

# Mastère Data Science pour l'ingénierie

## Rapport de stage

Présenté par : M. Abou CISSE

# Modélisation de la présence de bioagresseurs (insectes, maladies fongiques) des grandes cultures

Tuteur de Stage : M. Corentin BARBU

Tuteur Académique : M. Jean-Marc BOURINET

14 septembre 2022

1	Présentation de l'INRAE .....	4
2	Résumé.....	4
3	Abstract .....	4
4	Remerciements.....	5
5	Données sur les bioagresseurs.....	5
6	Présentation des bioagresseurs.....	8
6.1	Septoriose .....	8
6.2	Rouille brune du blé .....	8
6.3	Le puceron vert du pêcher ( <i>Myzus persicae</i> ).....	9
6.4	Le piétin-versé.....	10
6.5	Méligète du colza .....	11
6.6	Le charançon du bourgeon terminal .....	11
6.7	Le charançon de la tige du colza .....	12
6.8	Charançon de la tige du chou .....	13
6.9	Altise petite des crucifères .....	13
7	Inventaire des bioagresseurs dans les parcelles .....	14
7.1	Méthode d'évaluation des maladies dans les parcelles .....	14
7.1.1	Septoriose.....	15
7.1.2	Rouille brune du blé.....	15
7.1.3	Le piétin-versé.....	15
7.2	Technique de piégeage et méthode de comptage des insectes dans les parcelles .....	16
7.2.1	Piège coloré à liquide (Cuvette jaune) .....	16
7.2.2	Méthode de comptage des insectes dans les parcelles .....	16
7.3	Résumé des métriques utilisées pour chaque bioagresseur .....	17
8	Calcul de la pression annuelle des bioagresseurs dans une parcelle.....	17
8.1	Seuil médian .....	17
8.2	Pression annuelle des bioagresseurs d'une même espèce, dans une parcelle .....	17
8.3	Utilité de calculer la pression des bioagresseurs dans une parcelle .....	18
9	Localisation des parcelles liées aux bioagresseurs étudiés sur la carte de la France .....	18
9.1	Répartition des parcelles liées à la septoriose au niveau de la France selon les années ....	19
9.2	Répartition des parcelles liées à la rouille brune du blé au niveau de la France selon les années.....	20
9.3	Répartition des parcelles liées au puceron vert du pêcher au niveau de la France selon les années.....	21
9.4	Répartition des parcelles liées au piétin-versé au niveau de la France selon les années....	22
9.5	Répartition des parcelles liées au méligète du colza au niveau de la France selon les années.....	23
9.6	Répartition des parcelles liées au charançon du bourgeon terminal au niveau de la France selon les années .....	24

9.7	Répartition des parcelles liées au charançon de la tige du colza au niveau de la France selon les années .....	25
9.8	Répartition des parcelles liées au charançon de la tige du chou au niveau de la France selon les années .....	26
9.9	Répartition des parcelles liées à l'altise petite des crucifères au niveau de la France selon les années .....	27
10	Distribution annuelle de chaque bioagresseur sur l'ensemble des parcelles .....	27
10.1	Nombre de plantes mensuel touchées par la maladie de la septoriose sur l'ensemble des parcelles	28
10.2	Nombre de plantes mensuel touchées par la maladie de la rouille brune sur l'ensemble des parcelles .....	29
10.3	Nombre de pucerons verts du pêcher mensuel compté sur l'ensemble des parcelles .....	30
10.4	Nombre de plantes mensuel touchées par la maladie du piétin verse sur l'ensemble des parcelles	31
10.5	Nombre de méligrèthes mensuel compté sur l'ensemble des parcelles .....	32
10.6	Nombre de charançons du bourgeon terminal mensuel compté sur l'ensemble des parcelles	33
10.7	Nombre de charançons de la tige du colza mensuel compté sur l'ensemble des parcelles	34
10.8	Nombre de charançons de la tige du chou mensuel compté sur l'ensemble des parcelles	35
10.9	Nombre d'altises mensuel compté sur l'ensemble des parcelles .....	36
11	Pression annuelle des bioagresseurs d'espèce $i$ prédite dans une parcelle .....	36
11.1	Pression annuelle des bioagresseurs d'espèce $i$ , prédite en utilisant la moyenne pondérée	36
11.1.1	Tableau récapitulatif des hyperparamètres optimaux ( $h$ ) des bioagresseurs .....	48
11.2	Pression annuelle des bioagresseurs d'espèce $i$ , prédite en utilisant la méthode des $k$ -NN ( $k$ plus proches voisins) .....	49
11.2.1	Tableau récapitulatif des hyperparamètres optimaux ( $k$ ) des bioagresseurs .....	58
11.3	Pression annuelle des bioagresseurs d'espèce $i$ , simulée .....	59
12	Comparaison des modèles sur les observations de chaque année .....	59
12.1	Comparaison des résultats obtenus sur la septoriose .....	60
12.2	Comparaison des résultats obtenus sur la rouille brune du blé .....	61
12.3	Comparaison des résultats obtenus sur le puceron vert de pêcher .....	62
12.4	Comparaison des résultats obtenus sur le méligrète de colza .....	63
12.5	Comparaison des résultats obtenus sur le charançon du bourgeon terminal .....	64
12.6	Comparaison des résultats obtenus sur le charançon de la tige du colza .....	65
12.7	Comparaison des résultats obtenus sur le charançon de la tige du chou .....	66
12.8	Comparaison des résultats obtenus sur l'altise petite des crucifères .....	67
13	Conclusion et perspectives .....	68
14	Bibliographies .....	69

## **1 Présentation de l'INRAE**

L'Institut national de la recherche agronomique (INRA) est un organisme français de recherche agronomique qui a existé de 1946 à 2019. L'Institut a fusionné avec l'IRSTEA (Institut national de recherche en sciences et technologies pour l'environnement et l'agriculture) le 1er janvier 2020 pour former l'Institut national de recherche pour l'agriculture, l'alimentation et l'environnement (INRAE).

Il est constitué de :

- + de dix mille agents ;
- 18 centres de recherche en France ;
- 14 départements scientifiques spécialisés dans divers domaines (biologie, génétique, numérique, etc.).

INRAE a pour ambition d'être l'un des leaders mondiaux de la recherche répondant aux enjeux sociétaux :

- sécurité alimentaire et nutritionnelle ;
- transition des agricultures (agroécologie, réduction de la chimie) ;
- gestion des ressources naturelles et des écosystèmes (eau, sol, forêt) ;
- érosion de la biodiversité ;
- économie circulaire et risques naturels.

## **2 Résumé**

### **Sujet : Modélisation de la présence de bioagresseurs (insectes, maladies fongiques) des grandes cultures**

L'utilisation des produits phytosanitaires (fongicides, insecticides, pesticides, etc...) par les agriculteurs constitue une menace majeure pour la biodiversité, qui met également en danger l'agriculture elle-même. Les grandes cultures (blé, maïs, colza, ...) représentent plus de la moitié de l'utilisation de produits phytosanitaires en France. Les bioagresseurs (insectes, maladies fongiques) sont à l'origine de l'utilisation des produits phytosanitaires. Pour réduire l'utilisation de ces derniers, il est important de quantifier les bioagresseurs pour mesurer la gravité sur les cultures. C'est dans ce sens que chaque bioagresseur est associé à ce que nous appelons "pression bioagresseur" (définition ci-dessous). Il est important de montrer aussi la relation entre rendement agricole et présence de bioagresseurs.

Ce stage entre dans le cadre du projet MoCoRiBA qui met en relation les observations de présence de bioagresseurs (insectes et maladies fongiques) et des données de pratiques et de rendement à la parcelle. Les données d'observation à l'échelle du territoire métropolitain n'étant pas présentes aux mêmes endroits que les données de pratiques et de rendement, il est important d'avoir des modèles fiables d'interpolation de la présence de ces bioagresseurs. De tels modèles pourraient aussi être utilisés en cours de campagne agricole pour estimer la gravité des attaques et leur impact potentiel sur les cultures. Pour élaborer ces modèles, pour chacun des principaux bioagresseurs étudiés (neuf au total), nous utilisons des approches de fouille de données sur des données d'épidémiologie végétale (quelques centaines à quelques milliers de points-années par bioagresseur).

## **3 Abstract**

The use of phytosanitary products (fungicides, insecticides, pesticides, etc.) by farmers is a major threat to biodiversity, which also endangers agriculture itself. Field crops (wheat, corn, rapeseed, ...)

represent more than half of the use of phytosanitary products in France. Bioagressors (insects, fungal diseases) are at the origin of the use of phytosanitary products. In order to reduce the use of these products, it is important to quantify the pests and diseases in order to measure their severity on the crops. It is in this sense that each pest is associated with what we call "pest pressure" (definition below). It is also important to show the relationship between crop yield and pest presence.

This internship is part of the MoCoRiBA project, which links observations of pests (insects and fungal diseases) with data on practices and yields in the field. Since observation data at the metropolitan scale are not present in the same places as practice and yield data, it is important to have reliable models for interpolating the presence of these pests. Such models could also be used during the crop year to estimate the severity of attacks and their potential impact on crops. To develop these models, for each of the main pests studied (nine in total), we use data mining approaches on plant epidemiological surveillance data (a few hundred to a few thousand point-years per pest).

#### 4 Remerciements

Je remercie mon tuteur Corentin Barbu pour ses connaissances qui m'ont servi de travailler sur un projet intéressant. Corentin, merci pour ton encadrement, pour tes relectures du rapport de stage et ta disponibilité, pour les multiples discussions sur le sujet de stage, pour tout ce que tu as partagé avec moi pendant ces six mois.

Un merci particulier à Baptiste Rayon pour son aide précieuse et encouragement pendant le stage.

Merci aux stagiaires avec qui j'ai partagé le même bureau au cours des 6 mois, à savoir : Théo Froment, Claire Ayumi Bazille, Marie-Thérèse Morisson, Mariem Ben Rahal. Merci aussi à mes très chers amis Patrice Humblot, Simon Buresi.

Je remercie l'INRAE de Grignon de m'avoir accueilli au cours de ces moments studieux.

#### 5 Données sur les bioagresseurs

Les données sur les bioagresseurs viennent de l'Epiphyt et de la base de données Vigicultures. Dans le cadre du réseau de surveillance des bioagresseurs des grandes cultures, la base de données Vigicultures centralise une grande partie des données du réseau d'épidémiosurveillance national et a été développé par l'institut agricole français ARVALIS-Institut du végétal.

Nous avons extrait ces observations de ravageurs et de maladies pour 7 cultures différentes (Betterave, Blé tendre d'hiver, Colza d'hiver, Maïs, Orge d'hiver, Pomme de terre, Tournesol).

Voici une liste non exhaustive des attributs (colonnes) de notre dataset :

- **ID\_plot** : Numéro d'identification de la parcelle ;
- **Culture** : Type de culture de la parcelle ;
- **Campagne** : Année de récolte ;
- **Code\_departement** : Code département où se trouve la parcelle ;
- **Code\_insee** : Code élaboré par l'INSEE, à 5 chiffres. Concaténation du code département (2 chiffres) et du code commune (3 chiffres) ;

- **Latitude** : Latitude de la parcelle ;
- **Longitude** : Longitude de la parcelle ;
- **database** : le nom de la base de données (Vigiculture ou Epiphyt) d'où provient l'observation ;
- **Code\_obs\_gen** : ID du type d'observation (métrique) ;
- **Obs\_date** : Date d'observation ;
- **Semaine** : Semaine de l'année où s'est faite l'observation ;
- **Valeur** : Valeur de l'observation ‘(qui peut être un comptage ou un pourcentage) ;
- **stade\_sensible** : Selon le nombre d'une observation recensé dans la parcelle. Il lui est associé une valeur booléen (True pour dire que la parcelle est dans un stade sensible, False pour dire que la parcelle ne l'est pas) ;
- **Reseau** : Réseau d'observation (département, région) ;
- ...

Le dataset comporte **1 290 479 observations** et **27 colonnes**. J'ai fait le choix de représenter ci-dessous 11 colonnes et 13 observations pour donner une idée sur les données.

#### *Quelques observations et quelques colonnes de notre dataset*

ID_plot	Culture	Campagne	Latitude	Longitude	Code_departement	Code_obs_gen	database	Obs_date	Semaine	Reseau
131546	Colza d'hiver	2013	46.07621	4.880226	01	MeI%P	Vigiculture	2013-04-15	NA	RHONE-ALPES
712594	Blé tendre d'hiver	2019	49.01853	7.961990	67	SEPF3	Epiphyt	2019-05-21	21	Alsace
513873	Maïs	2017	46.54154	5.174190	71	PYR_PHE_NB_ADULTES	Epiphyt	2017-08-01	31	Bourgogne
30710	Tournesol	2010	46.05173	4.805574	01	PhaFeu	Vigiculture	2010-09-16	NA	RHONE-ALPES
502151	Blé tendre d'hiver	2017	46.26612	4.983794	01	OIDF3	Vigiculture	2017-04-24	17	Rhône-Alpes
514070	Orge d'hiver	2017	47.33491	3.598110	58	RHYNCF2	Epiphyt	2017-05-22	21	Bourgogne
146219	Colza d'hiver	2013	45.91814	4.832847	01	MeI%P	Vigiculture	2013-04-08	NA	RHONE-ALPES
160036	Maïs	2013	45.87196	5.186920	01	SES_PHE_NB_ADULTES	Vigiculture	2013-05-21	21	Rhône-Alpes
286588	Blé tendre d'hiver	2015	45.96433	5.127025	01	SEPF3	Vigiculture	2014-10-27	44	Rhône-Alpes
502151	Blé tendre d'hiver	2017	46.26612	4.983794	01	SEPF3	Vigiculture	2017-05-30	22	Rhône-Alpes
394348	Colza d'hiver	2016	45.91284	4.909756	01	LGA%P	Vigiculture	2015-11-02	NA	RHONE-ALPES
368240	Tournesol	2015	46.13849	5.304192	01	PucVCri	Vigiculture	2015-06-01	NA	RHONE-ALPES
502151	Blé tendre d'hiver	2017	46.26612	4.983794	01	HELMIN_F1	Vigiculture	2017-04-24	17	Rhône-Alpes

...with 1 290 466 more rows, and 16 more variables

Le tableau suivant donne les bioagresseurs (insectes, maladies fongiques) par culture :

*Bioagresseurs par culture*

Cultures	Bioagresseurs
Betterave	Puceron noir
	Pégomyie
	Puceron vert
	Cercosporiose
	Rouille
	Teigne de la betterave
Blé tendre d'hiver	Septoriose des céréales tritici
	Oïdium des céréales
	Helminthosporiose de l'orge D teres
	Rouille brune du blé
	Puceron
	Piétin verste
	Puceron vecteurs de viroses
Colza d'hiver	Méligète du colza
	Charançon de la tige du colza
	Altise Grosse d'hiver du Colza
	Charançon du bourgeon terminal
	Charançon de la tige du chou ( <i>Ceutorhynchus pallidactylus</i> )
	Puceron vert du pêcher <i>Myzus persicae</i>
	Sclérotiniose
	Altise petite des crucifères
	Altise
Maïs	Sésamie
	Taupin
	Pyrale du maïs
	Chrysomèle des racines du maïs <i>Diabrotica virgifera</i>
	Insectes foreurs de la tige
Orge d'hiver	Helminthosporiose de l'orge D teres
	Rhynchosporiose
	Oïdium des céréales
	Piétin verste
	Puceron vecteurs de viroses
	Taupin
Pomme de terre	Doryphores
	Alternariose de la pomme de terre
	Mildiou de la pomme de terre
	Puceron
Tournesol	Limace
	Puceron vert du prunier
	<i>Phoma macdonaldi</i> Maladie des tâches noires
	<i>Phomopsis</i> du tournesol
	Puceron noir de la fève

## **6 Présentation des bioagresseurs**

Dans cette partie, on présente les bioagresseurs (maladies et insectes) que nous avons choisis pour évaluer leur présence dans les champs de culture. Pour présenter les bioagresseurs, on s'est appuyé sur les écrits de Arvalis sur ce sujet, présents depuis leur site internet.

### **6.1 Septoriose**

#### Description :

La septoriose est une maladie foliaire caractérisée par l'apparition de taches rectangulaires ou ovales initialement jaunes (chlorose) puis brunes (nécrose).



**Symptôme de septoriose sur blé**

*Crédit Photo : ARVALIS – Institut du végétal*

#### Localisation :

La maladie est retrouvée partout en France et est présente sur les cultures de blé tendre, blé dur et triticale.

Les données dont nous disposons sur la septoriose concernent les cultures du blé tendre d'hiver.

#### Dégâts:

À l'échelle de la parcelle, elle crée en moyenne des pertes évaluées à 17 quintaux/hectare (jusqu'à 50 quintaux/ha dans les parcelles non traitées).

#### Facteurs favorables :

La maladie de la septoriose est favorisée par un temps chaud et humide. La pluie assure la propagation de la maladie par effet splashing.

#### Période de présence :

Mars, Avril, Mai, Juin

### **6.2 Rouille brune du blé**

#### Description :

La rouille brune du blé est une maladie foliaire caractérisée par l'apparition de pustules de couleur orange à brune.



### Rouille brune du blé

*Crédit Photo : ARVALIS – Institut du végétal*

#### Localisation :

Cette maladie est localisée dans les zones tempérées c'est pour cela qu'elle est beaucoup présent dans l'Ouest et la moitié Sud de la France. On retrouve cette maladie sur les cultures de blé, de l'orge et du seigle

Les données dont nous disposons sur la rouille brune concernent les cultures du blé tendre d'hiver

#### Dégâts :

Si elle est mal gérée, la rouille brune du blé peut causer à l'échelle de la parcelle, des pertes de rendement allant jusqu'à 40 quintaux/hectares.

#### Facteurs favorables :

Elle est favorisée par la pluie et des températures douces, c'est pour cela qu'elle est beaucoup présente dans l'Ouest et la moitié Sud de la France.

#### Période de présence :

Mars, Avril, Mai, Juin

## 6.3 Le puceron vert du pêcher (*Myzus persicae*)

#### Description :

Le puceron vert du pêcher dont le nom scientifique est *myzus persicae*. C'est un insecte ravageur dont l'adulte, avec ou sans aile, mesure de 1,5 à 2,5 mm et possède un abdomen vert taché de noir.



### Le puceron vert du pêcher (*myzus persicae*)

*Crédit Photo : ARVALIS – Institut du végétal*

### Localisation :

Il est visible au niveau des feuilles de colza dès septembre/octobre. Il revient ensuite au printemps. Les données dont nous disposons sur le puceron vert du pêcher concernent les cultures du colza.

### Dégâts :

À l'échelle de la parcelle, Il peut créer des pertes de rendement qui peuvent aller jusqu'à 8 quintaux/hectares.

### Facteurs favorables :

Le réchauffement des températures provoque leur multiplication.

## 6.4 Le piétin-verse

### Description :

Le piétin-verse est une maladie du blé tendre, de l'orge, etc... La maladie affecte uniquement les bas de tiges.



**Maladie du piétin-verse sur le blé tendre d'hiver**

*Credit Photo : ARVALIS – Institut du végétal*

### Localisation :

Le piétin-verse est courant dans les régions aux automnes et hiver doux et humides.

Les données dont nous disposons sur le piétin-verse concernent les cultures du blé tendre d'hiver.

### Dégâts :

À l'échelle de la parcelle, la maladie entraîne des pertes de rendement de 5 à 20 quintaux par hectare.

### Facteurs favorables :

De fortes précipitations et des températures douces en automne et en hiver favorisent le développement de la maladie.

### Période de présence :

Février, Mars, Avril, Mai, Juin

## 6.5 Méligète du colza

### Description :

Le méligète du colza est un insecte ravageur qui mesure entre 1.5 à 2.5 mm. Il a une couleur noir brillant et une forme aplatie.



**Méligète du colza adulte sur une inflorescence de colza**

*Crédit Photo : ARVALIS – Institut du végétal*

### Localisation géographique :

Le méligète du colza est présent dans la plupart des régions françaises. Les données dont nous disposons sur le méligète du colza concernent les cultures du colza d'hiver.

### Dégâts :

À l'échelle de la parcelle, les méligèthes du colza peuvent provoquer des pertes de rendement allant de 30 à 40 % sur colza d'hiver.

### Facteurs favorables :

Les températures faibles ralentissent l'activité des méligèthes du colza. Ils commencent à voler lorsque la température du sol dépasse 15°C. Donc leur attaque sur le colza dépend des conditions météorologiques (pluies, vent, températures faibles pour la saison).

## 6.6 Le charançon du bourgeon terminal

### Description :

C'est un insecte ravageur du colza qui mesure entre 2,5 et 3,7 mm. Il a un corps noir brillant avec des poils courts clairsemés, des taches rouges sur les flancs entre la poitrine et l'abdomen et des taches blanches sur le dos.



### **Charançon du bourgeon terminal du colza**

*Crédit Photo : ARVALIS – Institut du végétal*

#### Localisation géographique :

Le charançon du bourgeon terminal est surtout présent à l'Est et Centre de la France. Il est apparu plus récemment dans le Sud-Ouest de la France.

Les données dont nous disposons sur le charançon du bourgeon terminal concernent les cultures du colza d'hiver.

#### Dégâts :

À l'échelle de la parcelle, ils créent une perte de rendement à partir de 30 % de plantes à port buissonnant.

#### Facteurs favorables :

Les bonnes conditions climatiques (températures douces, ensoleillement) favorisent l'arrivée des charançons du bourgeon terminal.

## **6.7 Le charançon de la tige du colza**

#### Description :

C'est un des insectes ravageurs du colza, il a un corps gris cendré, long de 3-4mm, avec des stries fines séparées par des intervalles couverts de courts poils grisâtres. Il a des pattes noires.



### **Charançon de la tige du colza**

*Crédit Photo : ARVALIS – Institut du végétal*

#### Localisation géographique :

On retrouve le charançon de la tige du colza au printemps sur l'ensemble du territoire français.

Les données dont nous disposons sur le charançon de la tige du colza concernent les cultures du colza d'hiver.

#### Dégâts :

À l'échelle de la parcelle, les charançons de la tige du colza peuvent causer des pertes de rendement allant jusqu'à 50 %.

#### Facteurs favorables :

Les températures douces favorisent l'activité du charançon de la tige du colza.

## 6.8 Charançon de la tige du chou

#### Description :

C'est un des insectes ravageurs du colza, à ne peut confondre avec le charançon de la tige du colza. Le charançon de la tige du chou est plus petit, il a le bout des pattes roux et une tache blanchâtre dorsale entre le thorax et l'abdomen.



**Charançon de la tige du chou**

*Crédit Photo : ARVALIS – Institut du végétal*

#### Localisation géographique :

On le retrouve partout en France. Les données dont nous disposons sur le charançon de la tige du chou concernent les cultures du colza d'hiver.

#### Dégâts :

Il est moins dangereux que son cousin (le charançon de la tige du colza) même en cas de forte attaque les pertes de rendement sont inférieures à 20%.

#### Facteurs favorables :

Les températures douces favorisent l'activité du charançon de la tige du chou.

## 6.9 Altise petite des crucifères

#### Description :

C'est un des insectes ravageurs du colza. Il mesure entre 2mm à 3mm de long. Il a une couleur bleu métallique parfois noir brillant, chaque élytre présente des bandes longitudinales jaunes.



**Altise petite des crucifères**

*Crédit Photo : ARVALIS – Institut du végétal*

#### Localisation géographique :

L'altise petite des crucifères est répandue dans toute la France. Les données dont nous disposons sur l'altise petite des crucifères concernent les cultures du colza d'hiver.

#### Dégâts :

Jusqu'au stade 2 feuilles, les morsures des altises petites sur les tigelles et les cotylédons peuvent provoquer la mort de la plantule.

Les morsures sont plus nuisibles sur le colza de printemps mais aussi sur les semis précoces d'automne lorsque la nouvelle génération se multiplie.

#### Facteurs favorables :

Facteurs climatiques : Leur activité est favorisée par un temps chaud et sec.

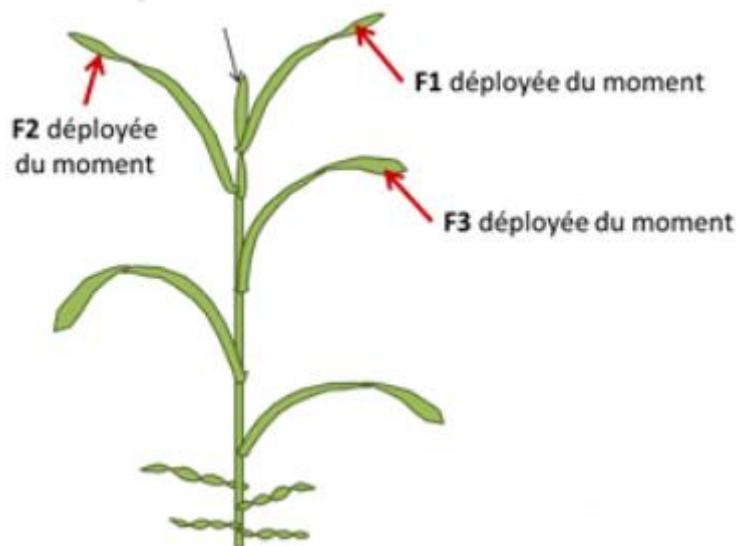
Facteurs agronomiques : Semi-précoces

## 7 Inventaire des bioagresseurs dans les parcelles

### 7.1 Méthode d'évaluation des maladies dans les parcelles

L'évaluation des maladies foliaires sur une plante repose sur l'observation de ses 3 dernières feuilles totalement sorties au moment de l'observation autrement dit les plus jeunes. Elles sont appelées F1, F2, F3 du moment.

- **F1** correspond à la dernière feuille sortie du moment (la plus jeune) ;
- **F2** correspond à celle qui est en-dessous de F1 ;
- **F3** correspond à celle qui est en-dessous de F2.



