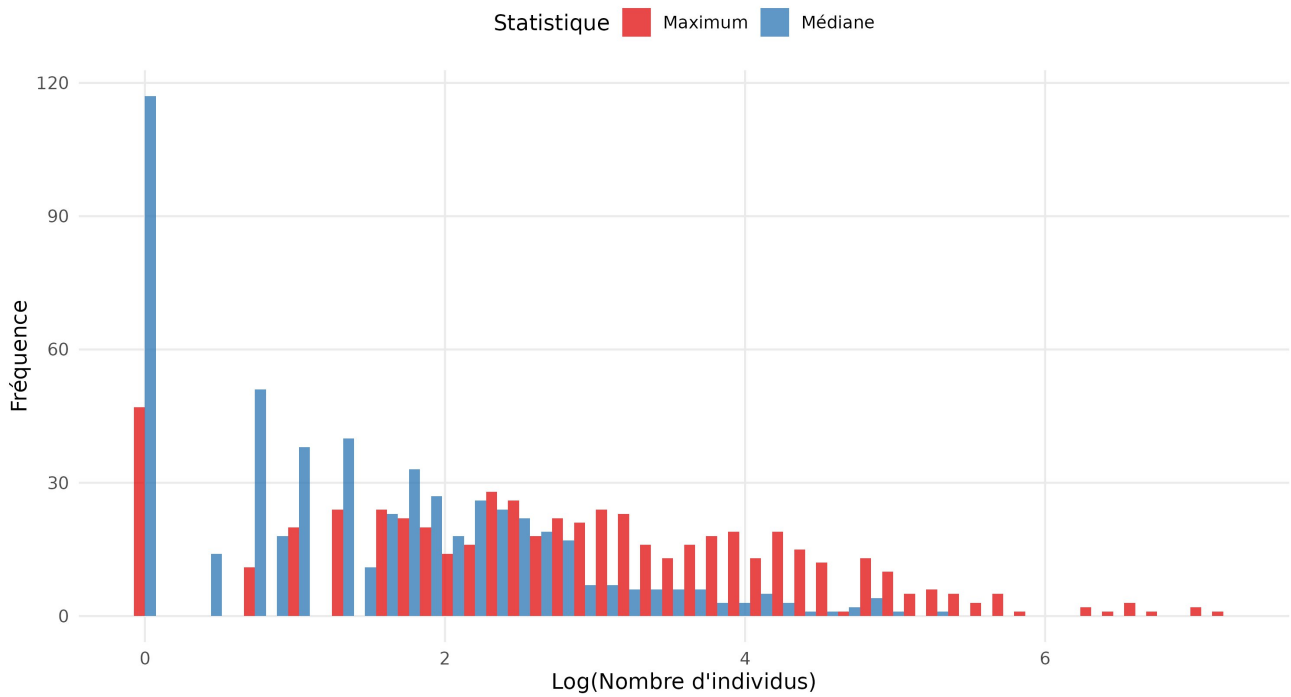
5.2.2. Punaises mirides



5.2.3. Sitones

Distribution des observations log-transformées de Sitone

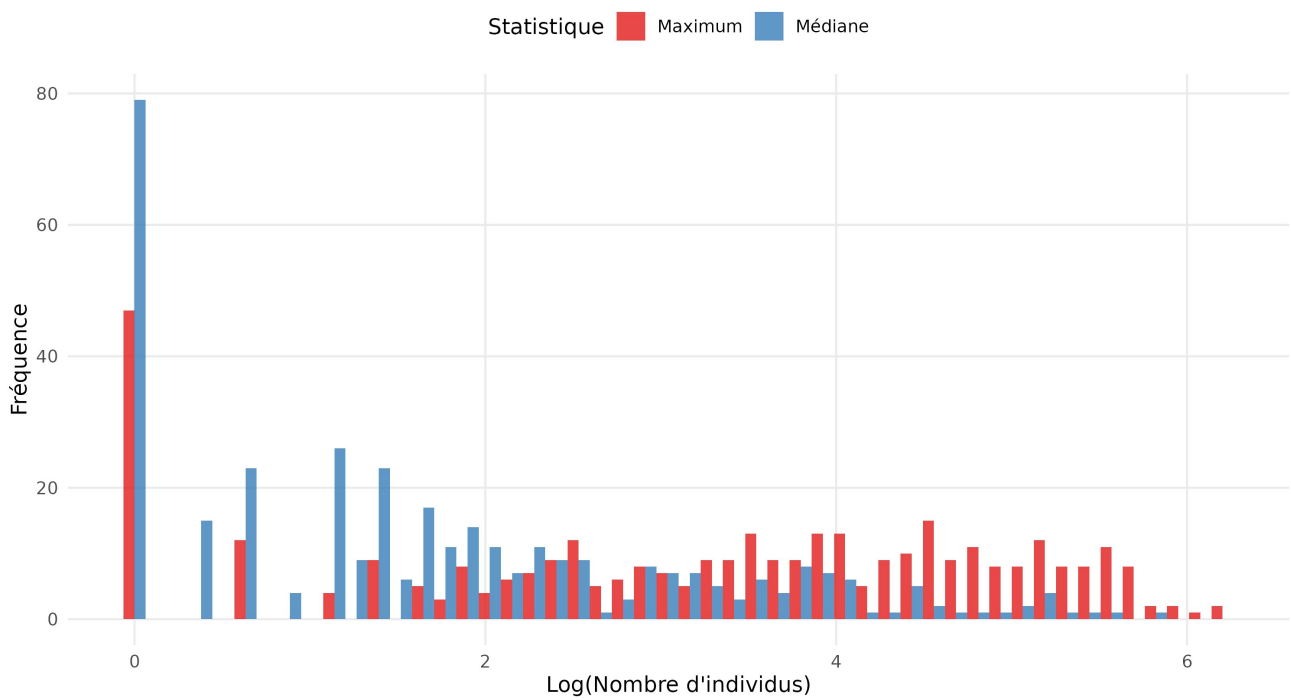
N = 560



5.2.4. Tordeuse

Distribution des observations log-transformées de Tordeuse

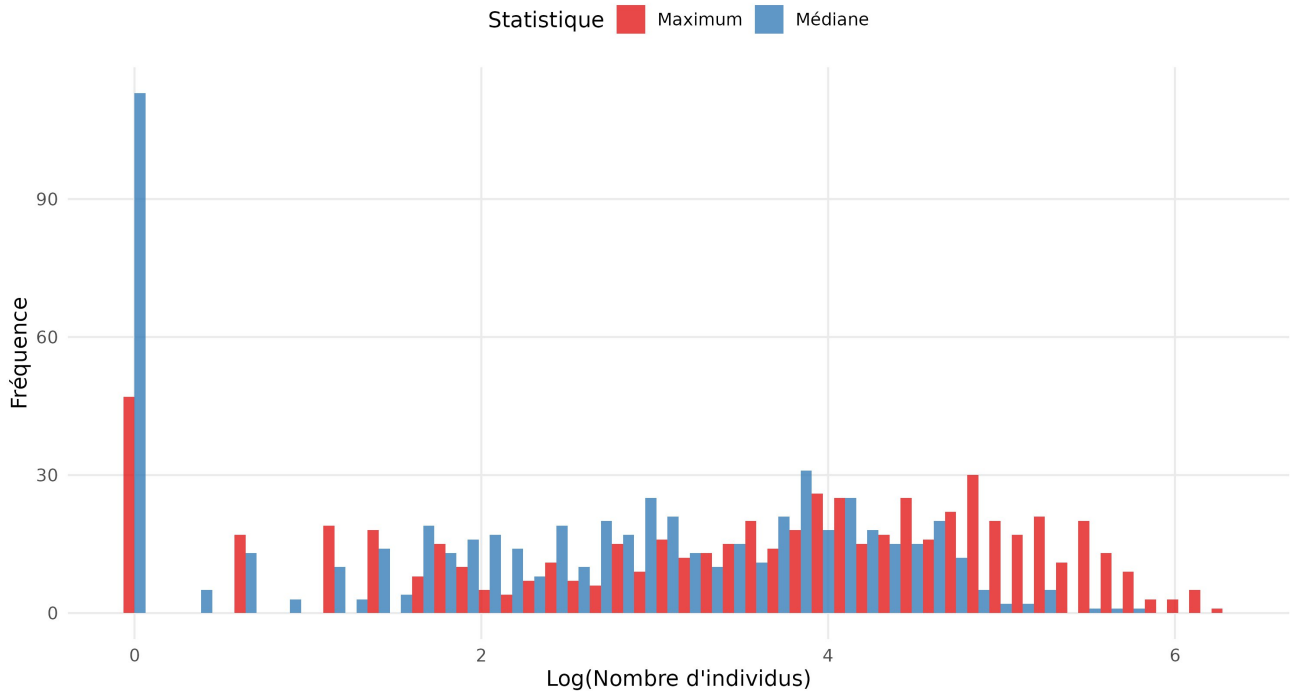
N = 361



5.2.5. Tychius