### Positionnement des feuilles F1, F2, F3 du moment sur une tige de céréale

*Crédit Photo : ARVALIS – Institut du végétal*

L'observation dans une parcelle a eu lieu chaque semaine, la méthodologie est la même pour toutes les maladies :

#### 7.1.1 Septoriose

Sur 20 plantes prélevées aléatoirement dans la parcelle

- Compter le nombre de feuilles F3 touchées

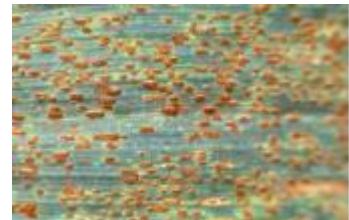


Symptôme de septoriose  
*Crédit Photo : ARVALIS – Institut du végétal*

#### 7.1.2 Rouille brune du blé

Sur 20 plantes prélevées aléatoirement dans la parcelle

- Compter le nombre de feuilles F3 touchées



Symptôme de rouille brune  
*Crédit Photo : ARVALIS – Institut du végétal*

#### 7.1.3 Le piétin-verse

Sur 20 tiges prélevées aléatoirement dans la parcelle

- Compter le nombre de tiges atteintes par la maladie



Symptôme du piétin verse  
*Crédit Photo : ARVALIS – Institut du végétal*

## 7.2 Technique de piégeage et méthode de comptage des insectes dans les parcelles

### 7.2.1 Piège coloré à liquide (Cuvette jaune)

Cette technique sert à capturer les insectes volants qui ont une attirance par la couleur. Ils vont venir se noyer dans le liquide au fond du piège. Le liquide est constitué de l'eau, du sel et du détergent. Le récipient est placé sur un piquet, est positionné à hauteur de la végétation.



Cuvette jaune : piégeage des insectes  
*Crédit Photo : ARVALIS – Institut du végétal*

### 7.2.2 Méthode de comptage des insectes dans les parcelles

Au niveau de chaque parcelle, une observation est faite chaque semaine. La méthode d'observation est la suivante :

- Disposer x pièges dans la parcelle ;
- Relever chaque semaine le nombre d'insectes dans chaque piège ;
- Ensuite faire la moyenne des nombres d'insectes relevés au niveau des pièges.

#### Remarque :

Les pièges restent fixes au cours de la campagne des observations.

### 7.3 Résumé des métriques utilisées pour chaque bioagresseur

Pour évaluer la présence des bioagresseurs dans une parcelle, diverses métriques sont utilisées au niveau des bases d'épidémiosurveillance. Pour les insectes on les quantifie en faisant la moyenne des nombres d'insectes dans les pièges, concernant les maladies on s'intéresse au nombre de feuilles F3 touchées ou de plantes touchées sur 20 plantes prélevées aléatoirement. La période d'observation parcellaire est de chaque semaine.

*Résumé des métriques utilisées pour chaque bioagresseur*

Bioagresseur	Type	Culture	Métrique	Période d'observation
Septoriose	maladie	Blé tendre d'hiver	Nbre de feuilles F3 touchées	Hebdomadaire
Rouille brune du blé	maladie	Blé tendre d'hiver	Nbre de feuilles F3 touchées	Hebdomadaire
Le puceron vert du pêcher	Insecte	Colza d'hiver	Moyenne des nbre d'insectes dans les pièges	Hebdomadaire
Piétin-versé	maladie	Blé tendre d'hiver	Nbre de tiges touchées	Hebdomadaire
Méligèthe du colza	Insecte	Colza d'hiver	Moyenne des nbre d'insectes dans les pièges	Hebdomadaire
Charançon du bourgeon terminal	Insecte	Colza d'hiver	Moyenne des nbre d'insectes dans les pièges	Hebdomadaire
Charançon de la tige du colza	Insecte	Colza d'hiver	Moyenne des nbre d'insectes dans les pièges	Hebdomadaire
Charançon de la tige du chou	Insecte	Colza d'hiver	Moyenne des nbre d'insectes dans les pièges	Hebdomadaire
Altise petite des crucifères	Insecte	Colza d'hiver	Moyenne des nbre d'insectes dans les pièges	Hebdomadaire

## 8 Calcul de la pression annuelle des bioagresseurs dans une parcelle

Pour quantifier la nuisibilité ou la présence des bioagresseurs, diverses métriques sont utilisées dans les bases d'épidémiosurveillance. Ces métriques sont évaluées, le plus souvent, à plusieurs reprises dans l'année. Pour obtenir une quantification unique de la pression pour l'ensemble de l'année qui soit similaire pour l'ensemble des bioagresseurs nous aggrégeons les observations comme définie précédemment (Delaune et al., 2021) et rappelé ci-dessous.

### 8.1 Seuil médian

Tout d'abord, chaque bioagresseur, est associé à un seuil médian défini comme suit :

**Seuil médian** : correspond à la médiane des observations de toutes les parcelles sur l'ensemble des années (2010 à 2020).

### 8.2 Pression annuelle des bioagresseurs d'une même espèce, dans une parcelle

La pression annuelle des bioagresseurs d'une même espèce, dans une parcelle est définie comme étant égal au nombre de fois où les valeurs de ses observations dépassent le seuil médian divisé par le nombre total de ses observations.

### Exemple de calcul de la pression annuelle des bioagresseurs d'une même espèce, dans une parcelle

Pour être plus explicite, ci-dessous est un exemple de calcul de pression annuelle du méligèthe du colza (insecte) dans une parcelle (ID parcelle = 391854) pour l'année 2016.

### *Suivi du méligrème du colza (bioagresseur) dans une parcelle*

ID_parcelle	Culture	Campagne	Bioagresseur	Observations	Seuil_median	Depasser	
1	391854	Colza d'hiver	2016	Méligème du colza	0	5	FALSE
2	391854	Colza d'hiver	2016	Méligème du colza	0	5	FALSE
3	391854	Colza d'hiver	2016	Méligème du colza	12	5	TRUE
4	391854	Colza d'hiver	2016	Méligème du colza	0	5	FALSE
5	391854	Colza d'hiver	2016	Méligème du colza	0	5	FALSE
6	391854	Colza d'hiver	2016	Méligème du colza	5	5	FALSE
7	391854	Colza d'hiver	2016	Méligème du colza	0	5	FALSE
8	391854	Colza d'hiver	2016	Méligème du colza	8	5	TRUE
9	391854	Colza d'hiver	2016	Méligème du colza	0	5	FALSE
10	391854	Colza d'hiver	2016	Méligème du colza	6	5	TRUE
11	391854	Colza d'hiver	2016	Méligème du colza	24	5	TRUE

Il y a eu 11 comptages (observations) liés au méligème du colza dans la parcelle parmi ces derniers 4 ont dépassé le seuil médian, d'après notre définition ci-dessus :

la pression annuelle du méligème de colza dans cette parcelle = 4/11

#### Remarque :

La pression des bioagresseurs d'une même espèce est un ratio compris entre 0 et 1. Plus elle est proche de 1 plus la présence du ravageur ou de la maladie est importante et on peut donc s'attendre à une baisse d'autant plus forte du rendement.

### **8.3 Utilité de calculer la pression des bioagresseurs dans une parcelle**

Connaitre la pression des bioagresseurs dans une parcelle donne une idée sur les dégâts qu'une culture peut subir. Elle permet de contrôler l'utilisation des produits phytosanitaires en indiquant si un traitement de la culture est nécessaire ou pas.

### **9 Localisation des parcelles liées aux bioagresseurs étudiés sur la carte de la France**

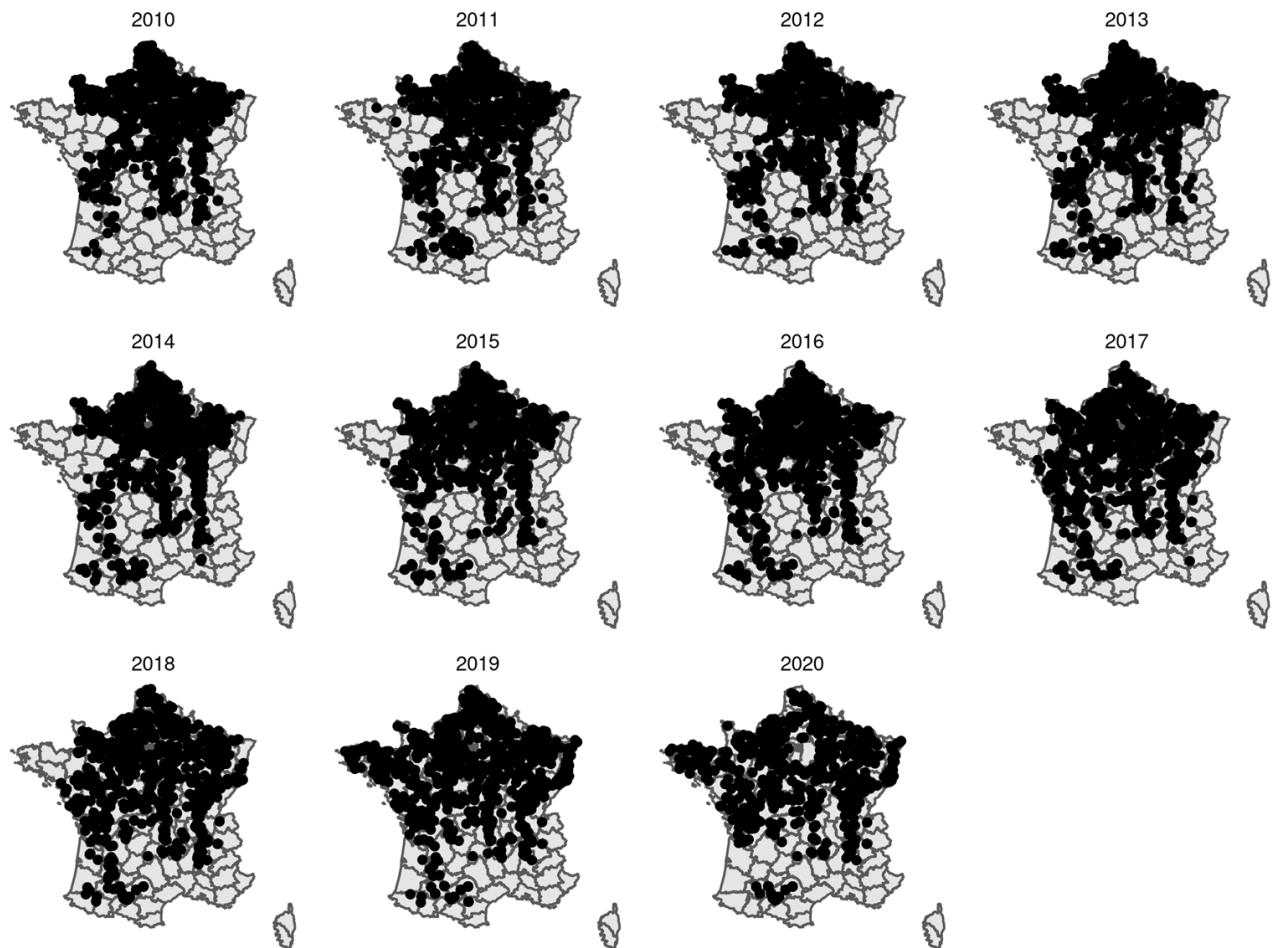
Les mêmes parcelles liées aux comptages des bioagresseurs ont été suivies quasiment sur une période de 10 ans. Certaines parcelles ont fait leur apparition au cours des années, quelques rares parcelles n'ont pas été suivies sur certaines années.

Grâce aux packages (frenchLandscape, spatDataManagement, mapview, sf) du programme RStudio, nous avons situé ces parcelles au niveau de la carte de la France.

Pour chaque bioagresseur nous avons représenté les parcelles associées au cours des années. Chaque point noir sur une carte représente une parcelle.

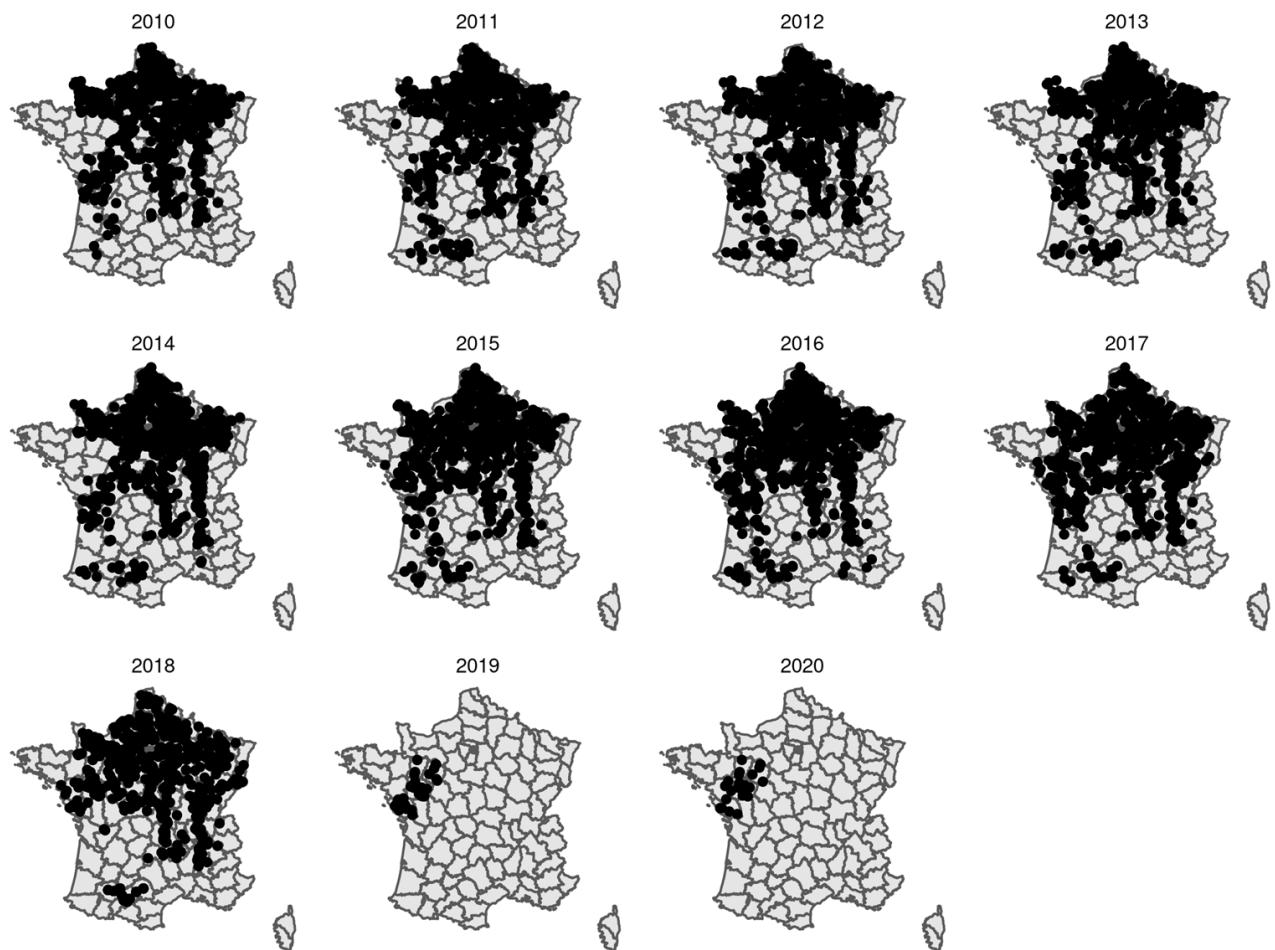
## 9.1 Répartition des parcelles liées à la septoriose au niveau de la France selon les années

*Répartition des parcelles liées à la septoriose au niveau de la France selon les années*



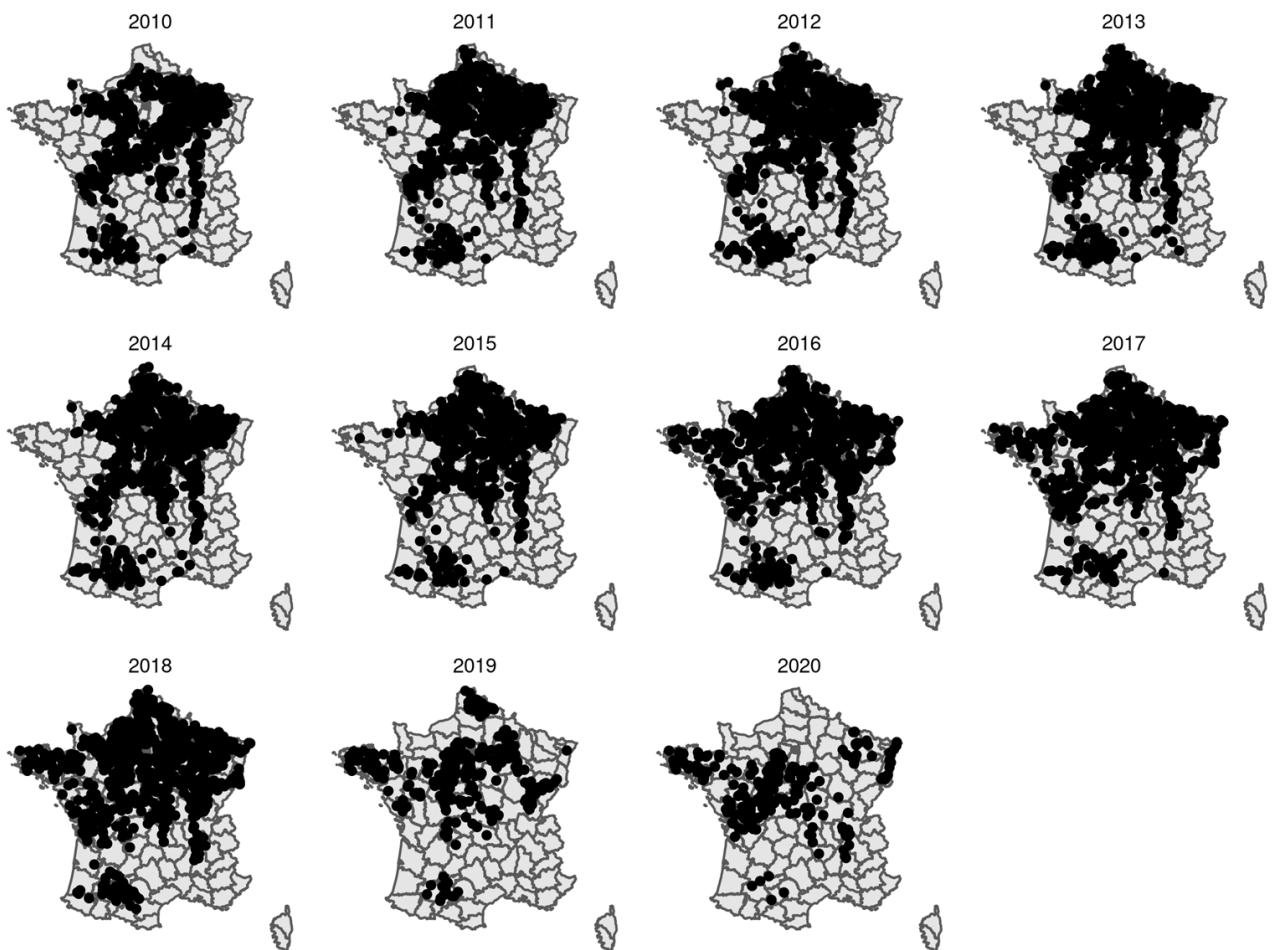
## 9.2 Répartition des parcelles liées à la rouille brune du blé au niveau de la France selon les années

*Répartition des parcelles liées à la rouille brune du blé au niveau de la France selon les années*



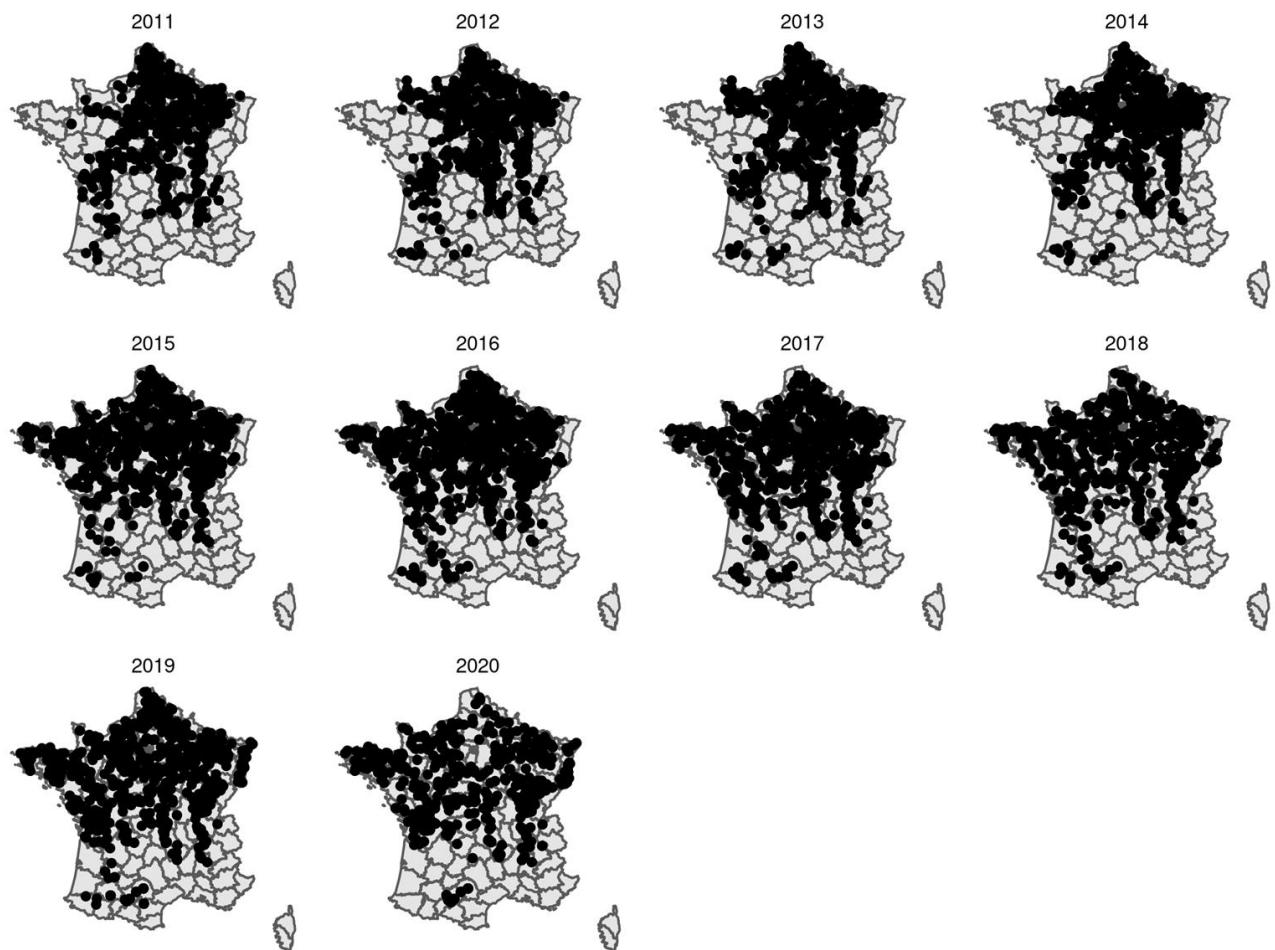
### 9.3 Répartition des parcelles liées au puceron vert du pêcher au niveau de la France selon les années

*Répartition des parcelles liées au puceron vert du pêcher au niveau de la France selon les années*



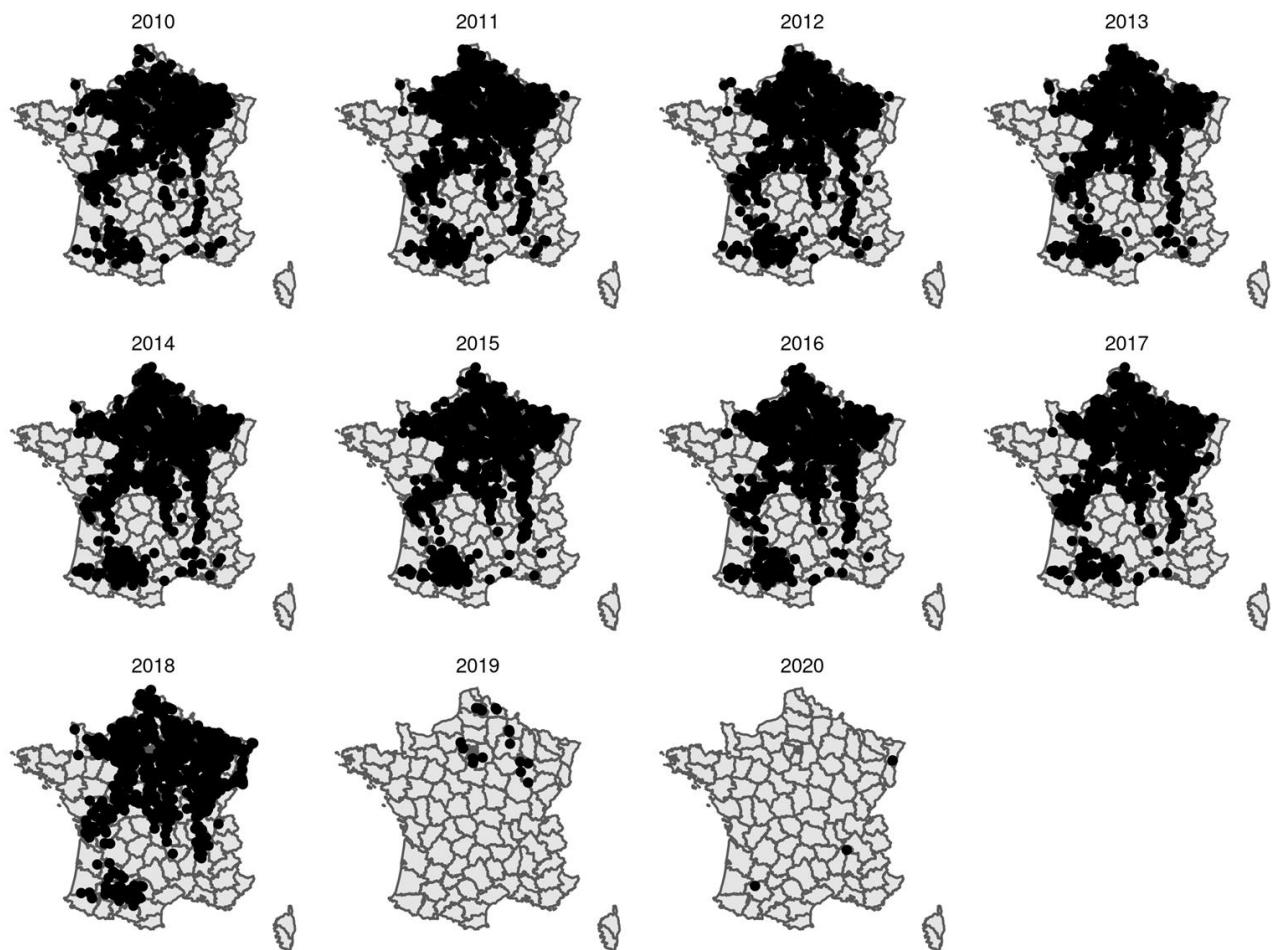
#### **9.4 Répartition des parcelles liées au piétin-verse au niveau de la France selon les années**

*Répartition des parcelles liées au piétin-verse au niveau de la France selon les années*



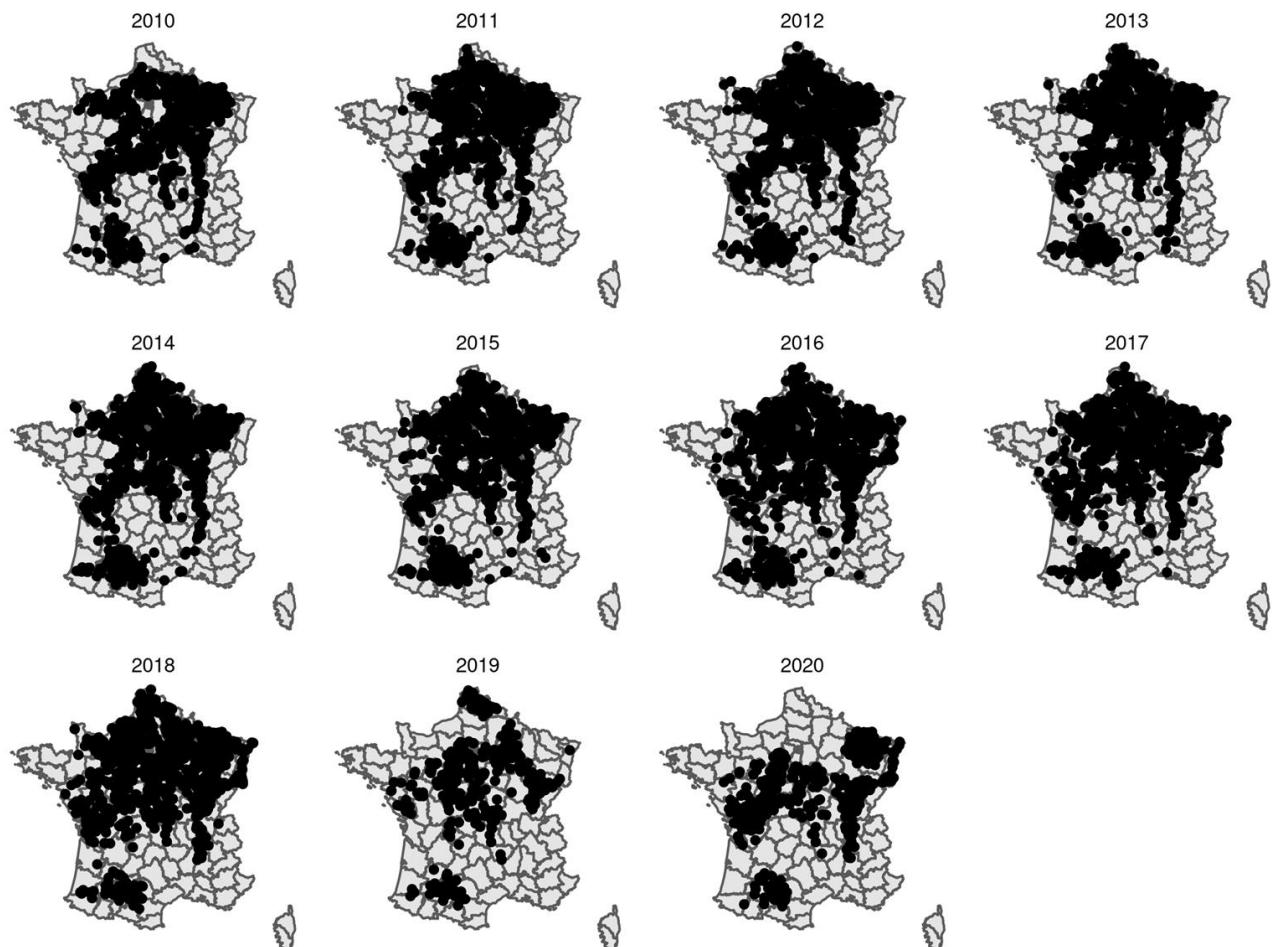
## 9.5 Répartition des parcelles liées au mélange du colza au niveau de la France selon les années

*Répartition des parcelles liées au mélange du colza au niveau de la France selon les années*



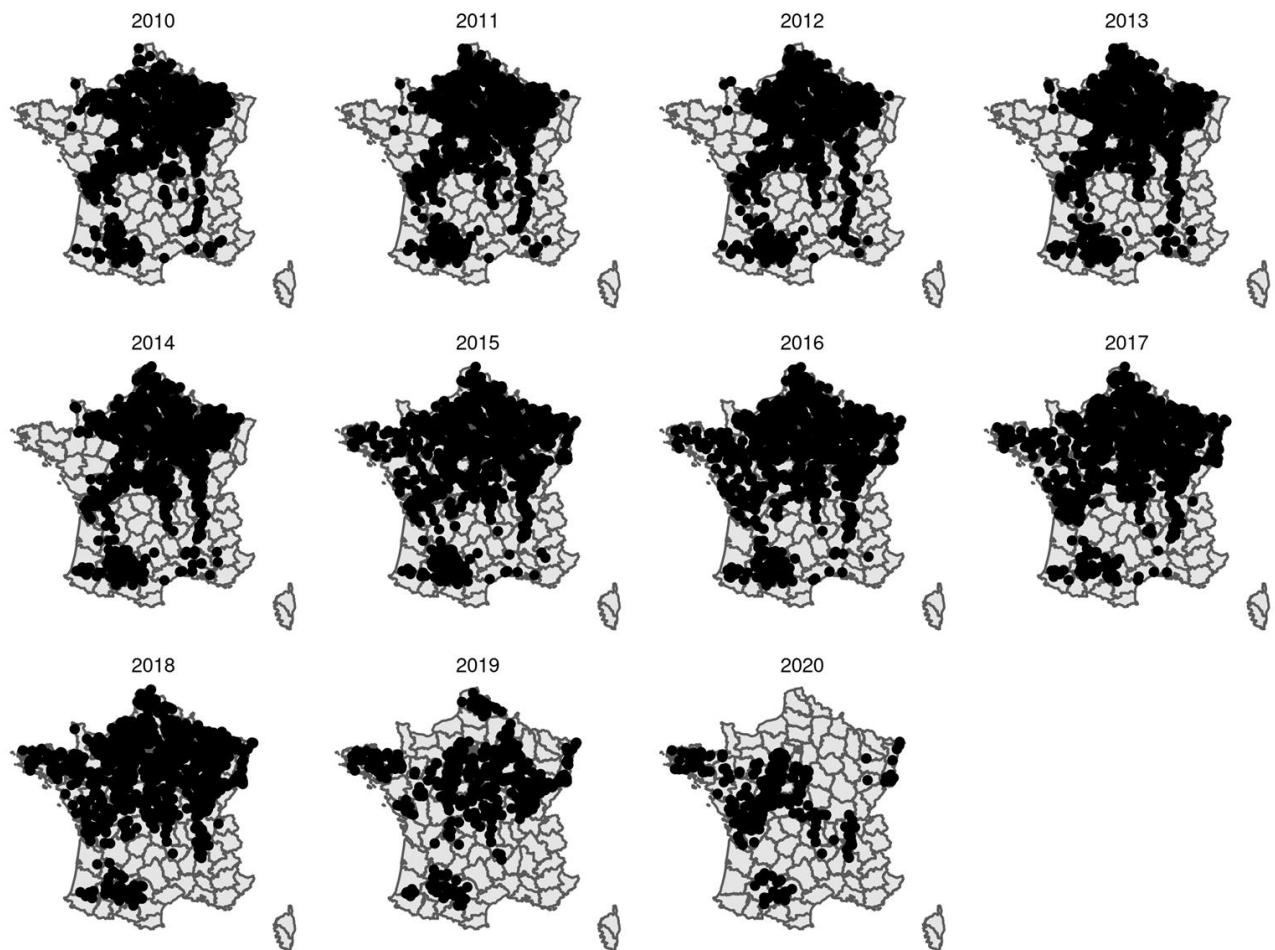
## 9.6 Répartition des parcelles liées au charançon du bourgeon terminal au niveau de la France selon les années

*Répartition des parcelles liées au charançon du bourgeon terminal au niveau de la France selon les années*



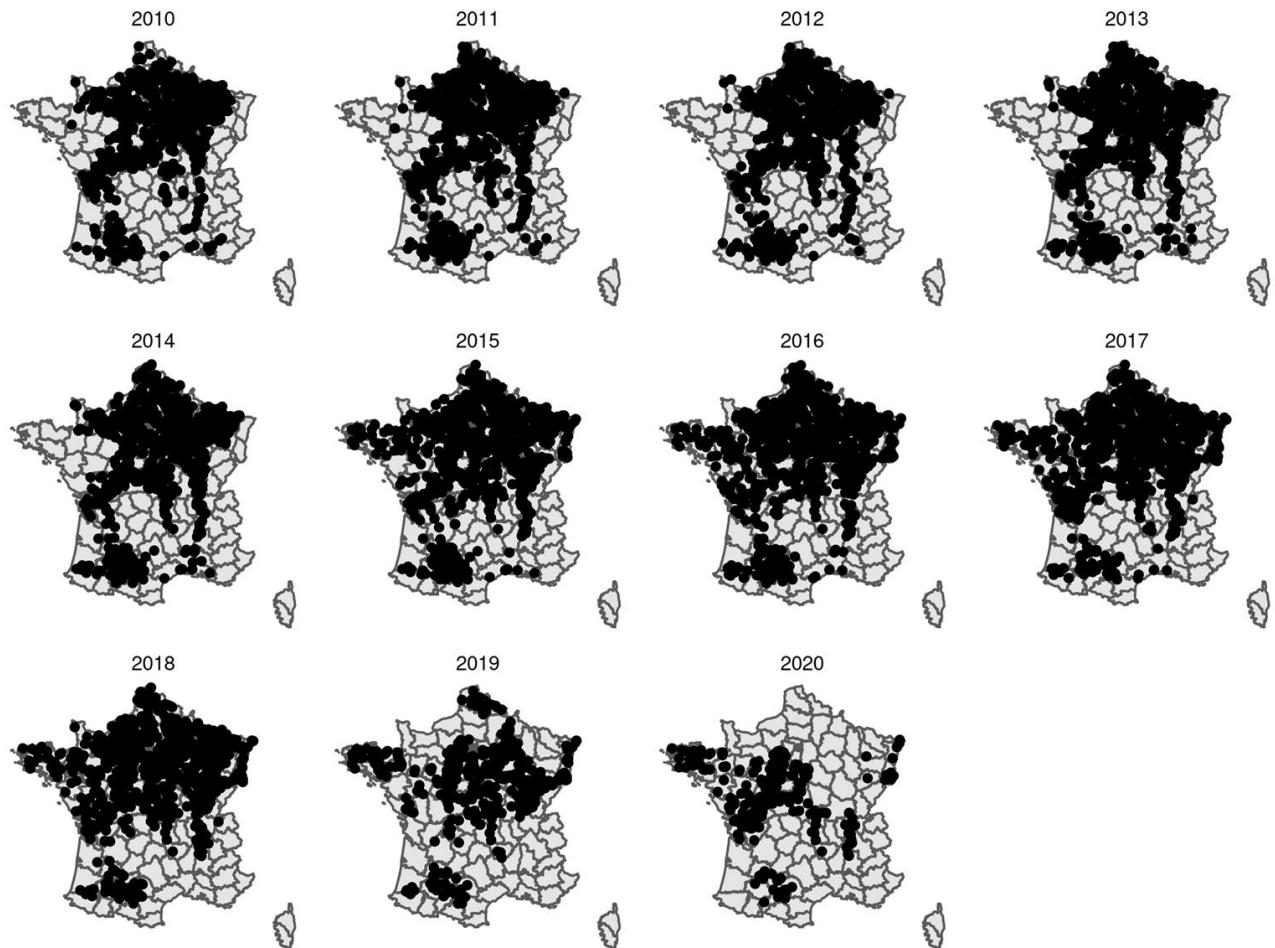
## 9.7 Répartition des parcelles liées au charançon de la tige du colza au niveau de la France selon les années

*Répartition des parcelles liées au charançon de la tige du colza au niveau de la France selon les années*



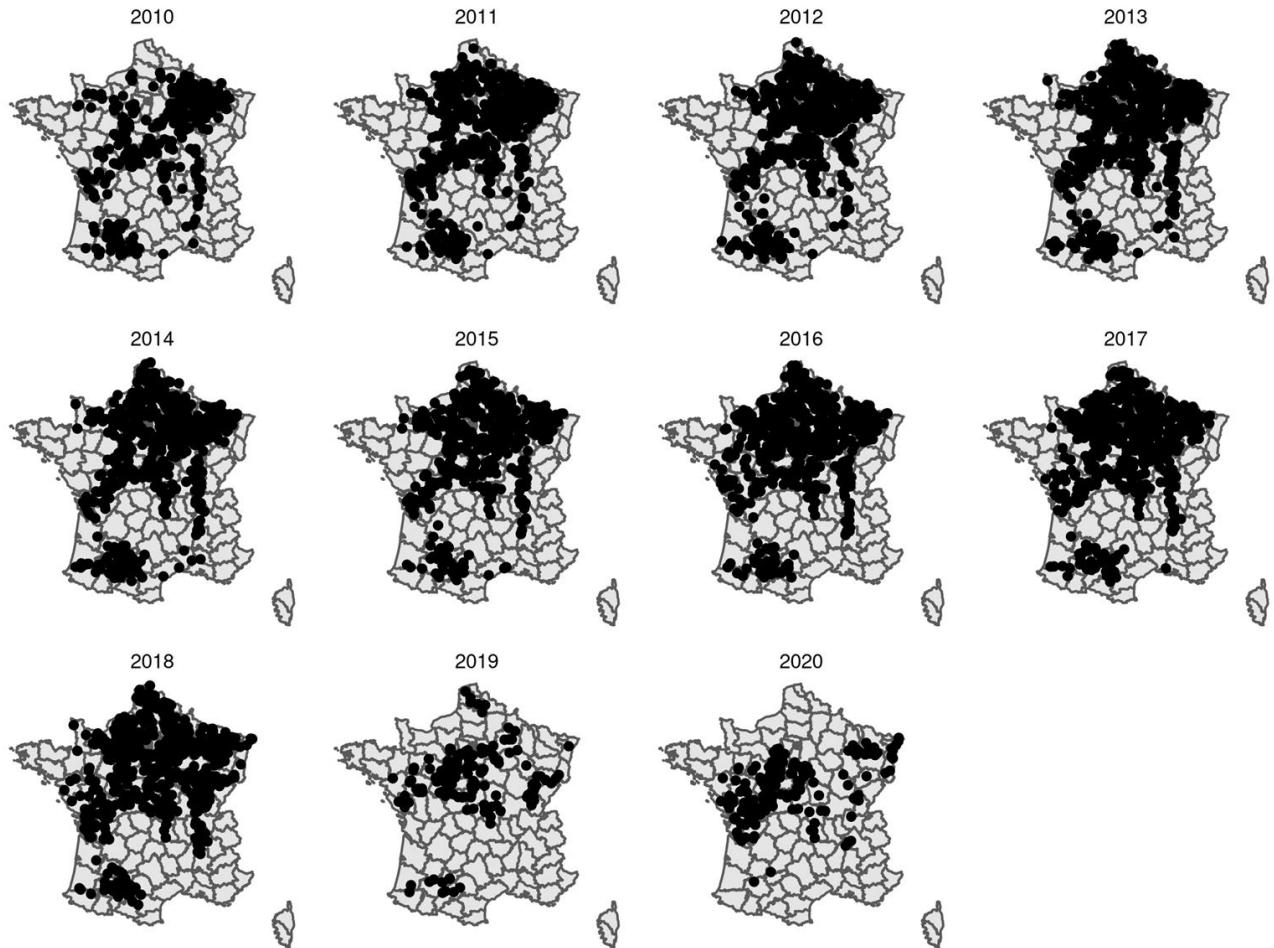
## 9.8 Répartition des parcelles liées au charançon de la tige du chou au niveau de la France selon les années

*Répartition des parcelles liées au charançon de la tige du chou au niveau de la France selon les années*



## 9.9 Répartition des parcelles liées à l'altise petite des crucifères au niveau de la France selon les années

*Répartition des parcelles liées à l'altise petite des crucifères au niveau de la France, selon les années*

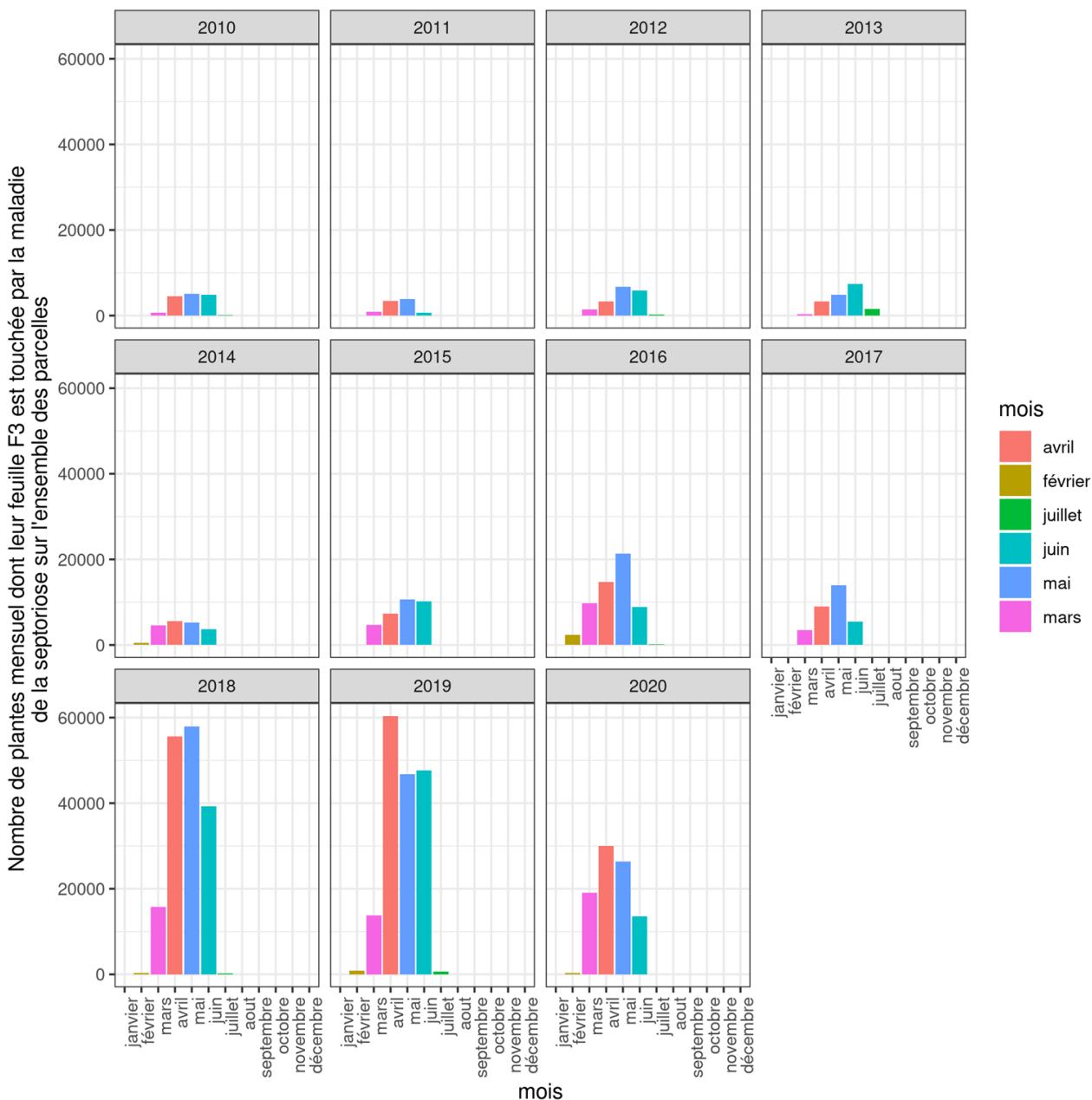


## 10 Distribution annuelle de chaque bioagresseur sur l'ensemble des parcelles

Avec l'aide des packages (ggplot2, dplyr, ...) du programme RStudio, on a représenté le nombre total mensuel de chaque insecte étudié ou le nombre de plantes touchées par une maladie, recensés sur toutes les parcelles associées, pour chaque année.

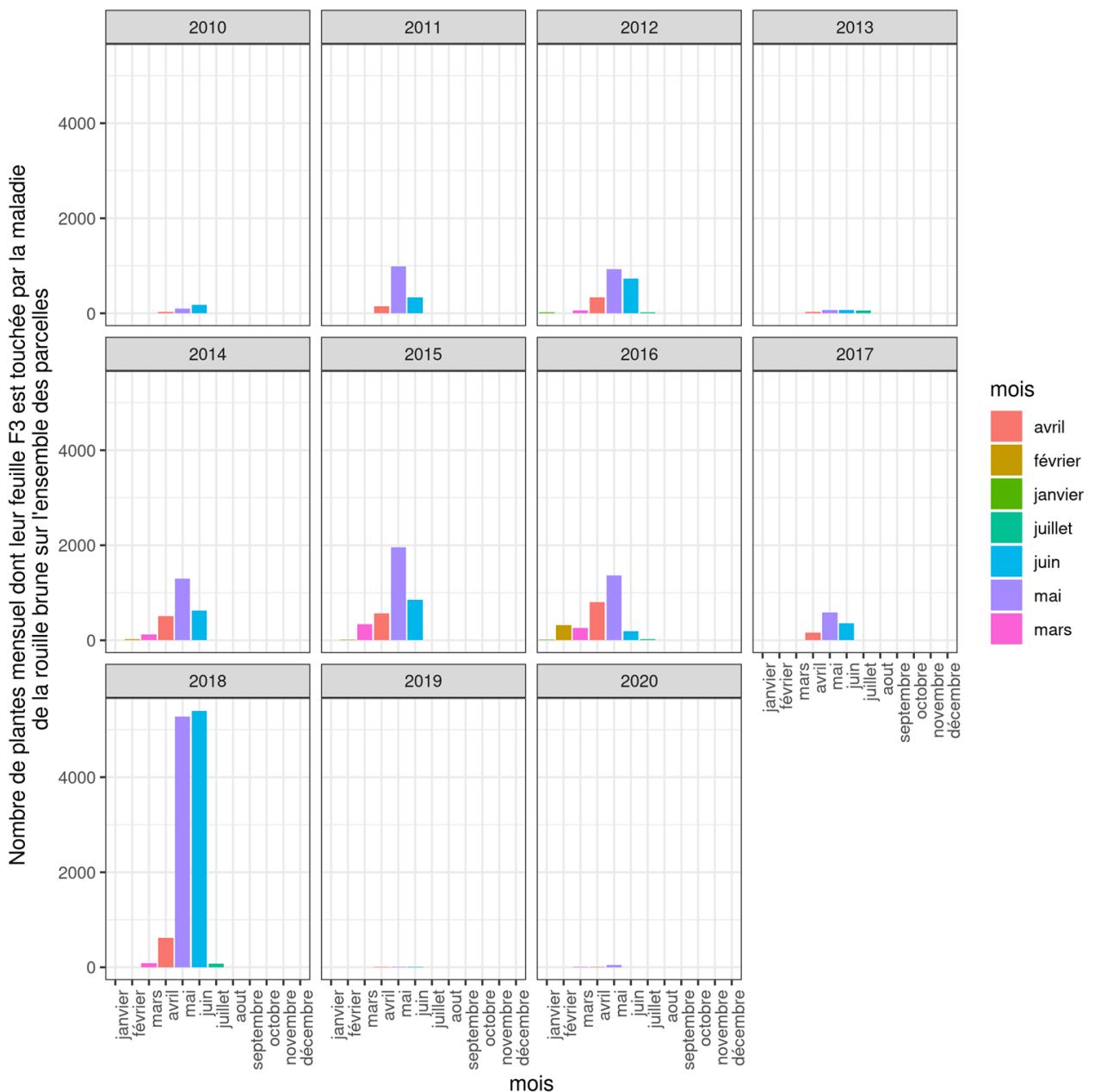
## **10.1 Nombre de plantes mensuel touchées par la maladie de la septoriose sur l'ensemble des parcelles**

*Nombre de plantes mensuel dont leur feuille F3 est touchée par la maladie de la septoriose sur l'ensemble des parcelles*