Distribution des observations log-transformées de Tychius

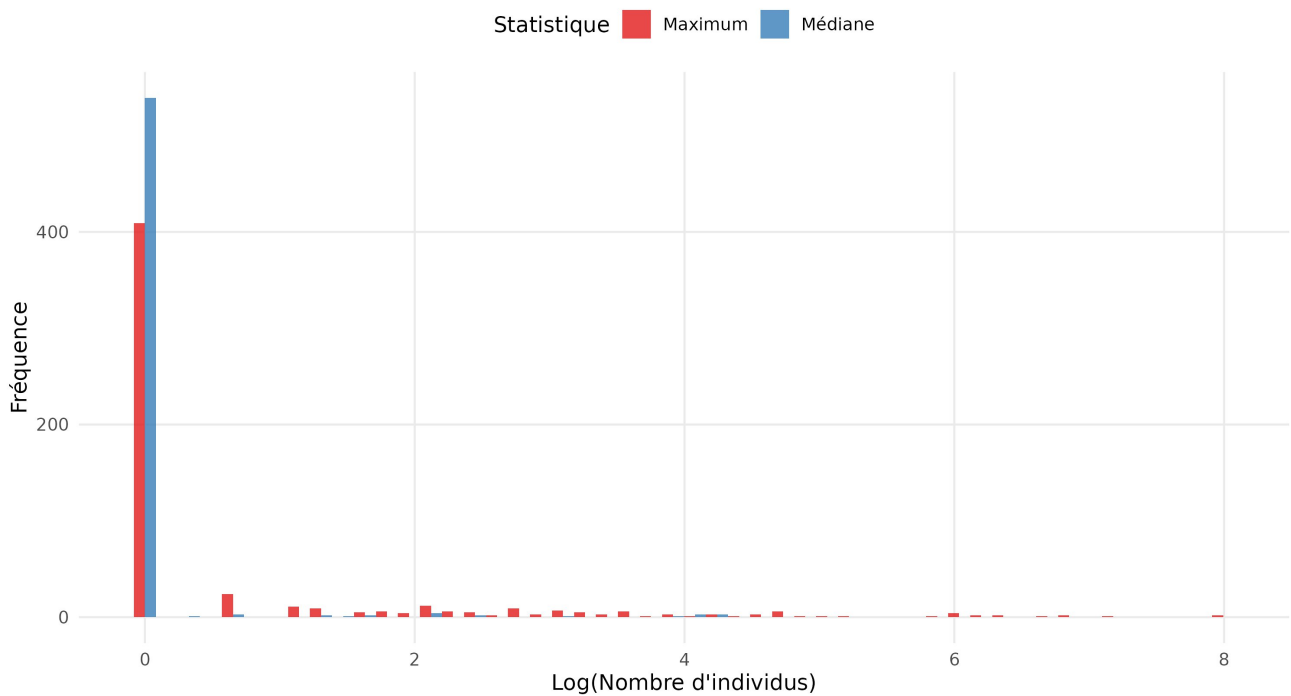
N = 605



5.2.6. Négril

Distribution des observations log-transformées de Negril

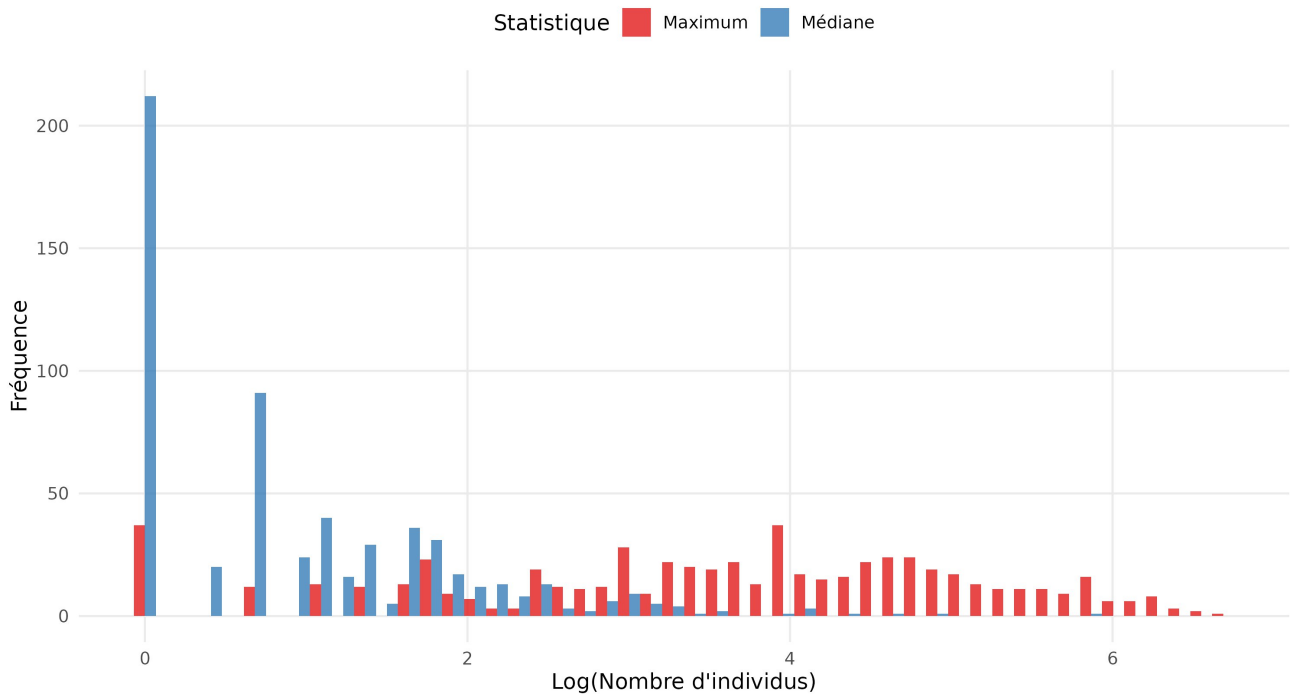
N = 562



5.2.7. Larves de phytonomes

Distribution des observations log-transformées de Phytonome

N = 607