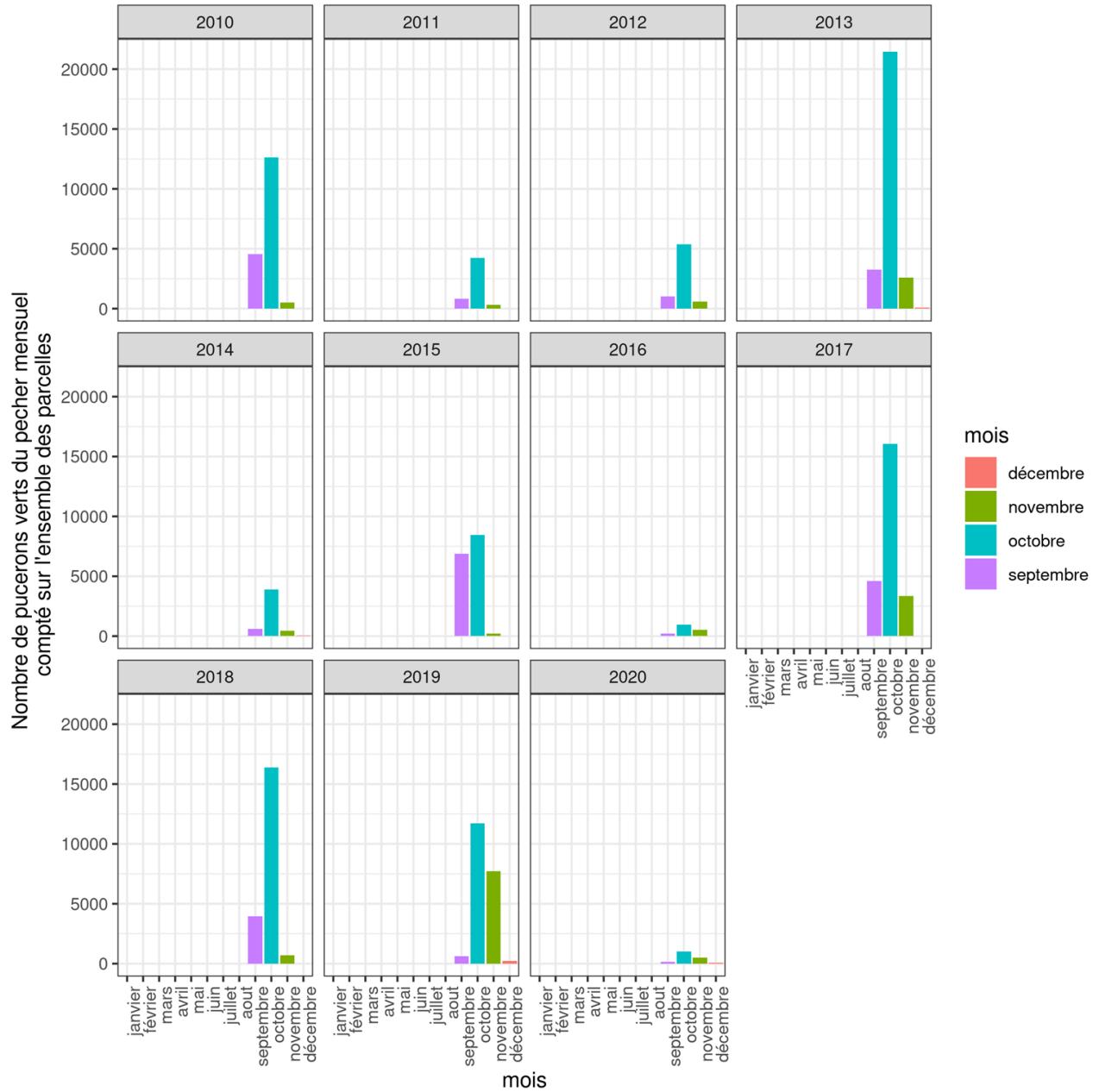
## **10.2 Nombre de plantes mensuel touchées par la maladie de la rouille brune sur l'ensemble des parcelles**

*Nombre de plantes mensuel dont leur feuille F3 est touchée par la maladie de la rouille brune  
sur l'ensemble des parcelles*



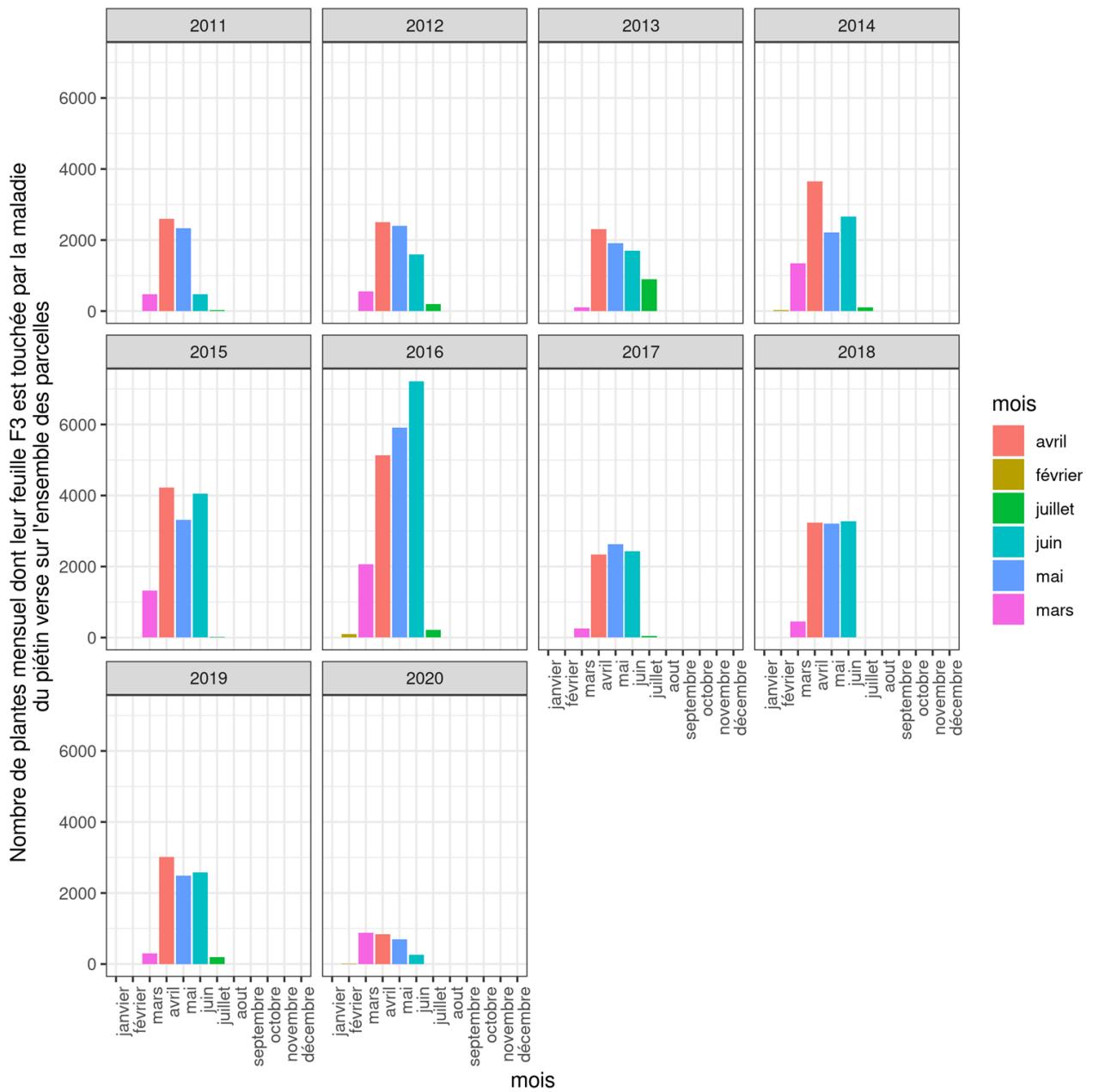
### 10.3 Nombre de pucerons verts du pêcher mensuel compté sur l'ensemble des parcelles

*Nombre de pucerons verts du pêcher mensuel compté sur l'ensemble des parcelles*



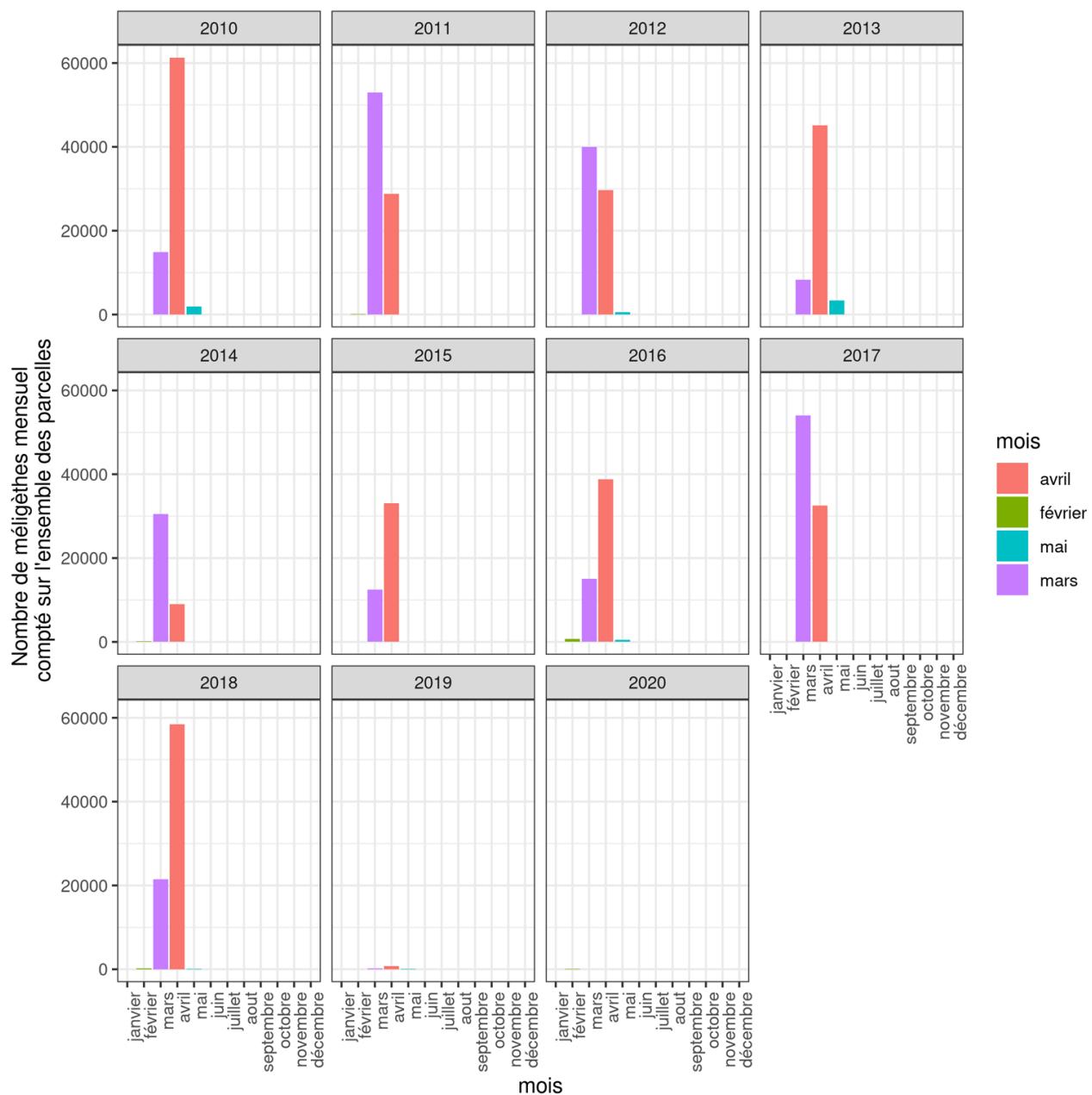
## 10.4 Nombre de plantes mensuel touchées par la maladie du piétin verse sur l'ensemble des parcelles

*Nombre de plantes mensuel dont leur feuille F3 est touchée par la maladie du piétin verse sur l'ensemble des parcelles*



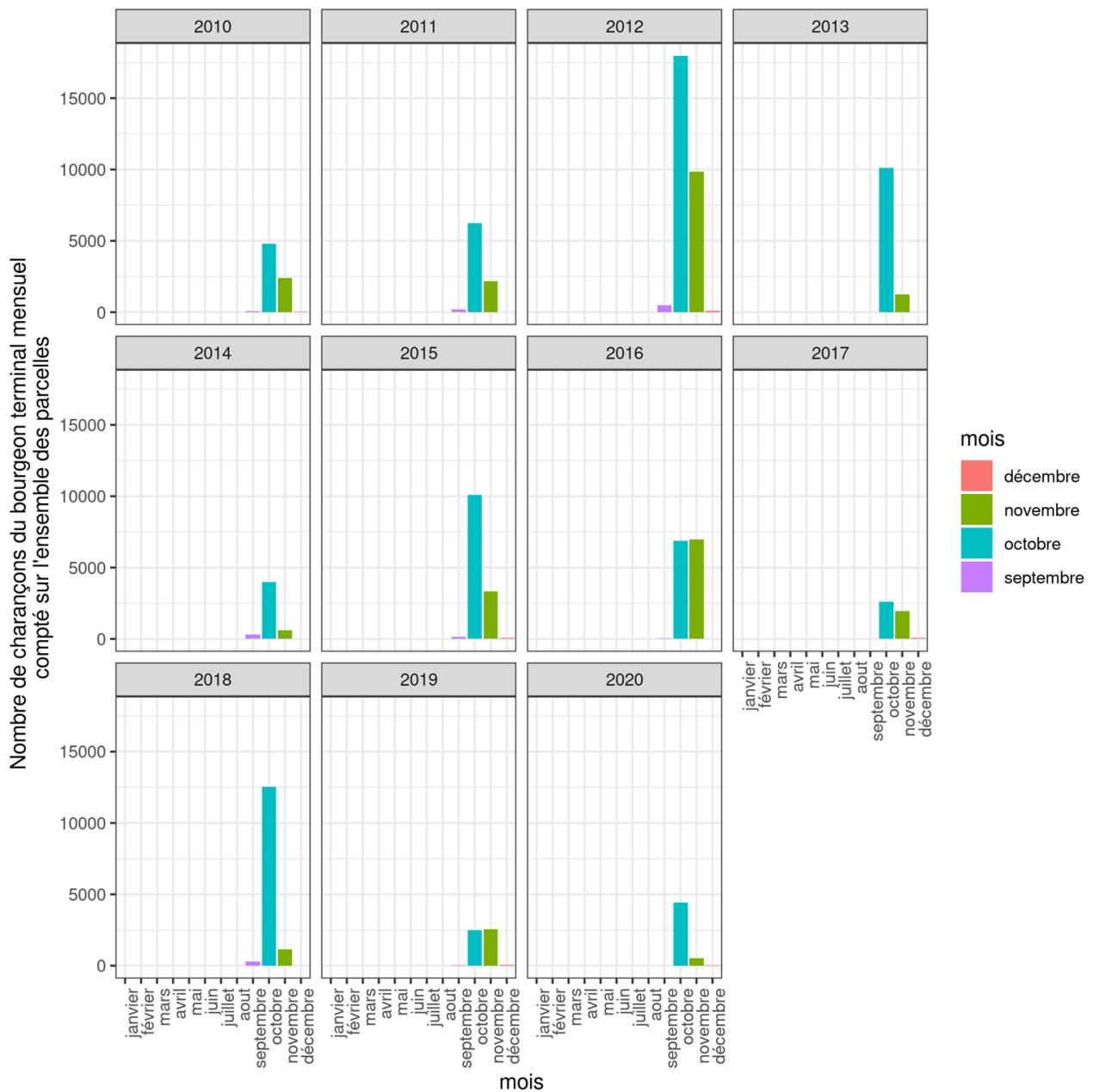
## 10.5 Nombre de mélègères mensuel compté sur l'ensemble des parcelles

*Nombre de mélègères mensuel compté sur l'ensemble des parcelles*



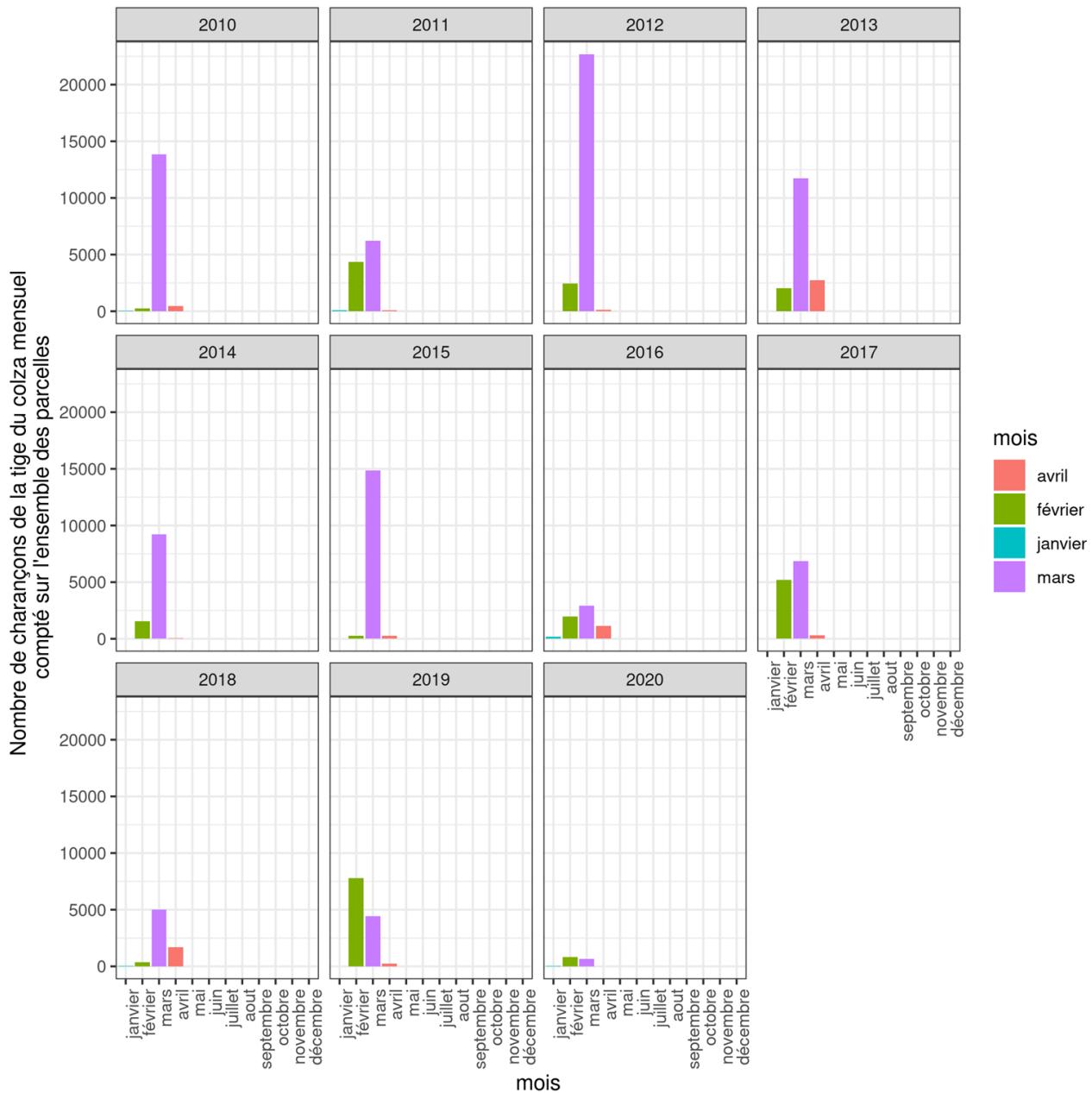
## 10.6 Nombre de charançons du bourgeon terminal mensuel compté sur l'ensemble des parcelles

*Nombre de charançons du bourgeon terminal mensuel compté sur l'ensemble des parcelles*



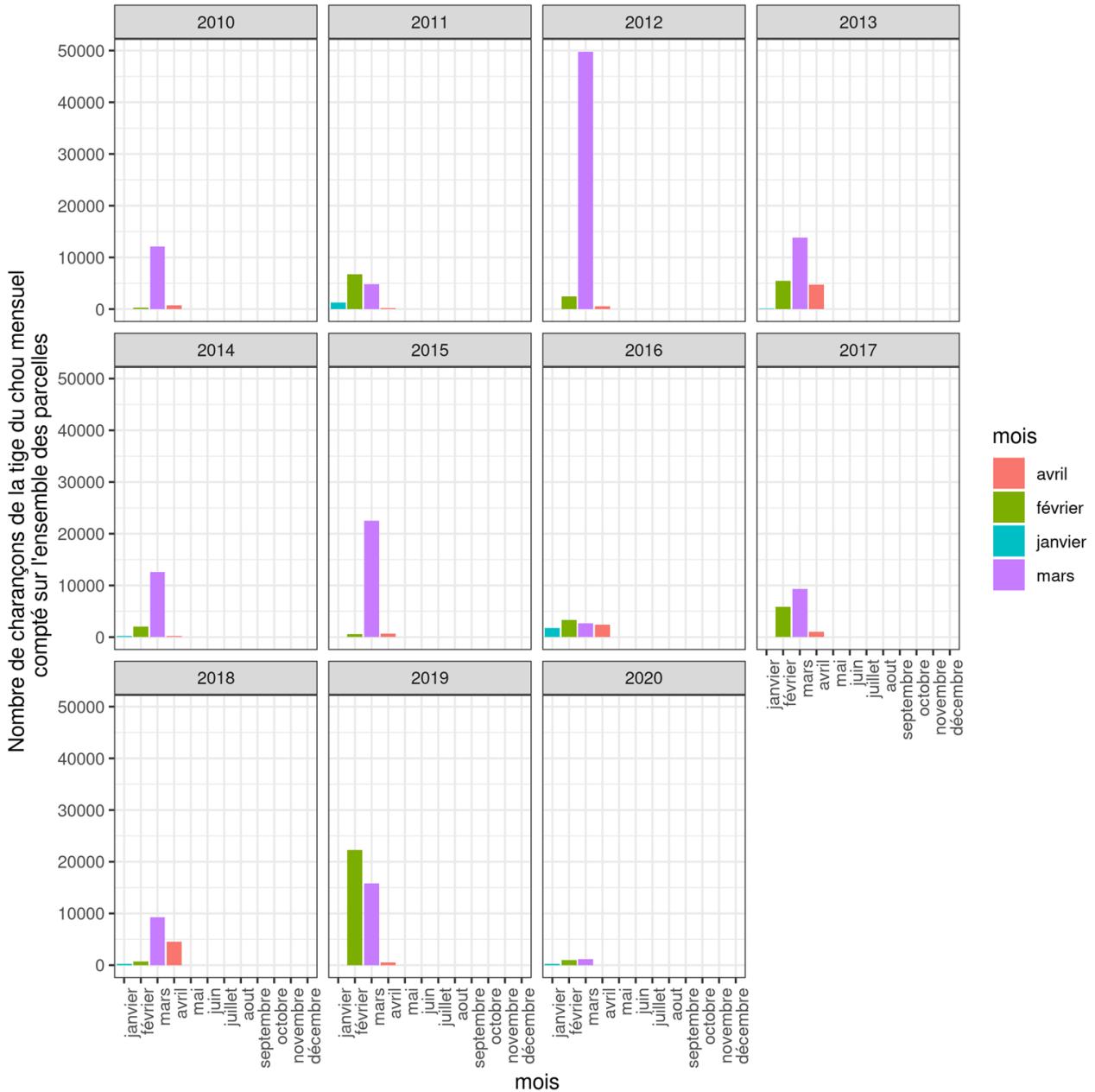
## **10.7 Nombre de charançons de la tige du colza mensuel compté sur l'ensemble des parcelles**

### *Nombre de charançons de la tige du colza mensuel compté sur l'ensemble des parcelles*



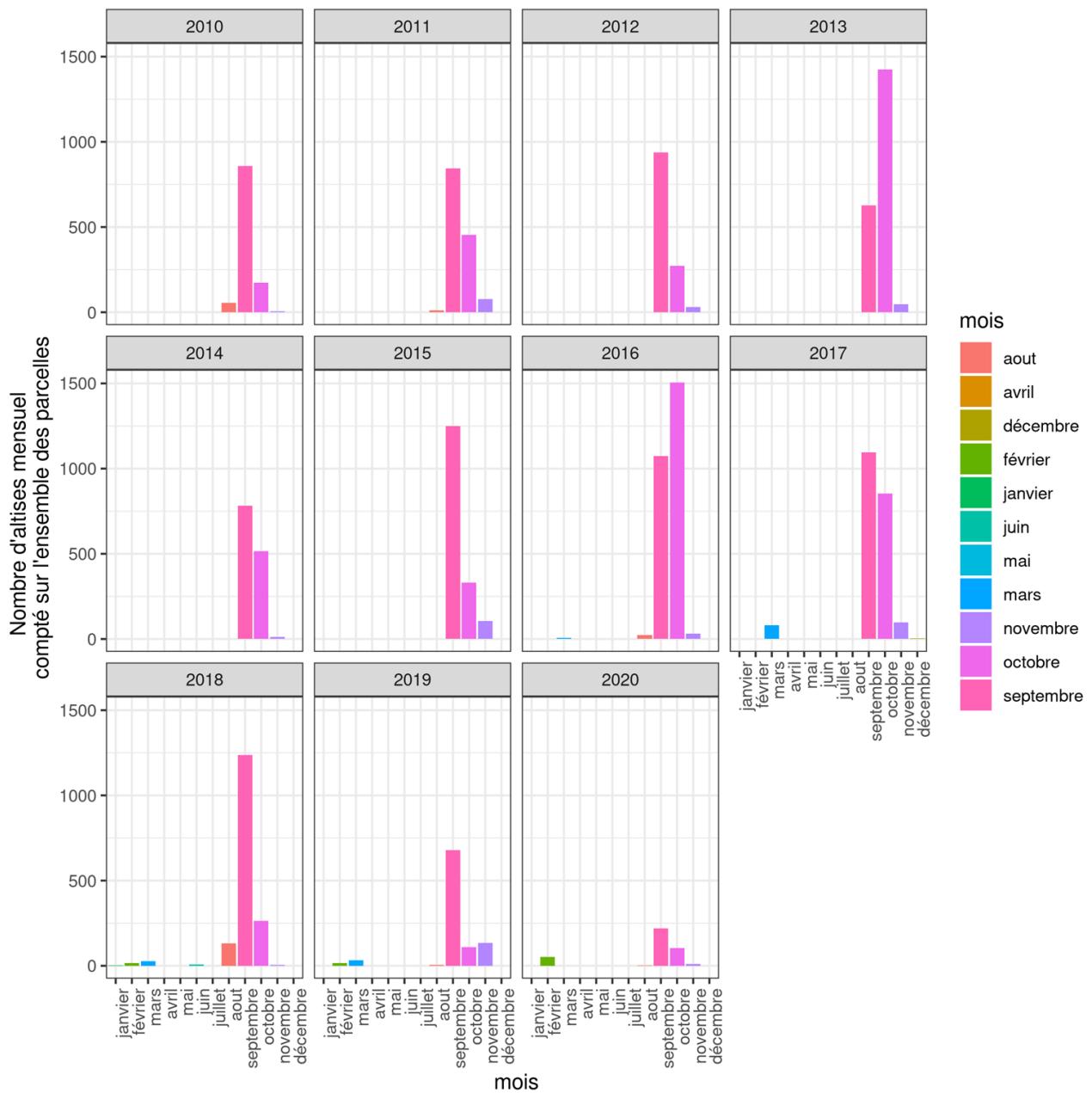
## **10.8 Nombre de charançons de la tige du chou mensuel compté sur l'ensemble des parcelles**

*Nombre de charançons de la tige du chou mensuel compté sur l'ensemble des parcelles*



## 10.9 Nombre d'altises mensuel compté sur l'ensemble des parcelles

*Nombre d'altises petites des crucifères mensuel compté sur l'ensemble des parcelles*



## 11 Pression annuelle des bioagresseurs d'espèce $i$ prédicté dans une parcelle

### 11.1 Pression annuelle des bioagresseurs d'espèce $i$ , prédicté en utilisant la moyenne pondérée

Dans le projet MoCoRiBA nous étudions l'impact de la présence des bioagresseurs sur le rendement. Cependant, les données sur les bioagresseurs ne sont pas présentes dans les parcelles où nous avons les rendements et les pratiques des agriculteurs. Pour étudier l'impact des bioagresseurs sur le rendement nous cherchons donc un modèle d'interpolation performant pour estimer la pression des bioagresseurs d'espèce  $i$  dans une parcelle d'intérêt.

Supposons que nous cherchons à prédire la pression des bioagresseurs d'espèce  $i$ , dans une parcelle nommée  $p$ , qui est dans un ensemble de  $m+1$  parcelles. Pour les  $m$  autres parcelles, nous

connaissons chacune leur pression des bioagresseurs d'espèce  $i$ , notée  $P_{p'}^{réelle}(bio\_i)$ , pour  $p' = 1, \dots, m$

La pression des bioagresseurs d'espèce  $i$ , prédite dans la parcelle  $p$ , notée  $P_p^{prédictive}(bio\_i)$  est donnée par :

$$P_p^{prédictive}(bio\_i) = \frac{\sum_{p'=1}^m e^{\frac{-d_{pp'}}{h}} P_{p'}^{réelle}(bio\_i)}{\sum_{p'=1}^m e^{\frac{-d_{pp'}}{h}}} \quad (11.1)$$

où :

$d_{pp'}$  : distance qui sépare la parcelle  $p$  et la parcelle  $p'$  avec  $1 \leq p' \leq m$ ;

$h$  : hyperparamètre.

$P_p^{prédictive}(bio\_i)$  correspond à une moyenne des  $P_{p'}^{réelle}(bio\_i)$ , pondérée par l'exponentielle de l'opposé de la distance entre les parcelles  $p$  et  $p'$ . Plus une parcelle  $p'$  est proche de la parcelle  $p$ , plus elle a un poids élevé. C'est normal parce que les parcelles les plus proches devront être les plus représentatives.

### Recherche de l'hyperparamètre $h$ optimal des bioagresseurs d'espèce $i$

Le but de notre modèle d'interpolation (11.1) est de pouvoir l'utiliser pour prédire la pression des bioagresseurs d'espèce  $i$  dans une parcelle dont nous n'avons pas de données. Pour cela nous devons trouver l'hyperparamètre optimal  $h$  (donne une meilleure performance de prédiction) associé à chaque espèce de bioagresseur.

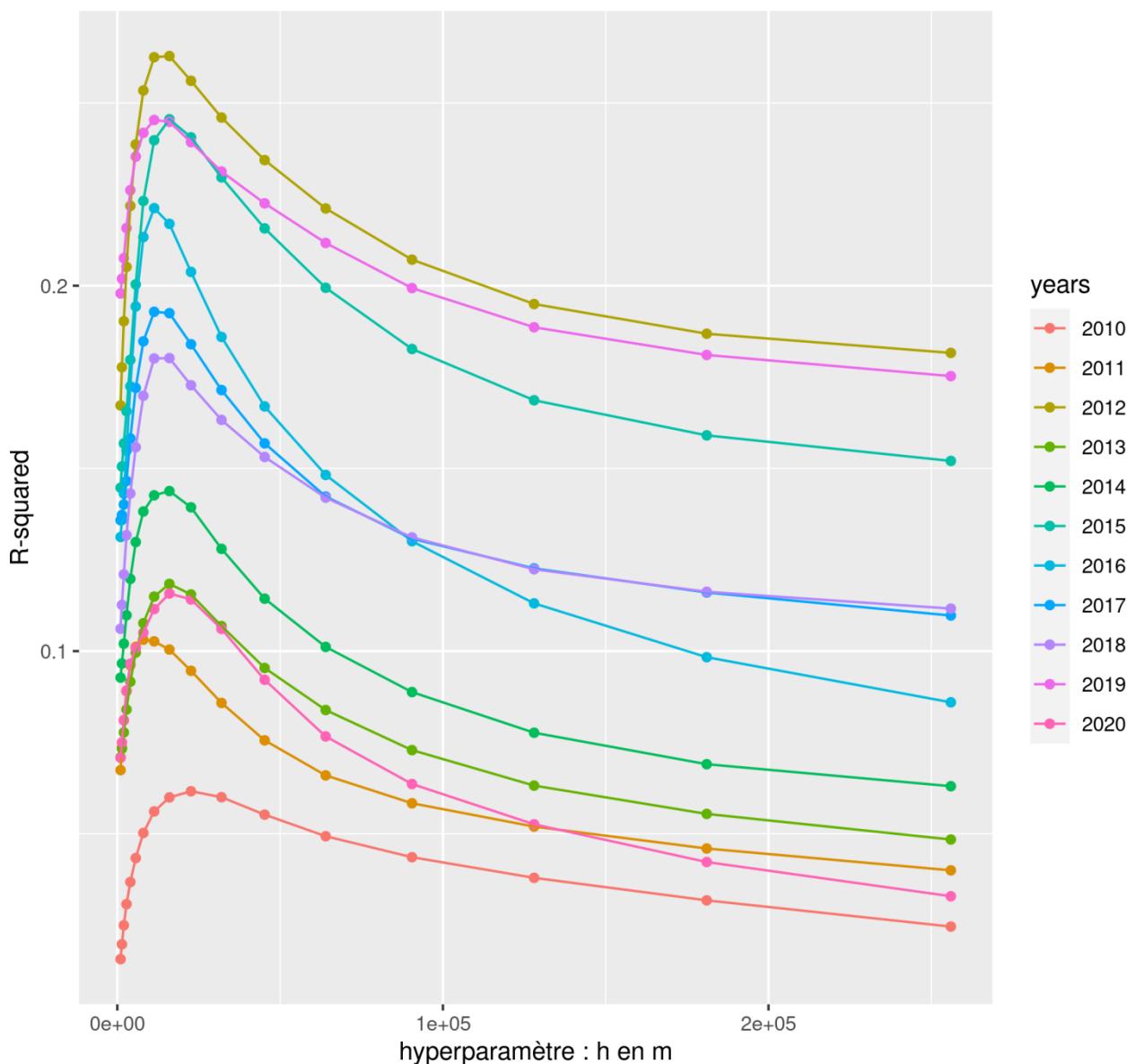
#### Procédure de recherche de l'hyperparamètre $h$ optimal :

- Varier l'hyperparamètre  $h$  dans un intervalle ;
- Pour chaque valeur de  $h$ , prédire la pression des bioagresseurs d'espèce  $i$ ,  $P_p^{prédictive}(bio\_i)$  au niveau de chaque parcelle ;
- Calculer le coefficient de détermination  $R^2$  entre les pressions réelles  $P_{p'}^{réelle}(bio\_i)$  et les pressions prédites  $P_p^{prédictive}(bio\_i)$  correspondant à l'hyperparamètre  $h$ .

Chaque graphique parmi les suivants représente le coefficient de détermination  $R^2$  entre les pressions réelles  $P_{p'}^{réelle}(bio\_i)$  et les pressions prédites  $P_p^{prédictive}(bio\_i)$  par le modèle (11.1), en fonction de l'hyperparamètre  $h$ .

Le modèle est appliqué séparément sur les données des années (2010 à 2020) avec les mêmes hyperparamètres. Ensuite, on a calculé le  $R^2$  correspondant aux observations d'une même année prédites avec le même hyperparamètre.

### Coefficient de détermination ( $R^2$ ) en fonction de l'hyperparamètre $h$ (cas de la septoriose)



Le tableau suivant indique les hyperparamètres optimaux selon les années et les coefficients de détermination  $R^2$  optimaux (max des courbes) correspondants.

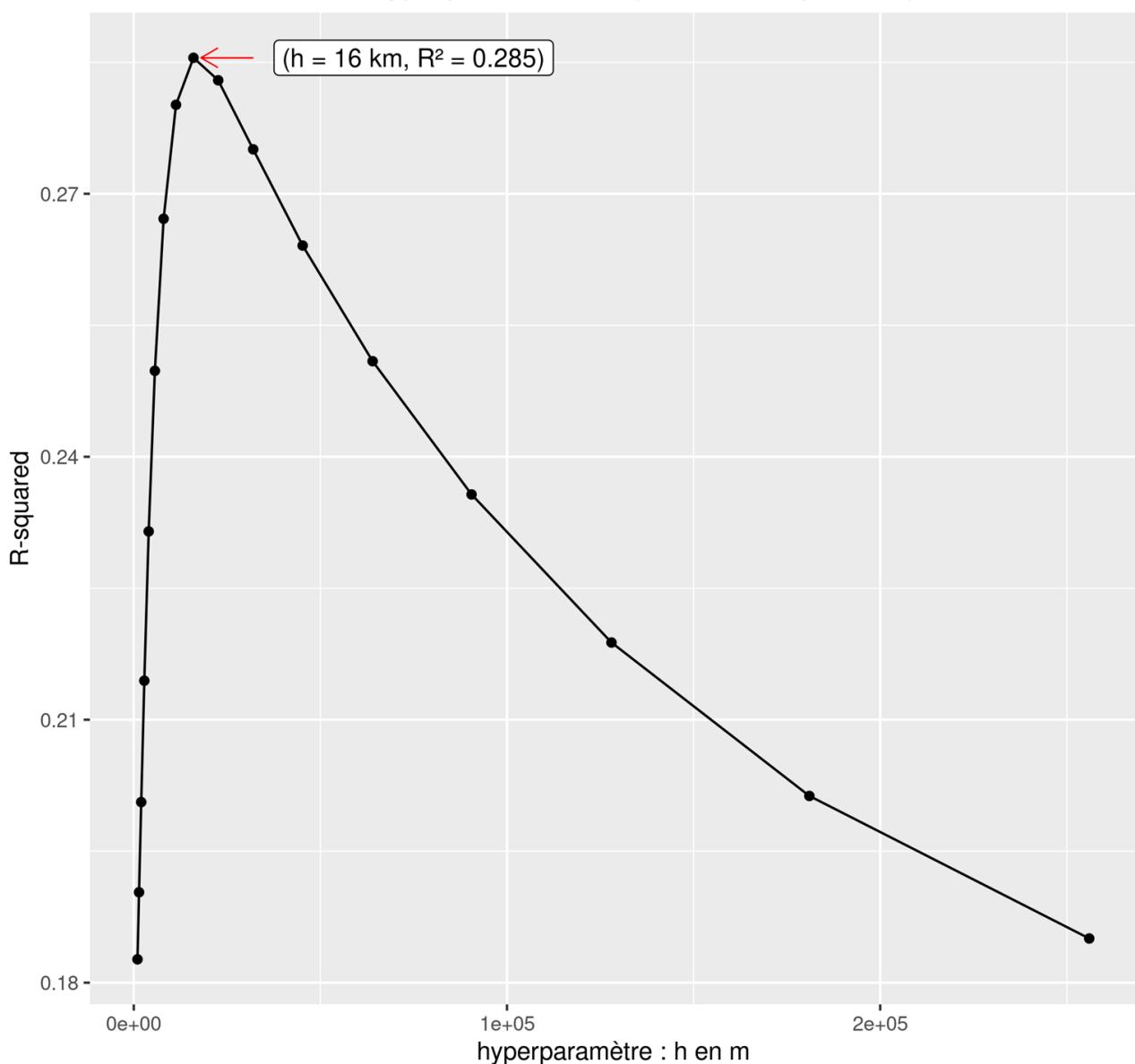
**Tableau 11.1 :** Hyperparamètres optimaux ( $h$ ) et coefficients de détermination optimaux ( $R^2$ ) selon les années

<b>Années</b>	<b>Hyperparamètres optimaux (h)</b>	<b>Coefficients de détermination optimaux (<math>R^2</math>)</b>
2010	$h = 22.627 \text{ km}$	$R^2 = 0.0617$
2011	$h = 8 \text{ km}$	$R^2 = 0.103$
2012	$h = 16 \text{ km}$	$R^2 = 0.263$
2013	$h = 16 \text{ km}$	$R^2 = 0.118$
2014	$h = 16 \text{ km}$	$R^2 = 0.144$
2015	$h = 16 \text{ km}$	$R^2 = 0.246$
2016	$h = 11.314 \text{ km}$	$R^2 = 0.221$
2017	$h = 11.314 \text{ km}$	$R^2 = 0.193$
2018	$h = 16 \text{ km}$	$R^2 = 0.18$
2019	$h = 11.314 \text{ km}$	$R^2 = 0.245$
2020	$h = 20 \text{ km}$	$R^2 = 0.116$

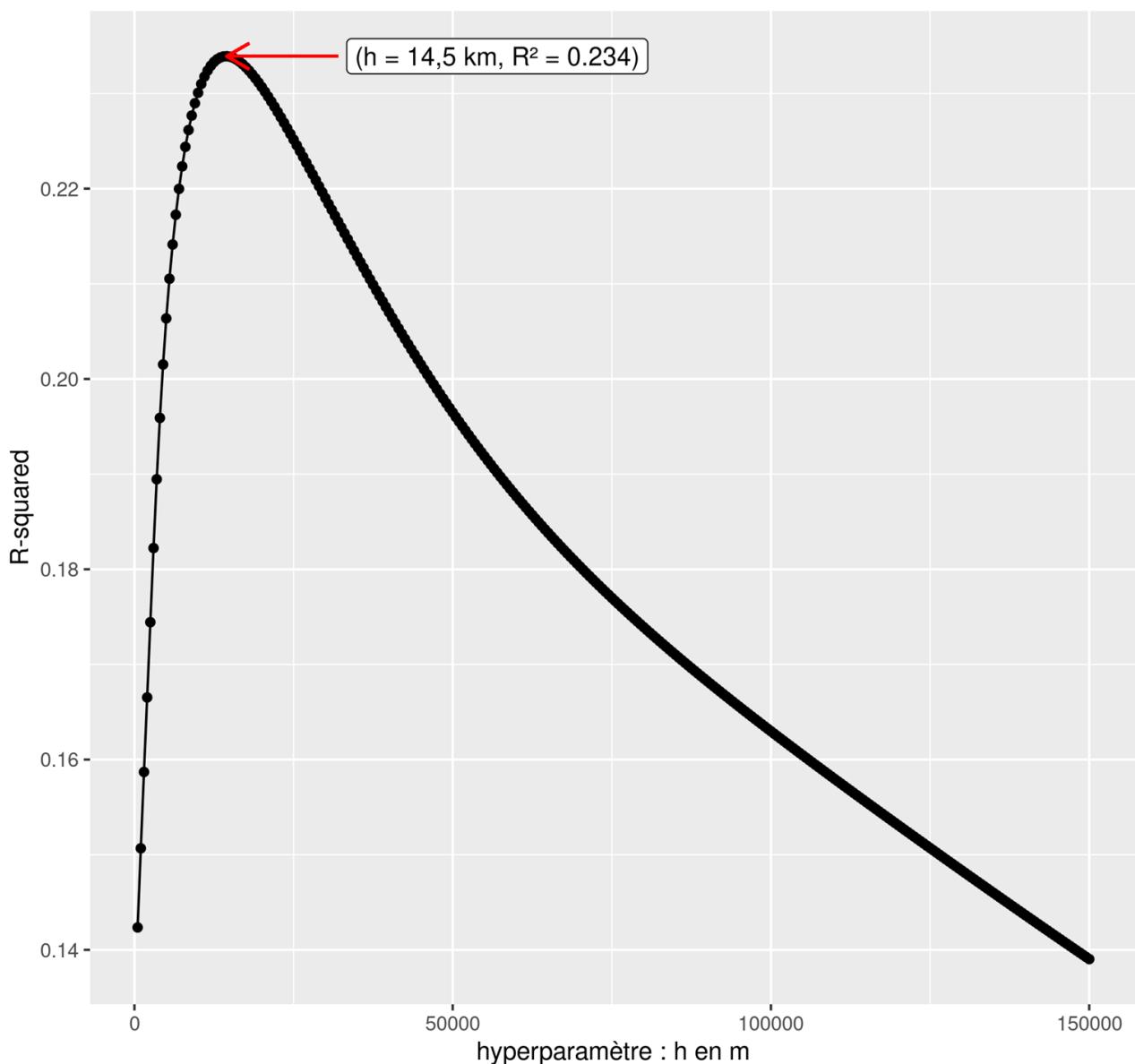
Comme précédemment, on a utilisé les mêmes hyperparamètres  $h$  pour entraîner séparément les données de chaque année (2010 à 2020). Mais cette fois-ci, on a calculé le  $R^2$  correspondant aux observations des années (2010 à 2020) prédictes avec le même hyperparamètre : celui qui donne le meilleur  $R^2$  sur l'ensemble des points-années.

Les graphiques suivants représentent les résultats obtenus sur chaque bioagresseur étudié :

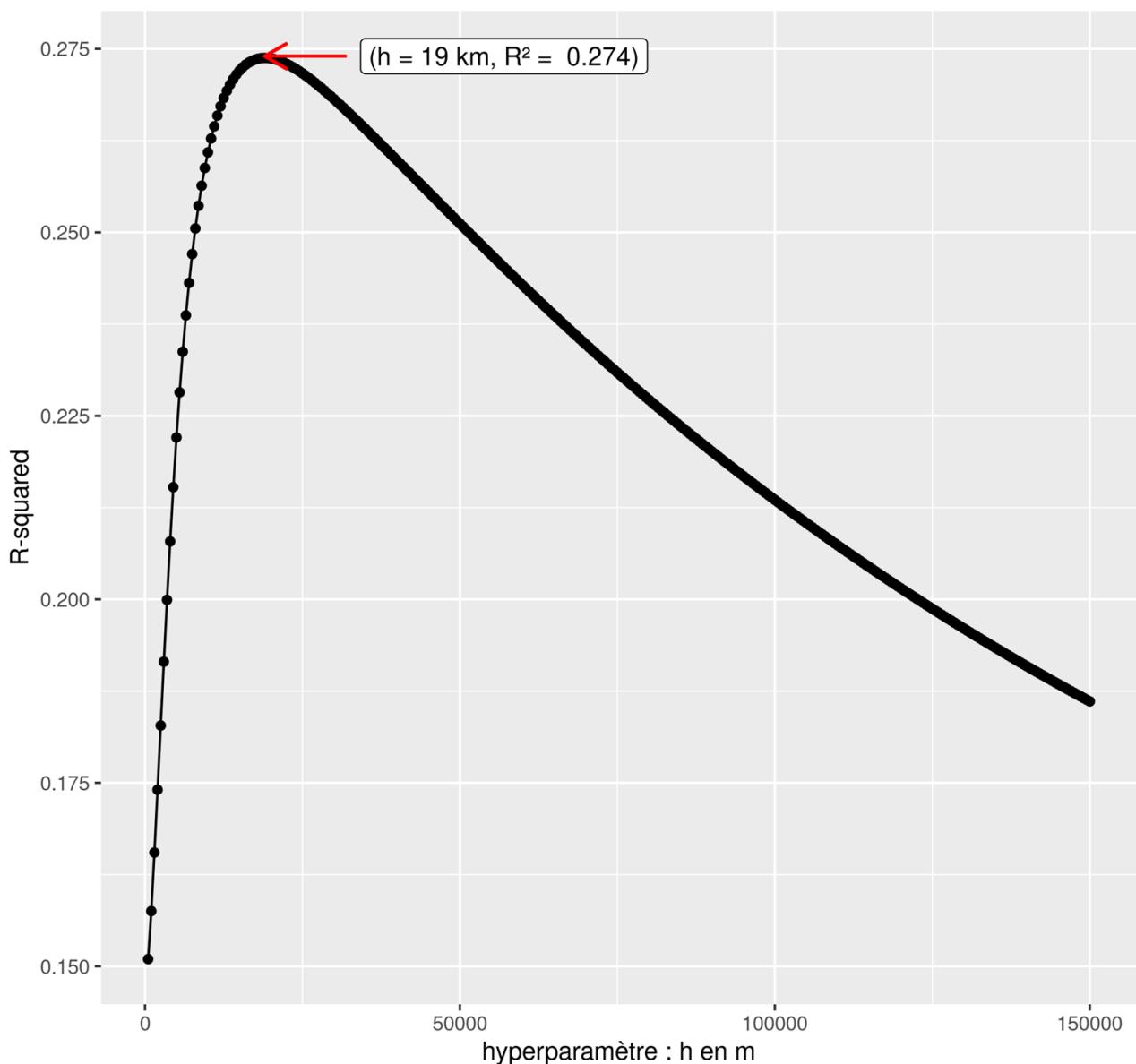
Coefficient de détermination ( $R^2$ ) en fonction  
de l'hyperparamètre  $h$  (cas de la septoriose)



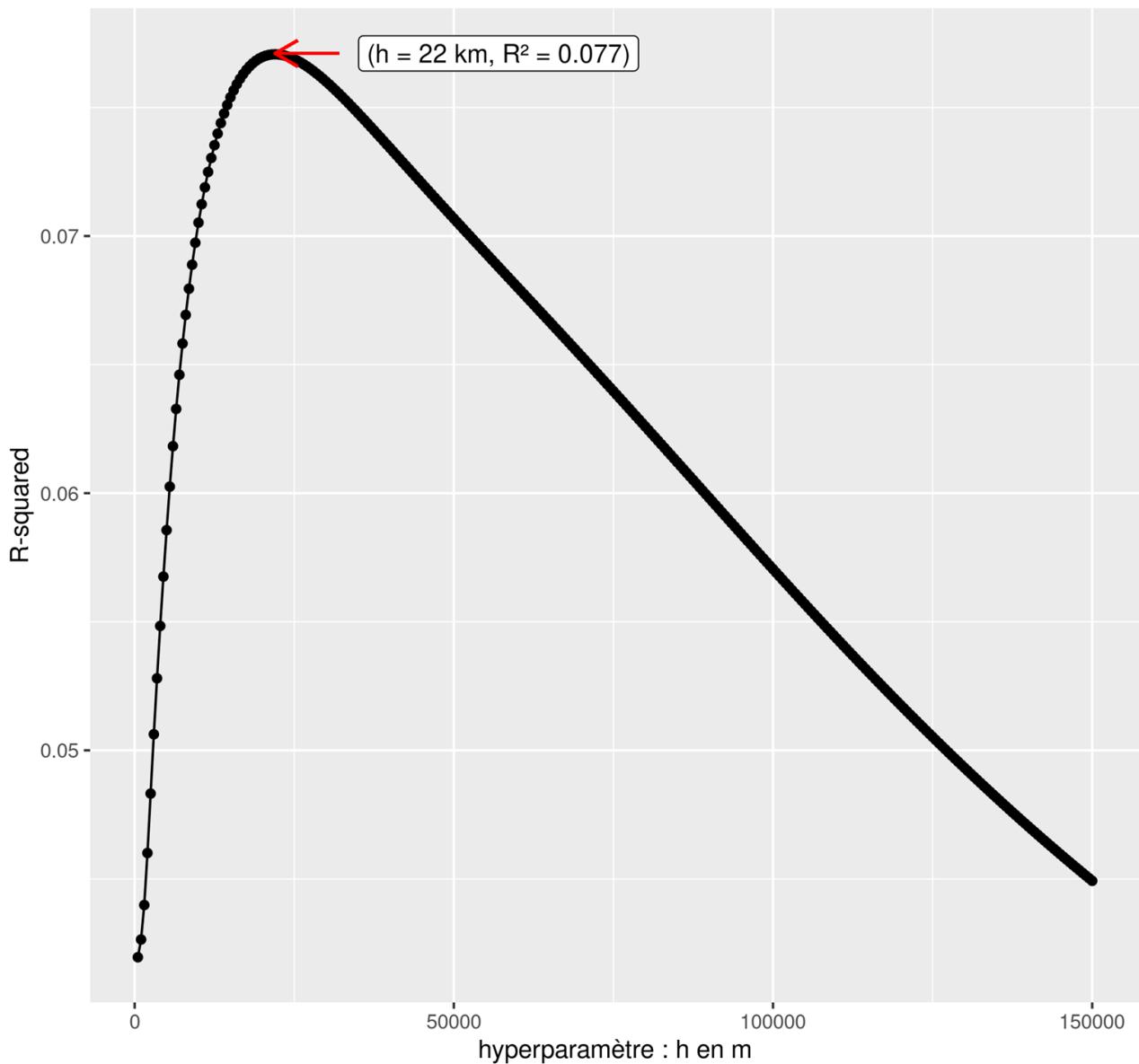
Coefficient de détermination ( $R^2$ ) en fonction  
de l'hyperparamètre  $h$  (cas de la Rouille brune du blé (RBF3))