5.3. Représentations graphiques de l'ajustement

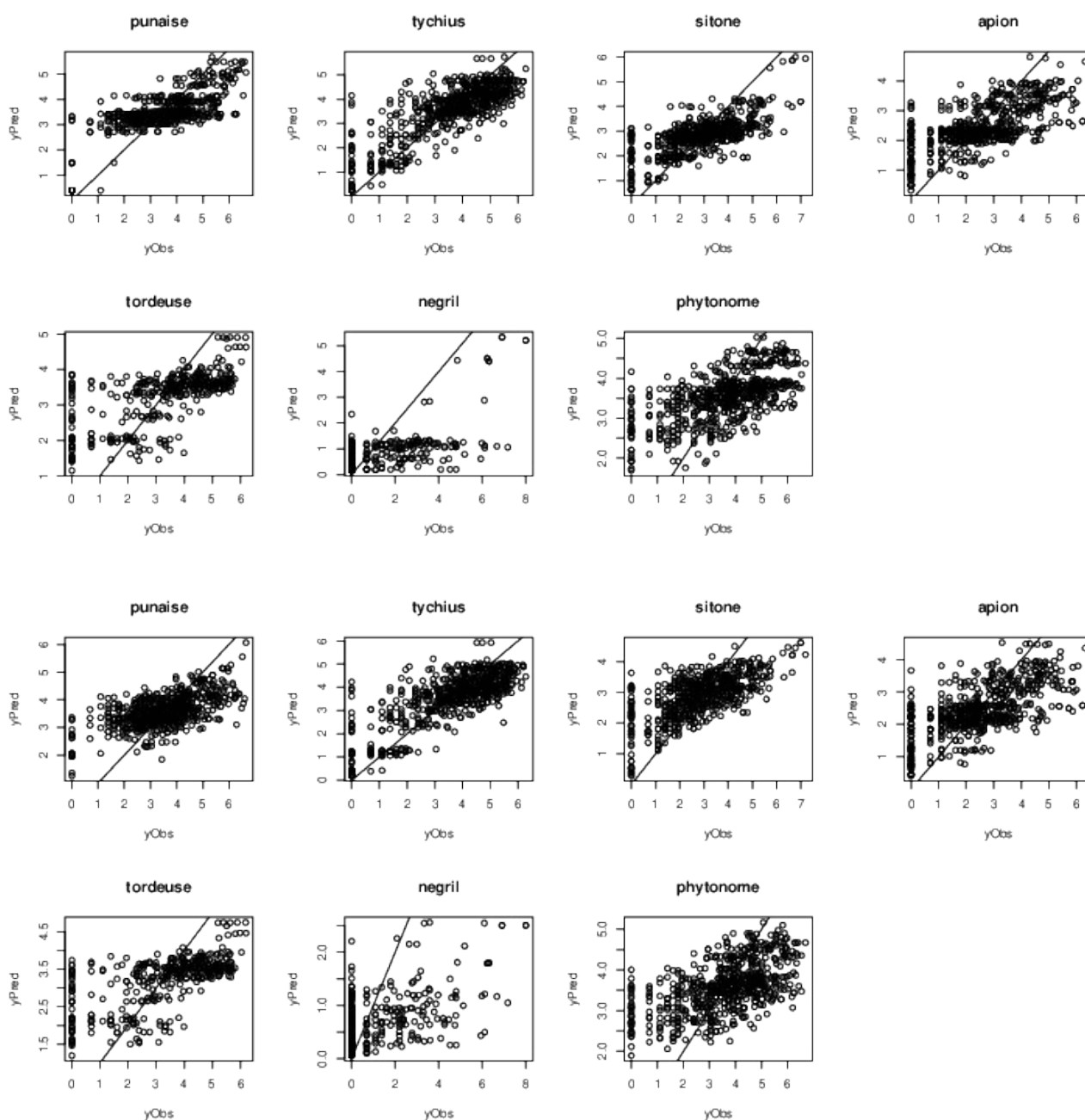


Figure supplémentaire : Valeurs observées et prédites par les modèles « Init » (haut) et « Météo » (bas) des log des maximum d'abondance pour chacun des bioagresseurs

Dans l'ensemble les prédictions des hauts niveaux d'infestation sont assez fiables. En revanche, les prédictions de niveaux moyens peuvent correspondre à des observations faibles ou nulles de ravageurs. On notera aussi la difficulté à bien décrire les hautes abondances de punaises simplement avec les variables météo (bas) alors que l'effet fixe « région-année » est efficace. Cela indique d'une part que les pressions hautes sont assez homogènes à l'échelle des régions une année donnée et

d'autre part qu'il y a de la variabilité annuelle non expliquée directement par les variables météo prises en compte. Cette variabilité pourrait être explorée dans des études à venir sur ces données en testant par exemple d'autres variables météo ou en introduisant des variables de présence d'auxiliaires de culture.