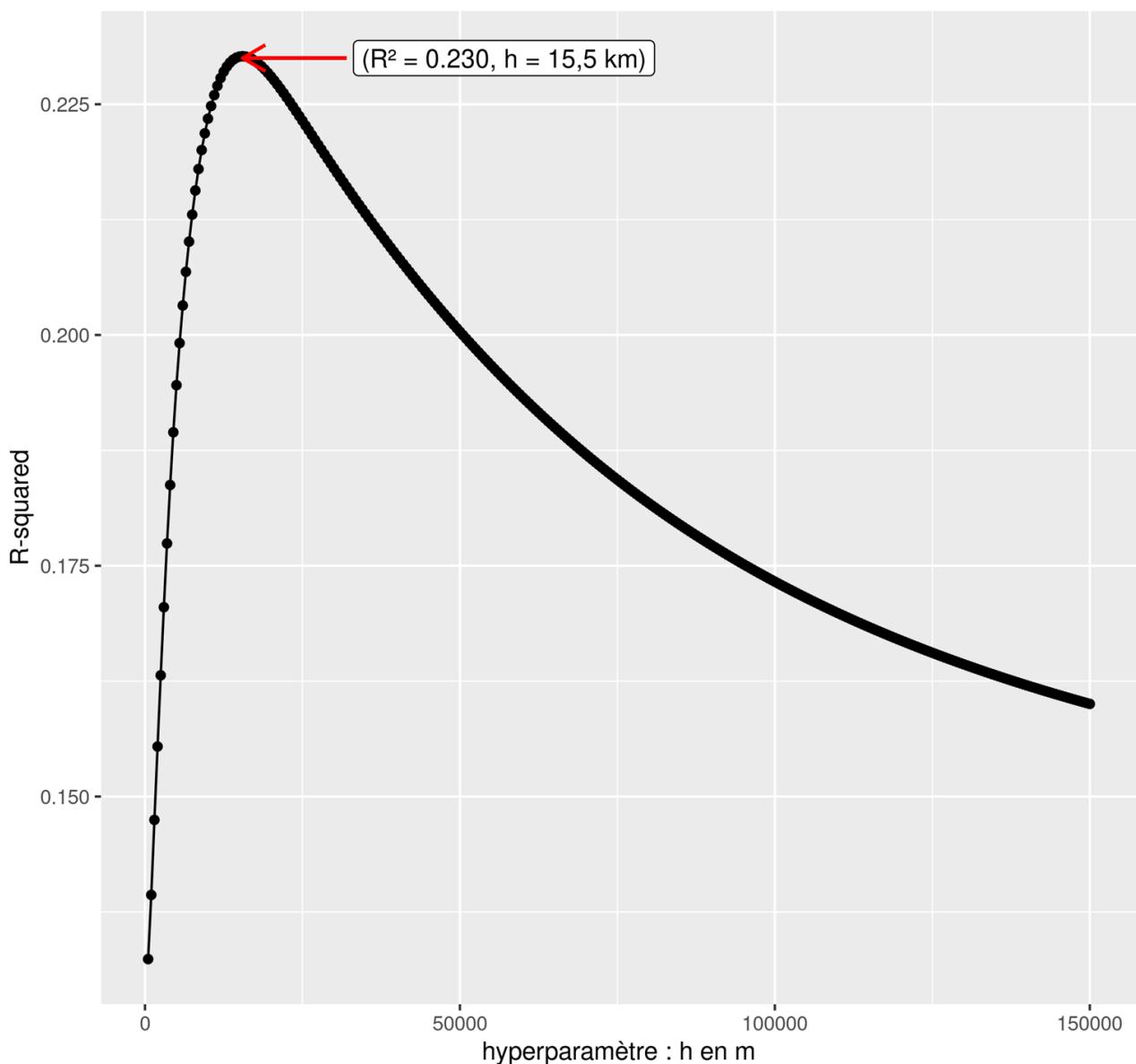
Coefficient de détermination ( $R^2$ ) en fonction de  
l'hyperparamètre  $h$  (cas du Puceron vert du pêcher *myzus persicae* (PV%P))



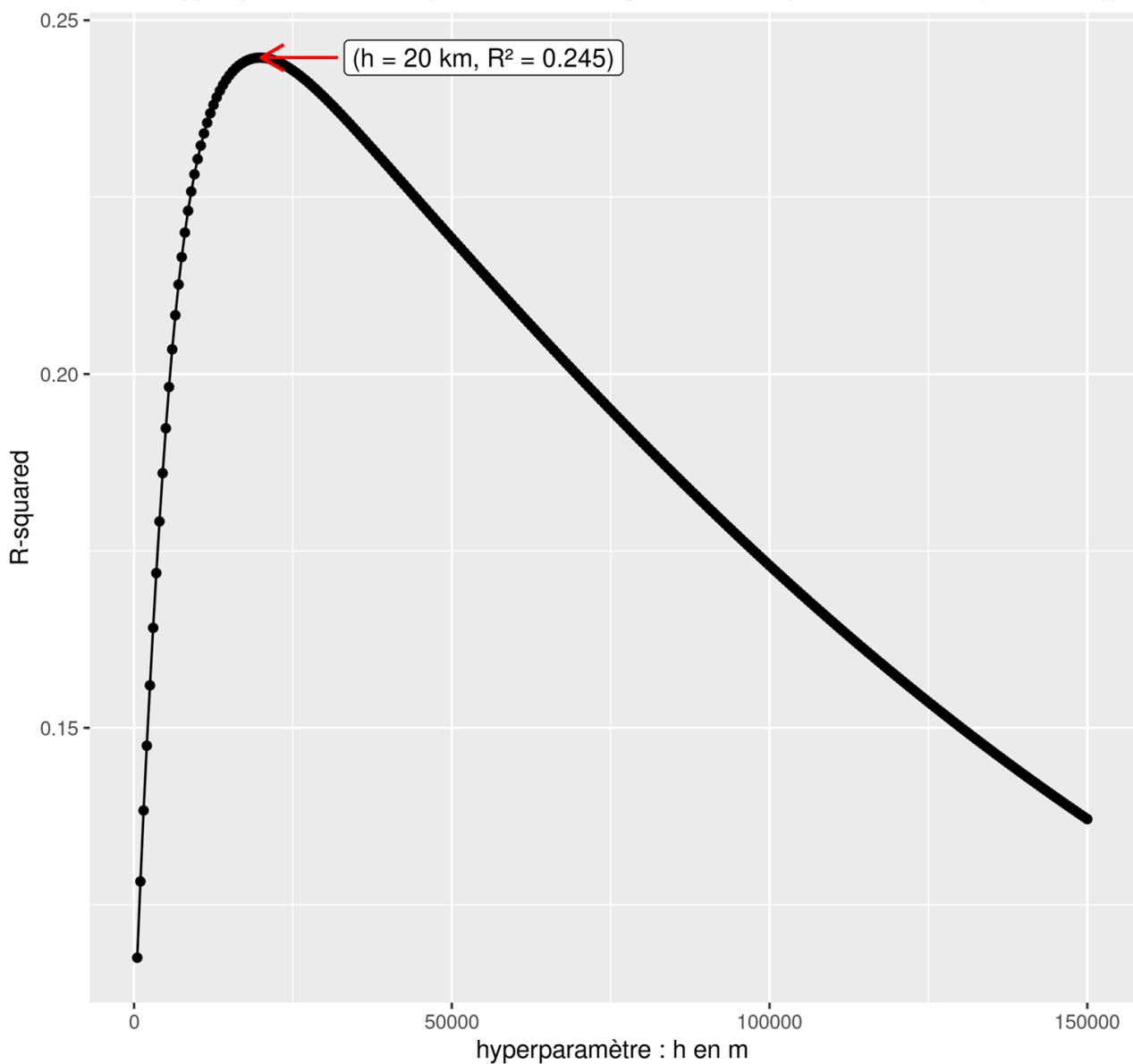
Coefficient de détermination ( $R^2$ ) en fonction  
de l'hyperparamètre  $h$  (cas du Piétin verse (PV\_%))



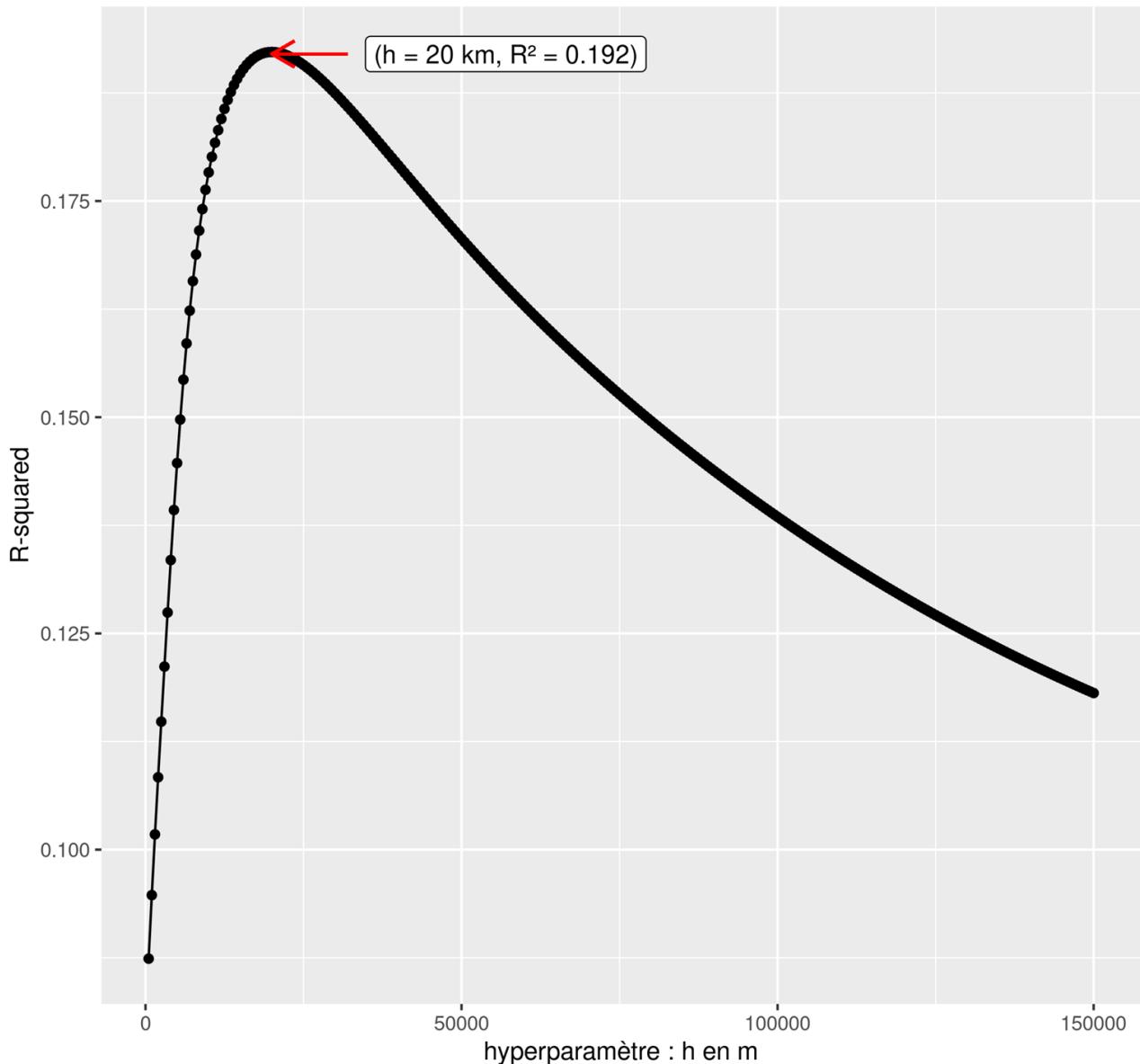
Coefficient de détermination ( $R^2$ ) en fonction  
de l'hyperparamètre  $h$  (cas du méligète du colza (MeP%P))



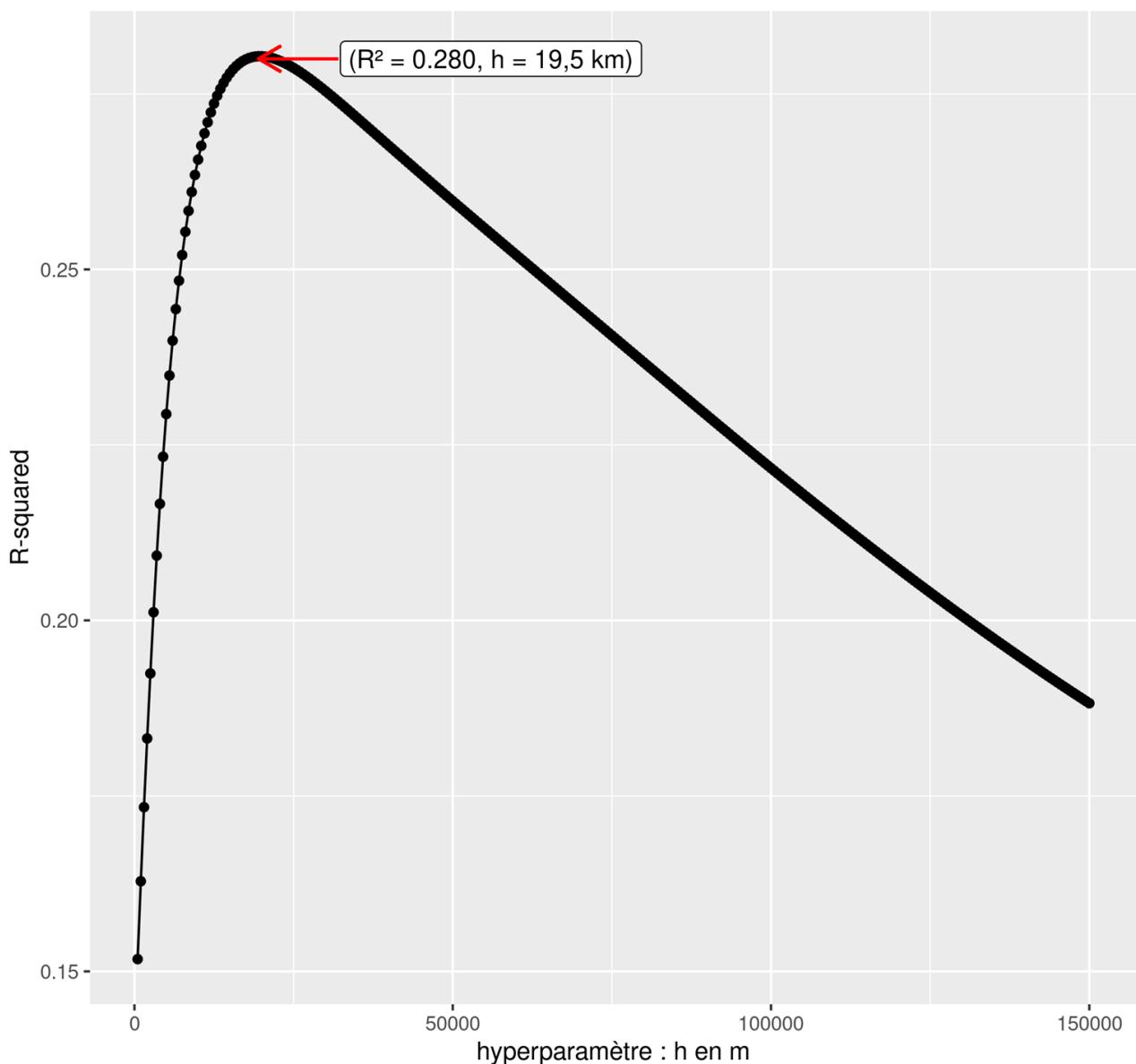
Coefficient de détermination ( $R^2$ ) en fonction  
de l'hyperparamètre  $h$  (cas du charançon du bourgeon terminal (CBTNbV))



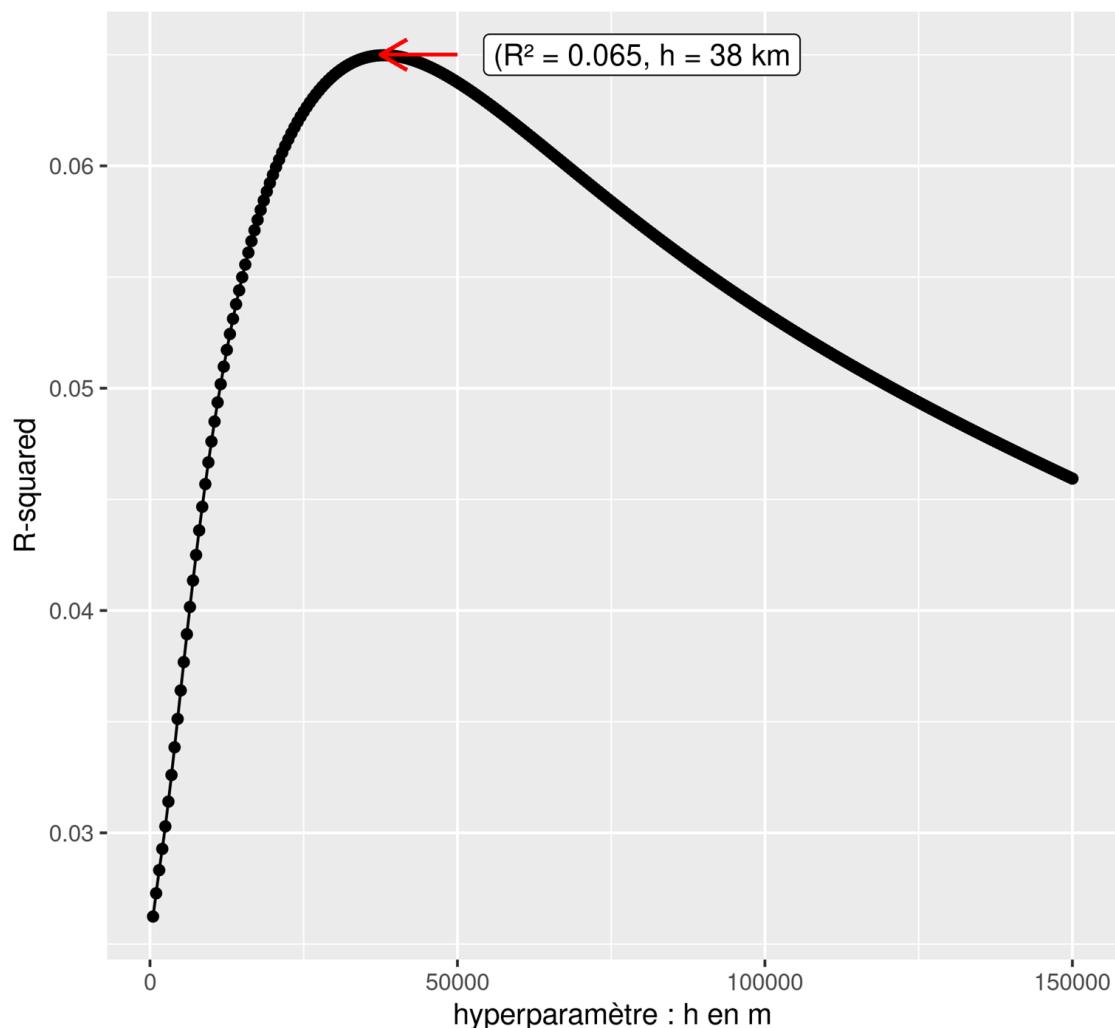
Coefficient de détermination ( $R^2$ ) en fonction  
de l'hyperparamètre  $h$  (cas du charançon de la tige du colza (ChTNbV))



Coefficient de détermination ( $R^2$ ) en fonction  
de l'hyperparamètre  $h$  (cas du charançon de la tige du chou (ChCNbV))



Coefficient de détermination ( $R^2$ ) en fonction de l'hyperparamètre  $h$  (cas d'altise petite des crucifères (PANbV))



### 11.1.1 Tableau récapitulatif des hyperparamètres optimaux ( $h$ ) des bioagresseurs

*Tableau 11.2* : Hyperparamètres optimaux ( $h$ ) et coefficients de détermination optimaux ( $R^2$ ) des bioagresseurs

Bioagresseurs	Hyperparamètres optimaux ( $h$ )	Coefficients de détermination optimaux ( $R^2$ )
Septoriose	$h = 16$ km	$R^2 = 0.285$
Rouille brune du blé	$h = 14.5$ km	$R^2 = 0.234$
Puceron vert du pêcher	$h = 19$ km	$R^2 = 0.274$
Piétin verste	$h = 22$ km	$R^2 = 0.077$
Méligèthe du colza	$h = 15.5$ km	$R^2 = 0.23$
Charançon du bourgeon terminal	$h = 20$ km	$R^2 = 0.245$
Charançons de la tige du colza	$h = 20$ km	$R^2 = 0.192$
Charançon de la tige du chou	$h = 19.5$ km	$R^2 = 0.28$
Altise petite des crucifères	$h = 38$ km	$R^2 = 0.065$

## 11.2 Pression annuelle des bioagresseurs d'espèce $i$ , prédicté en utilisant la méthode des $k$ -NN ( $k$ plus proches voisins)

Dans cette partie la pression des bioagresseurs d'espèce  $i$ , prédicté dans une parcelle  $p$  est obtenue en utilisant la méthode des  $k$ -NN ( $k$  plus proches voisins). Ici les  $k$  plus proches voisins représentent les  $k$  plus proches parcelles de celle où nous cherchons à prédire la pression des bioagresseurs d'espèce  $i$ . Cette dernière est calculée en faisant la moyenne arithmétique des pressions des bioagresseurs des  $k$  plus proches parcelles. Avec la méthode des  $k$ -NN, la pression des bioagresseurs d'espèce  $i$ , prédicté dans une parcelle  $p$  s'écrit :

$$P_p^{prédicté}(bio\_i) = \frac{\sum_{p'=1}^k P_{p'}^{réelle}(bio\_i)}{k} \quad (11.2)$$

où :

$k$  : représente le nombre de parcelles les plus proches de la parcelle  $p$ ;

$P_{p'}^{réelle}(bio\_i)$ : est la pression réelle des bioagresseur d'espèce  $i$  dans une parcelle  $p'$  parmi les  $k$  plus proches parcelles avec  $1 \leq p' \leq k$ .

### Recherche de l'hyperparamètre $k$ optimal des bioagresseurs d'espèce $i$

Pour pouvoir prédire la pression des bioagresseurs d'espèce  $i$  dans une parcelle avec la méthode des  $k$ -NN, il faut avant tout trouver l'hyperparamètre  $k$  optimal associé à cette espèce  $i$  de bioagresseurs à partir des données d'entraînement. L'hyperparamètre  $k$  optimal associé à une espèce  $i$  de bioagresseurs correspond à la valeur de  $k$  parmi les possibles, qui maximise le score de prédiction ( $R^2$ ).

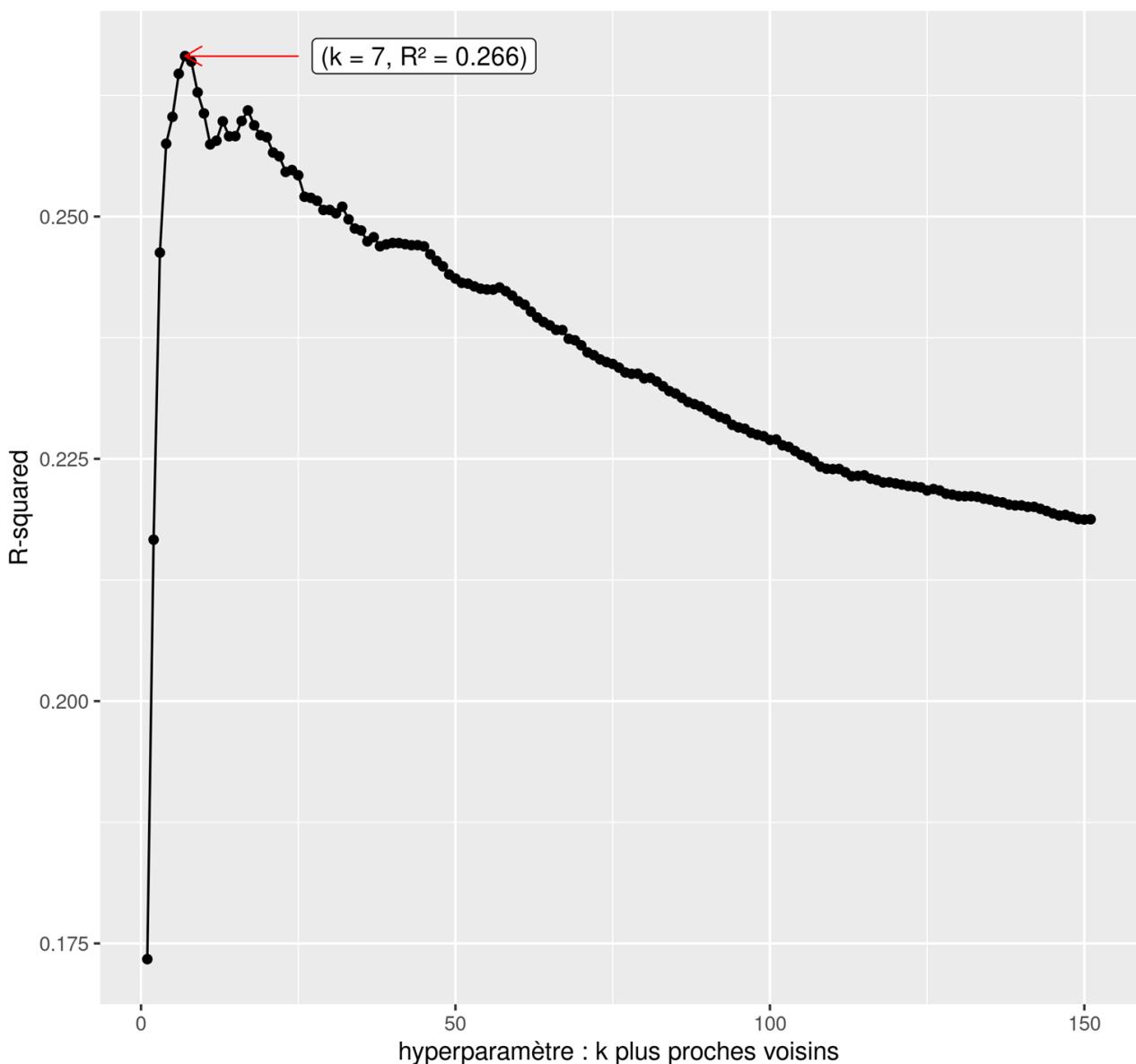
Procédure de recherche de l'hyperparamètre  $k$  optimal :

- Varier l'hyperparamètre  $k$  dans une plage de valeurs ;
- Pour chaque valeur de  $k$ , prédire la pression des bioagresseurs d'espèce  $i$  au niveau de chaque parcelle ;
- Calculer le coefficient de détermination ( $R^2$ ) entre les pressions réelles  $P_{p'}^{réelle}(bio\_i)$  et les pressions prédictes  $P_p^{prédicté}(bio\_i)$  correspondant à l'hyperparamètre  $k$ .

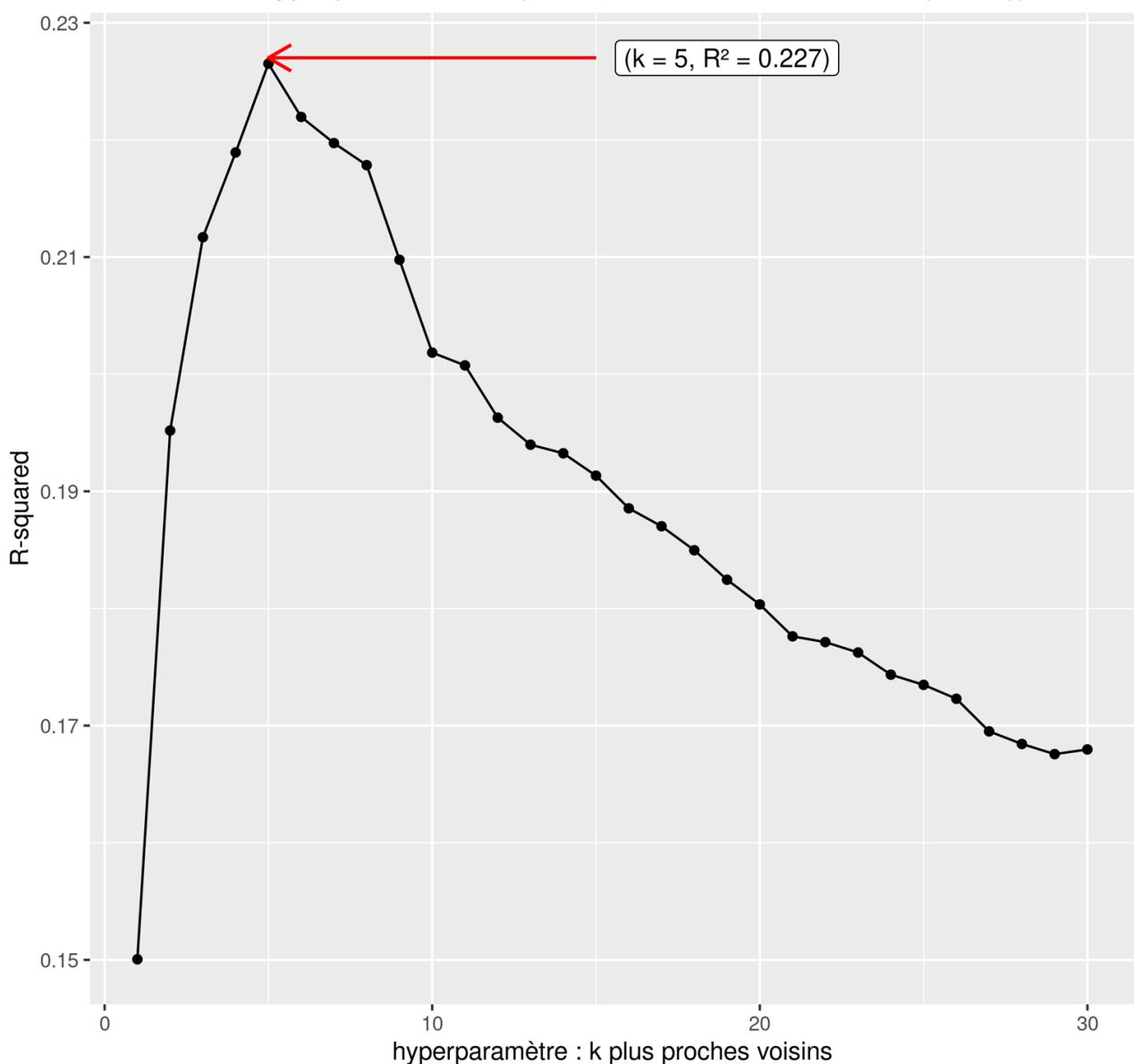
Nous avons utilisé les mêmes hyperparamètres  $k$  pour entraîner séparément les données de chaque année (2010 à 2020). Ensuite, on a calculé le  $R^2$  correspondant aux observations des années (2010 à 2020) prédictes avec le même hyperparamètre.

Les graphiques suivants représentent les résultats obtenus sur chaque bioagresseur étudié :

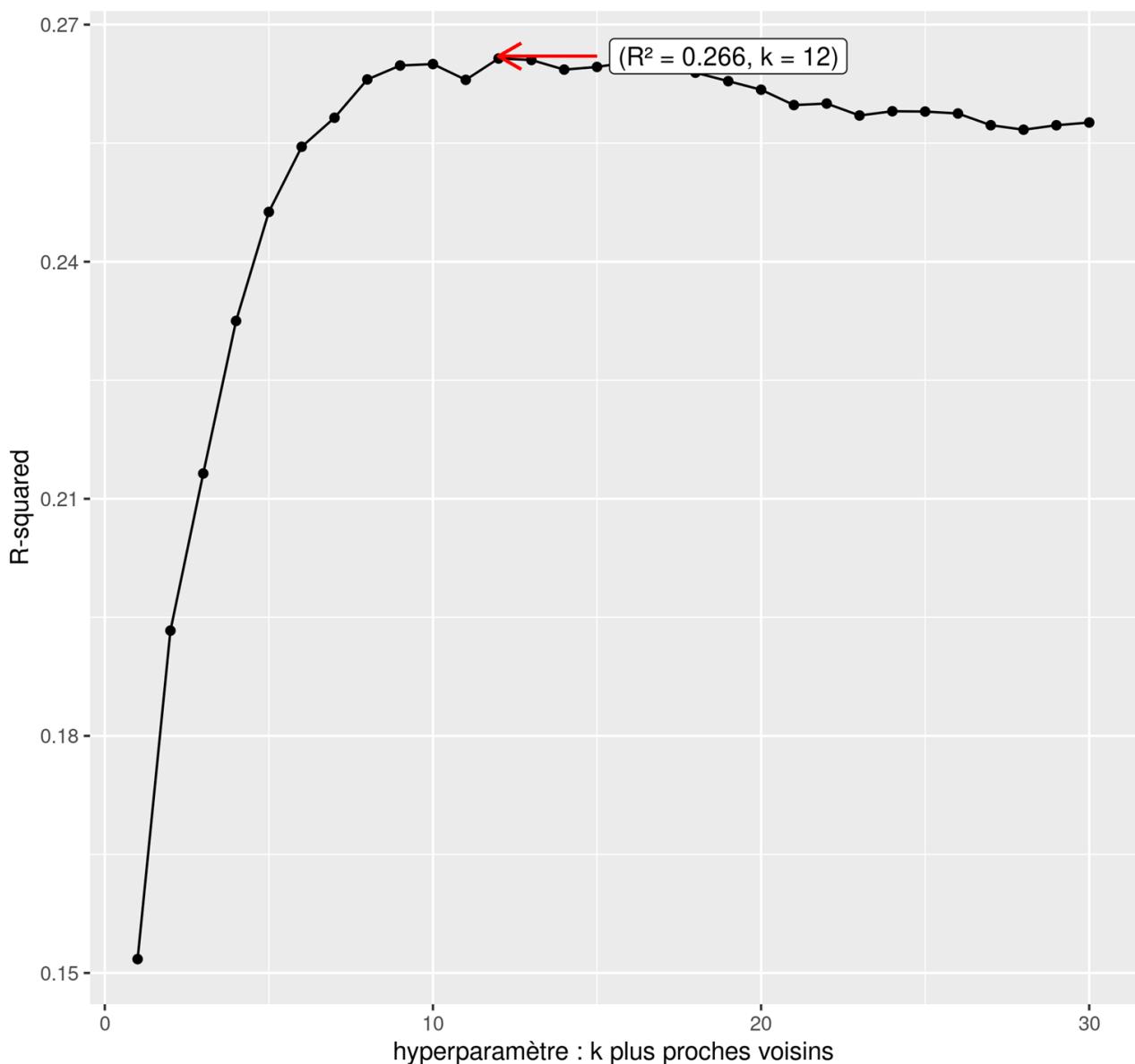
Coefficient de détermination ( $R^2$ ) en fonction  
de l'hyperparamètre k (cas de la septoriose)



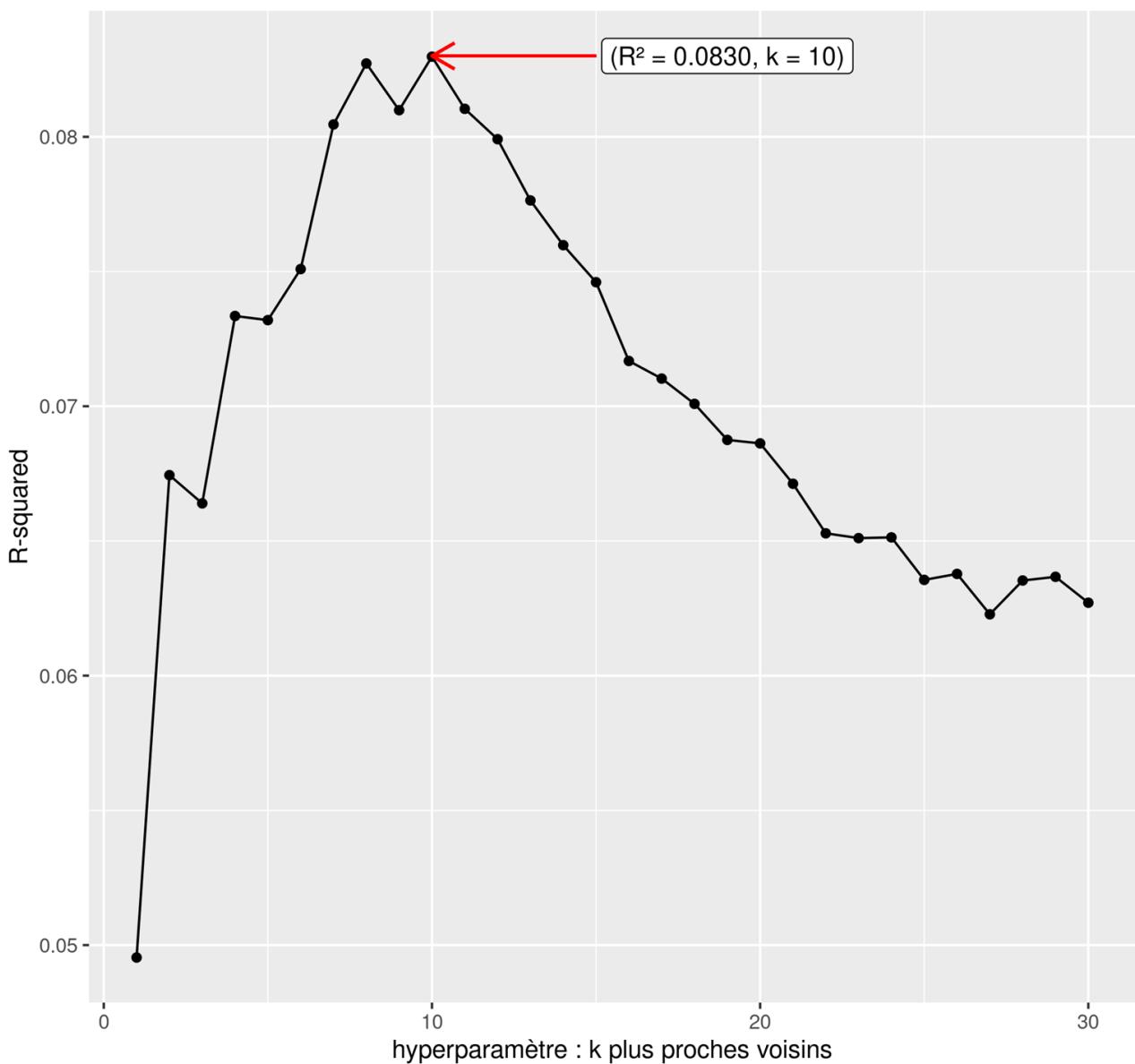
Coefficient de détermination ( $R^2$ ) en fonction  
de l'hyperparamètre  $k$  (cas de la rouille brune du blé (RBF3))



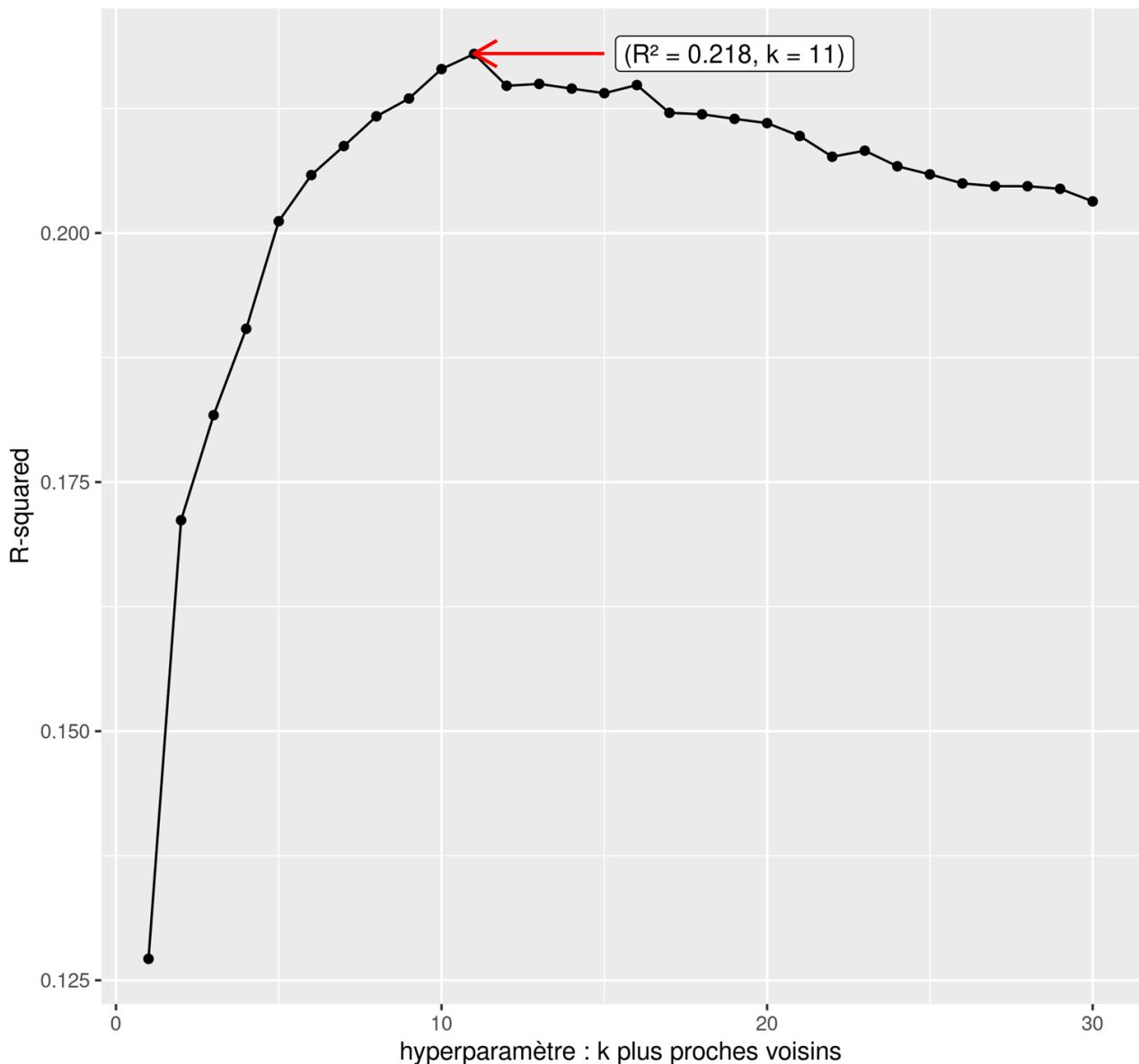
Coefficient de détermination ( $R^2$ ) en fonction  
de l'hyperparamètre k (Puceron vert du pêcher myzus persicae (PV%P))



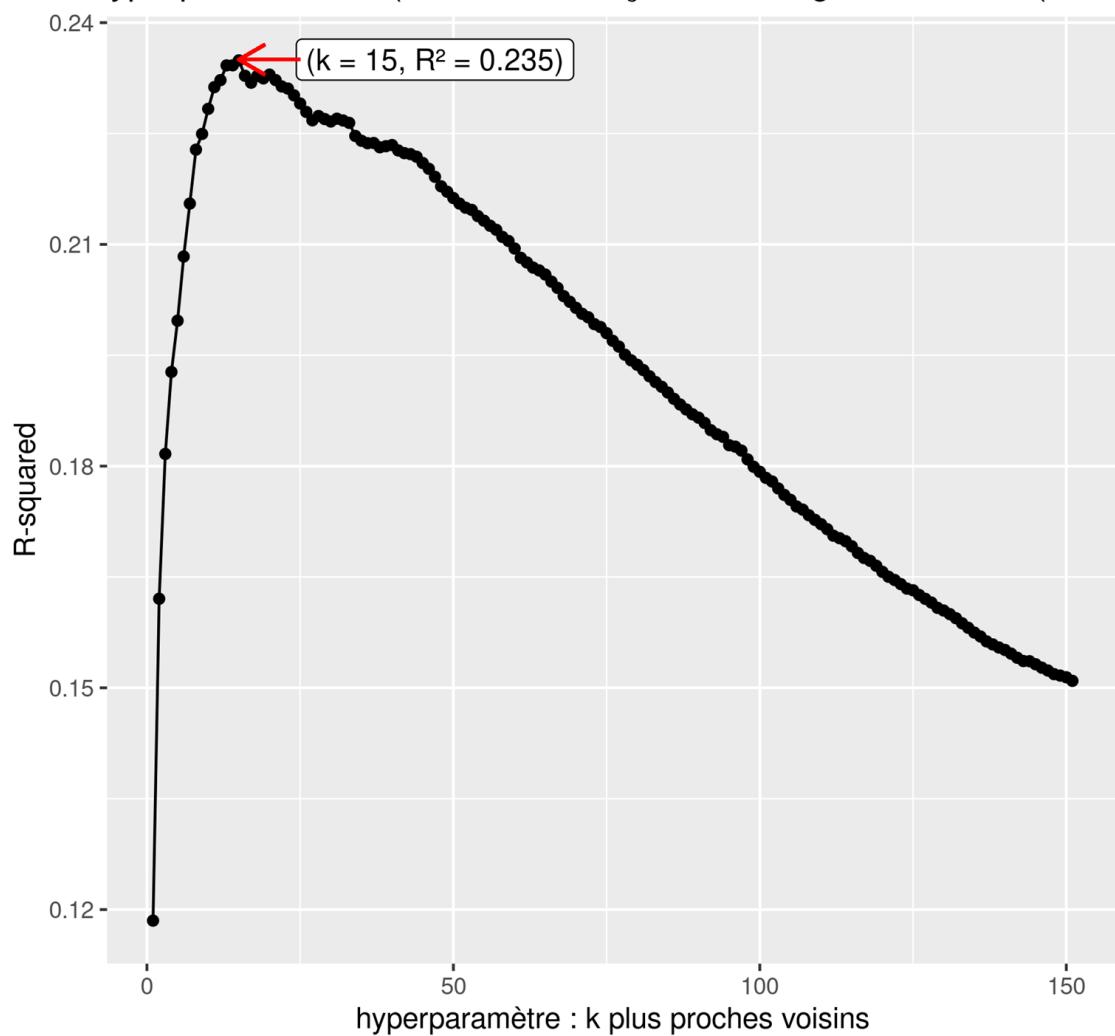
Coefficient de détermination ( $R^2$ ) en fonction  
de l'hyperparamètre k (Piétin verse (PV\_%))



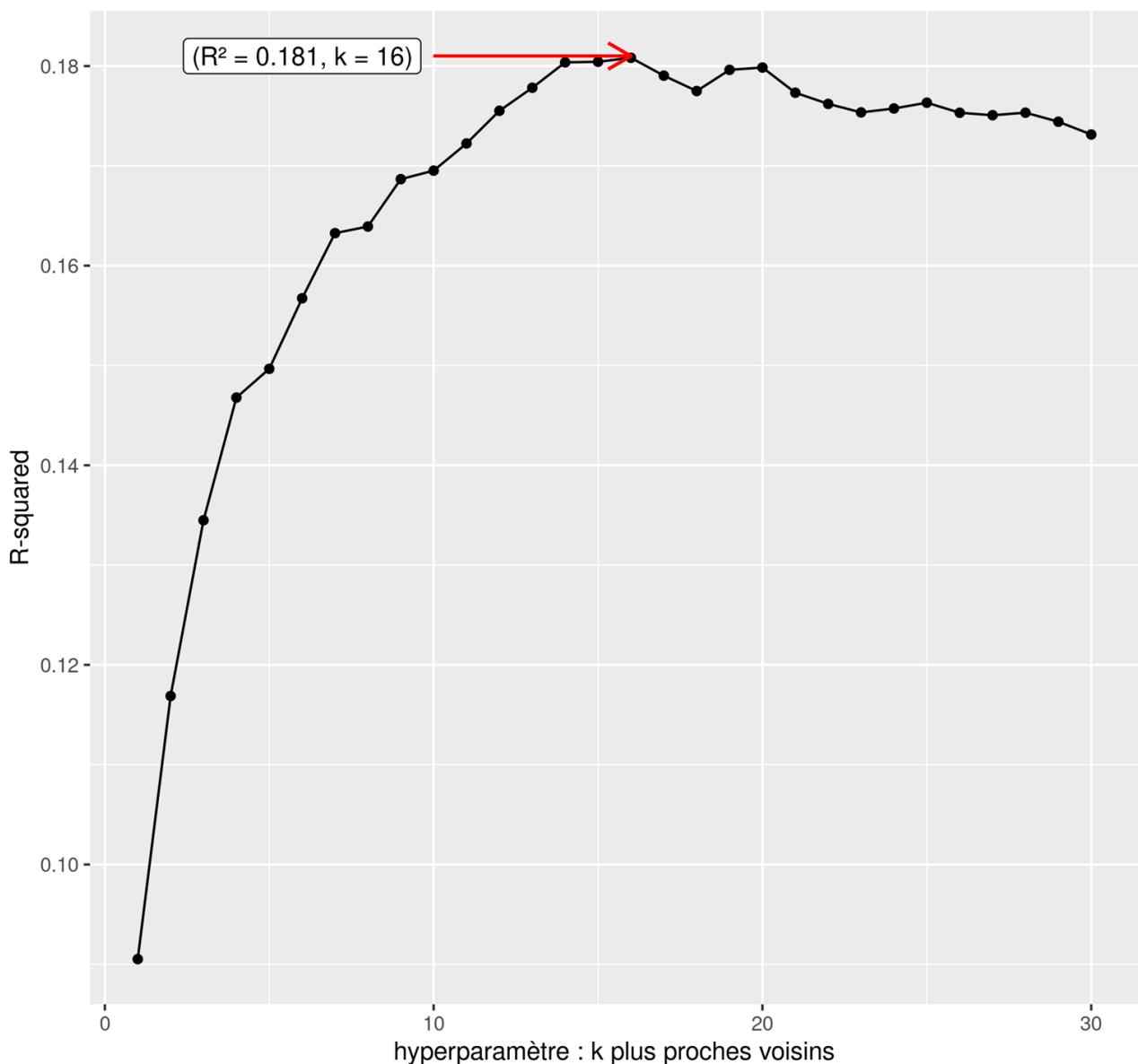
Coefficient de détermination ( $R^2$ ) en fonction  
de l'hyperparamètre k (Méligèthe du colza (Mel%P))



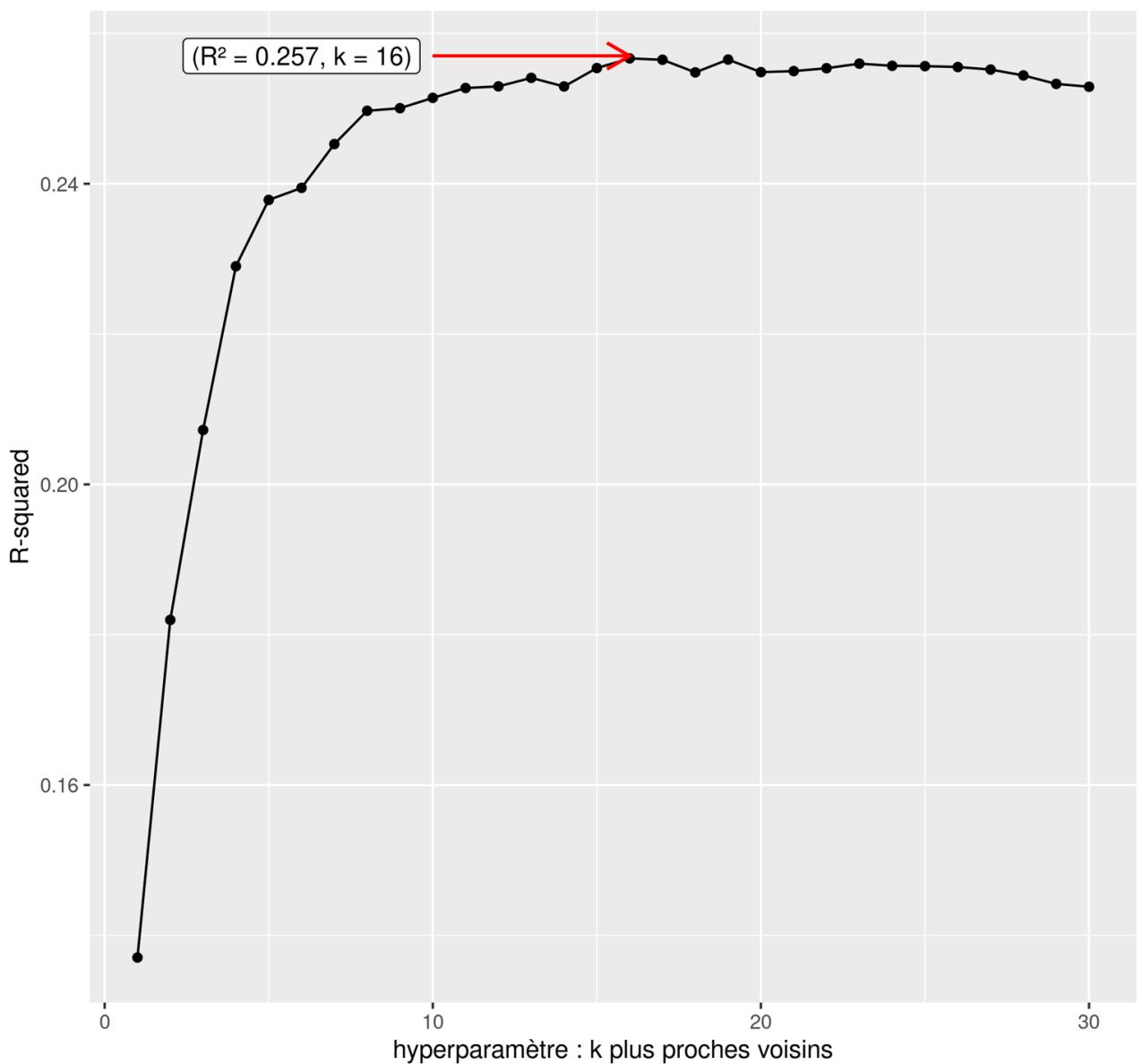
Coefficient de détermination ( $R^2$ ) en fonction  
de l'hyperparamètre  $k$  (cas du charançon du bourgeon terminal (CBTN))



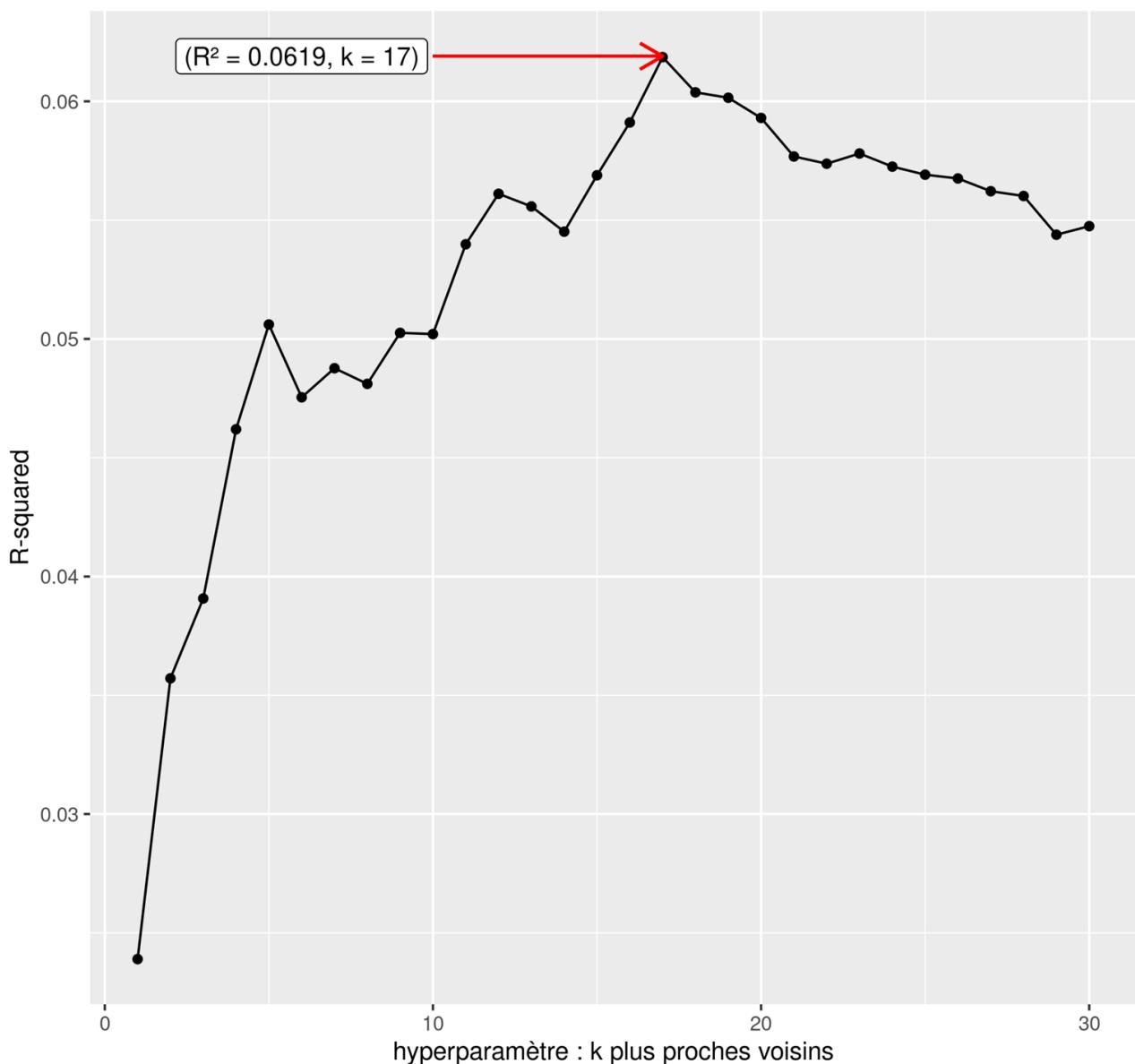
Coefficient de détermination ( $R^2$ ) en fonction  
de l'hyperparamètre k (cas du charançon de la tige du colza (ChTNbV))



Coefficient de détermination ( $R^2$ ) en fonction  
de l'hyperparamètre k (cas du charançon de la tige du chou (ChCNbV))



Coefficient de détermination ( $R^2$ ) en fonction  
de l'hyperparamètre k (Altise petite des crucifères (PANbV))



### 11.2.1 Tableau récapitulatif des hyperparamètres optimaux (k) des bioagresseurs

*Tableau 11.3* : Hyperparamètres optimaux ( $k$ ) et coefficients de détermination optimaux ( $R^2$ ) des bioagresseurs

Bioagresseurs	Hyperparamètres optimaux (k)	Coefficients de détermination optimaux ( $R^2$ )
Septoriose	$k = 7$	$R^2 = 0.266$
Rouille brune du blé	$k = 5$	$R^2 = 0.227$
Puceron vert du pêcher	$k = 12$	$R^2 = 0.266$
Piétin verste	$k = 10$	$R^2 = 0.083$
Méligèthe du colza	$k = 11$	$R^2 = 0.218$
Charançon du bourgeon terminal	$k = 15$	$R^2 = 0.235$
Charançon de la tige du colza	$k = 16$	$R^2 = 0.181$
Charançon de la tige du chou	$k = 16$	$R^2 = 0.257$
Altise petite des crucifères	$k = 17$	$R^2 = 0.0619$

### 11.3 Pression annuelle des bioagresseurs d'espèce $i$ , simulée

Nous avons deux issues possibles pour un comptage de bioagresseurs dans une parcelle  $p$  soit nous avons une valeur qui dépasse le seuil médian (défini à la partie 8.1) ou soit en dessous du seuil médian donc chaque comptage suit la loi de Bernoulli avec  $prob$  = la probabilité d'avoir un succès. Nous allons avoir une loi Binomiale si on répète l'expérience dans une même parcelle. Ainsi nous pouvons simuler pour chaque parcelle  $p$ , la pression des bioagresseurs d'espèce  $i$ , notée  $P_p^{simulée}(bio_i)$ .

Soit  $n$  le nombre de fois où un comptage de bioagresseurs d'espèce  $i$ , est effectué dans la parcelle  $p$ .

Parce que les observations sont souvent en petit nombre les valeurs possibles de pressions sont discrètes et les  $R^2$  ne peuvent pas atteindre des valeurs proches de 1. Pour estimer la valeur potentielle du  $R^2$  si les estimations étaient parfaites, nous allons générer aléatoirement  $n$  ratios compris entre 0 et 1 à partir d'une probabilité  $prob$  (=probabilité de dépasser le seuil pour chaque observation). La pression de bioagresseurs d'espèce  $i$ , simulée dans une parcelle  $p$  s'écrit :

$$P_p^{simulée}(bio_i) = \frac{\text{nombre de succès obtenus dans les } n \text{ expériences}}{\text{nombre d'expériences, } n} \quad (11.3)$$

**Les coefficients de détermination  $R^2$  entre les  $P_p^{simulée}(bio_i)$  et les  $P_p^{prédictive}(bio_i)$**  (obtenues avec les hyperparamètres  $h$  optimaux, du **Tableau 11.2 : Hyperparamètres optimaux (h) et coefficients de détermination optimaux ( $R^2$ ) des bioagresseurs**)

- Nous avons simulé les pressions des bioagresseurs d'espèce  $i$ ,  $P_p^{simulée}(bio_i)$  au niveau de chaque parcelle en utilisant les pressions de bioagresseurs prédictives,  $P_p^{prédictive}(bio_i)$ ;
- Ensuite nous avons calculé le coefficient de détermination  $R^2$  entre les  $P_p^{simulée}(bio_i)$  et les  $P_p^{prédictive}(bio_i)$ ;
- Si on simule plusieurs fois  $P_p^{simulée}(bio_i)$  au niveau de chaque parcelle, on se retrouve avec un  $R^2$  associé à chaque simulation.

Remarque : Nous réalisons deux types de simulations :

Dans les simulations de type 1,  $prob = P_p^{réelle}(bio_i)$  : on utilise le ratio observé pour faire les simulations binomiales;

Dans les simulations de type 2,  $prob = P_p^{prédictive}(bio_i)$  : on utilise le ratio estimé pour faire les simulations binomiales.

### 12 Comparaison des modèles sur les observations de chaque année

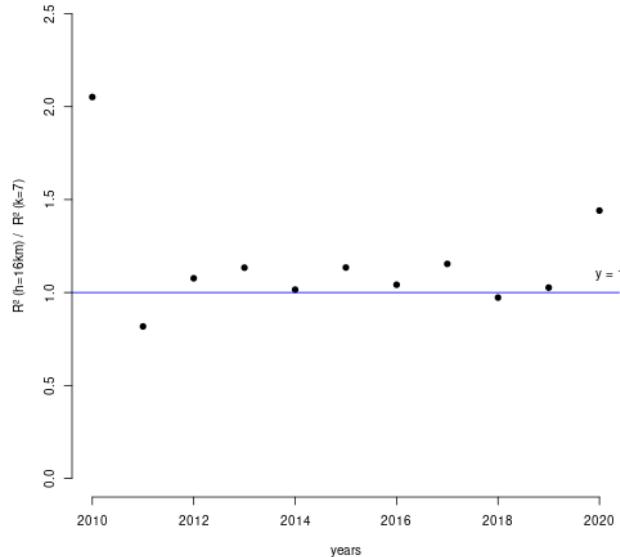
Cette partie concerne la comparaison des différents modèles utilisés. On a essayé de comparer le modèle de la moyenne pondérée avec les autres modèles en faisant le rapport des  $R^2$  correspondant aux observations de chaque année (2010 à 2020).

Nous avons pris comme modèle de référence celui de la moyenne pondérée.

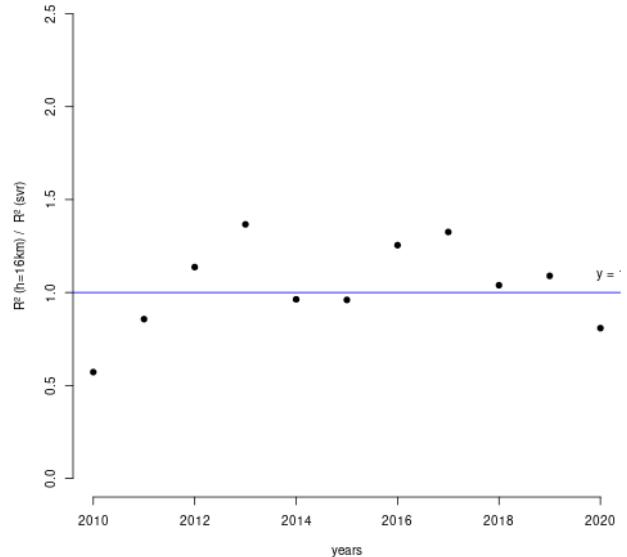
Les graphiques suivants représentent les résultats obtenus pour chaque bioagresseur.

## 12.1 Comparaison des résultats obtenus sur la septoriose

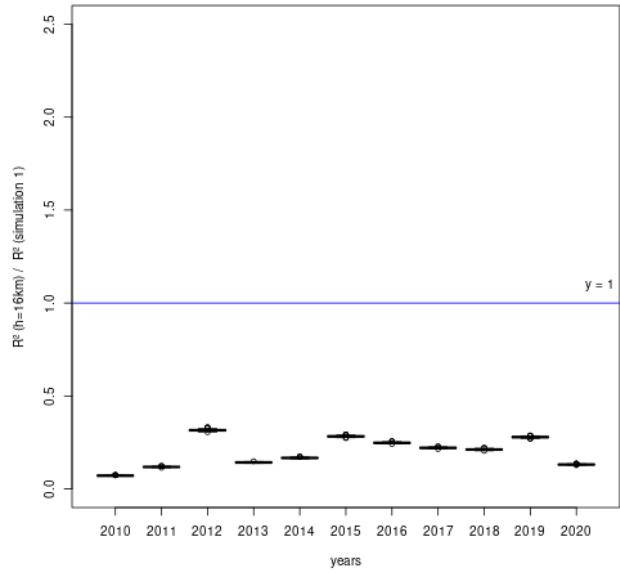
Rapport entre  $R^2$  ( $h=16\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2$  ( $k=7$ ) obtenu avec la méthode des  $k$  plus proches voisins



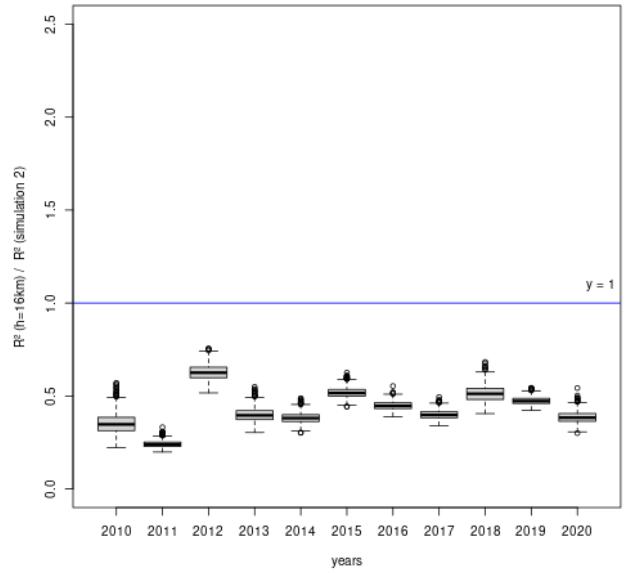
Rapport entre  $R^2$  ( $h=16\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2_{\text{svr}}$  par année ajusté sur toutes années avec variable année



Rapport entre  $R^2$  ( $h=16\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2$  obtenu avec la simulation 1

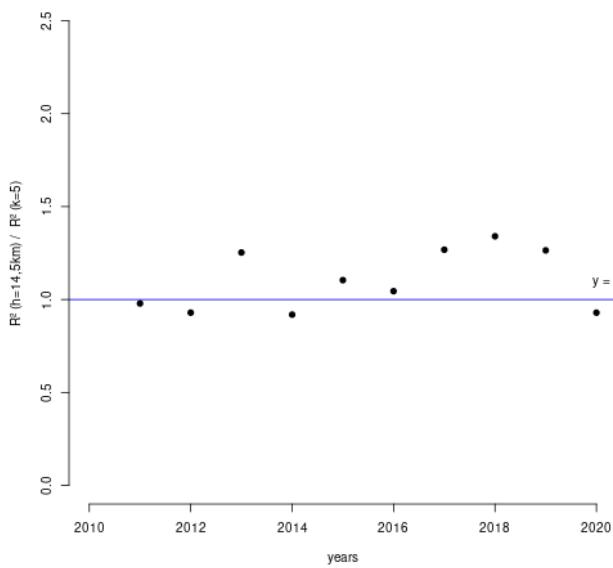


Rapport entre  $R^2$  ( $h=16\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2$  obtenu avec la simulation 2

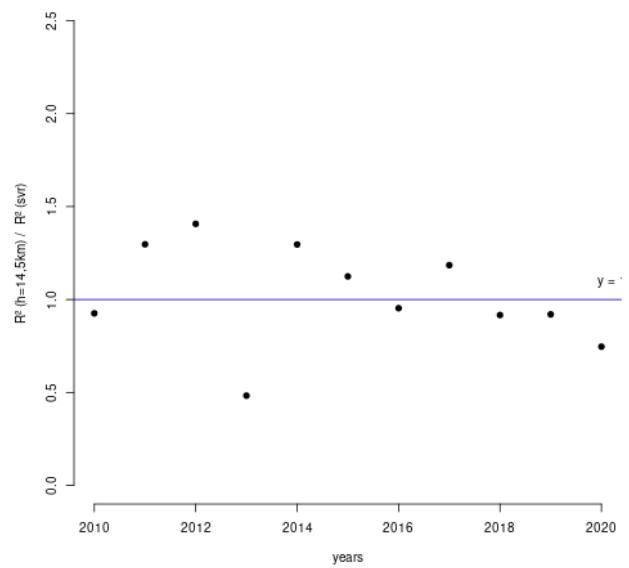


## 12.2 Comparaison des résultats obtenus sur la rouille brune du blé

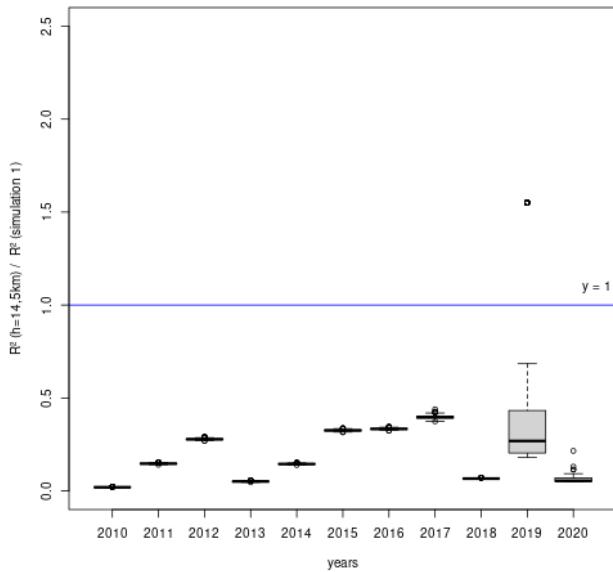
Rapport entre  $R^2$  ( $h=14,5\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2$  ( $k=5$ ) obtenu avec la méthode des  $k$  plus proches voisins



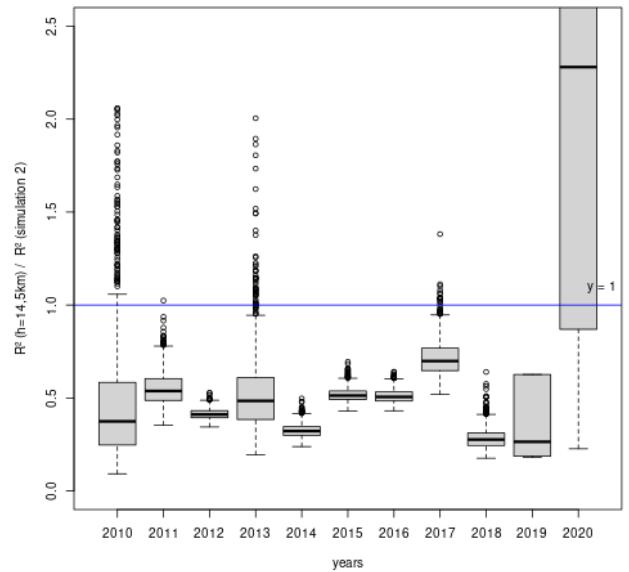
Rapport entre  $R^2$  ( $h=14,5\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2_{\text{svr}}$  par année ajusté sur toutes années avec variable année



Rapport entre  $R^2$  ( $h=14,5\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2$  obtenu avec la simulation 1

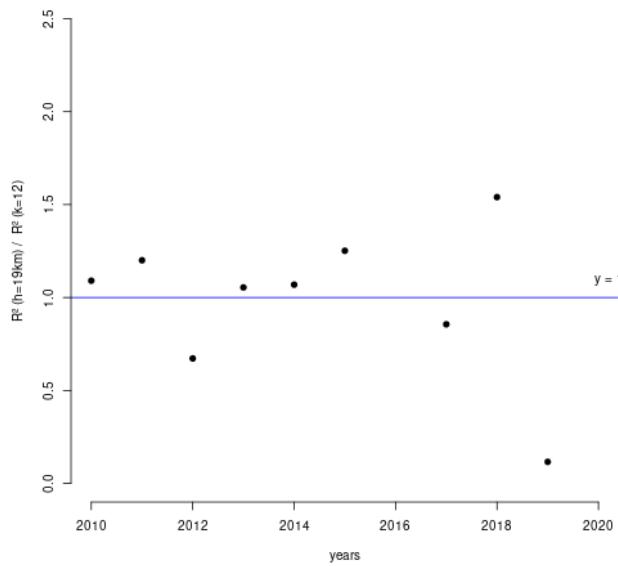


Rapport entre  $R^2$  ( $h=14,5\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2$  obtenu avec la simulation 2

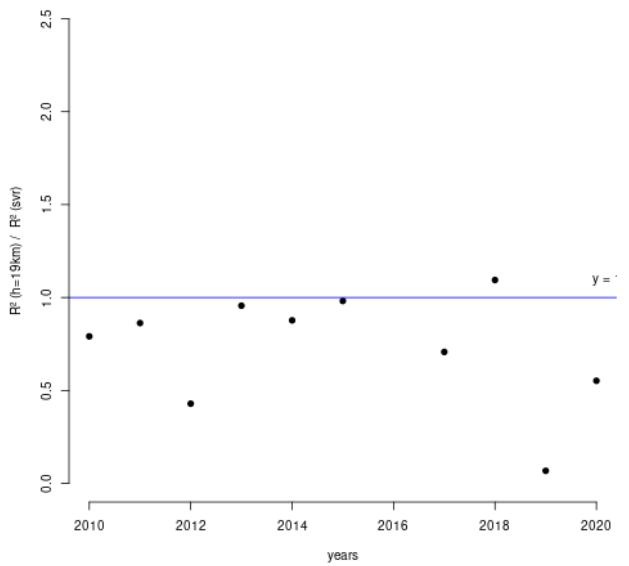


## 12.3 Comparaison des résultats obtenus sur le puceron vert de pêcher

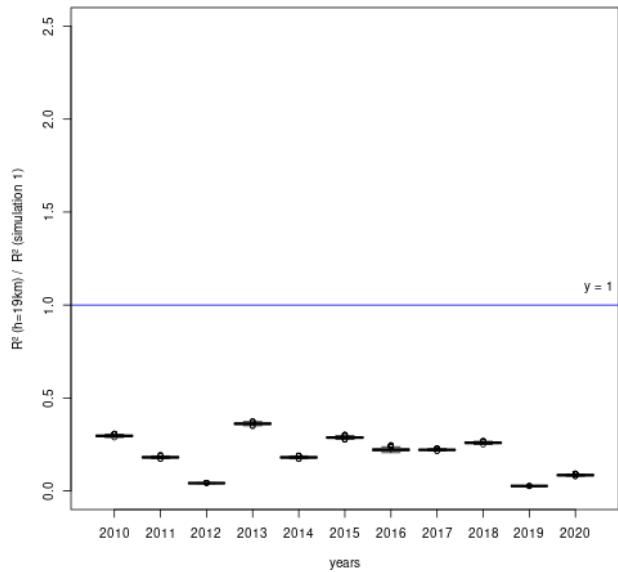
Rapport entre  $R^2$  ( $h=19\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2$  ( $k=12$ ) obtenu avec la méthode des  $k$  plus proches voisins



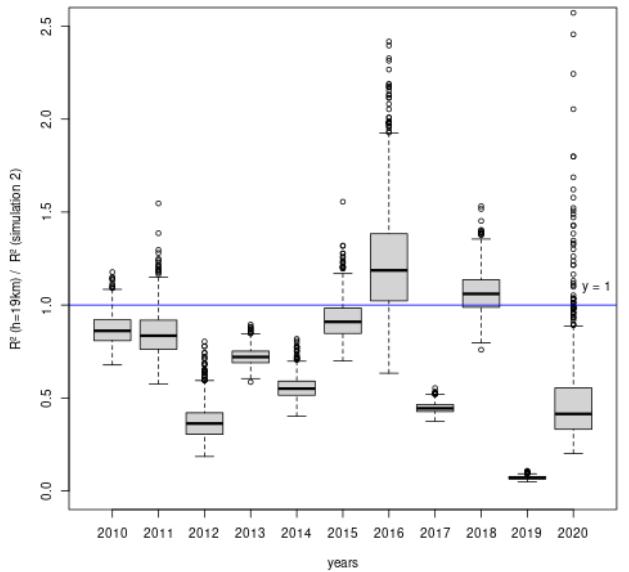
Rapport entre  $R^2$  ( $h=19\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2,svr$  par année ajusté sur toutes années avec variable année



Rapport entre  $R^2$  ( $h=19\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2$  obtenu avec la simulation 1

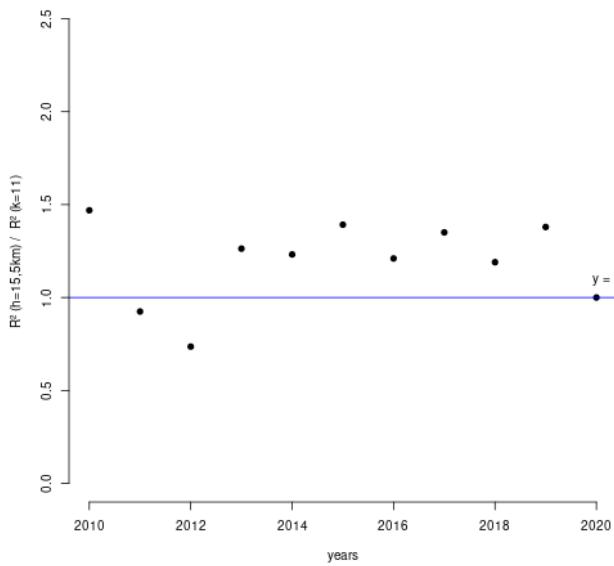


Rapport entre  $R^2$  ( $h=19\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2$  obtenu avec la simulation 2

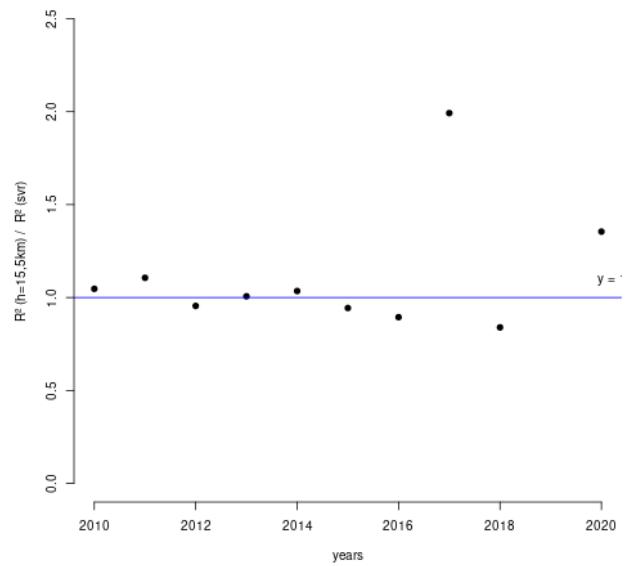


## 12.4 Comparaison des résultats obtenus sur le mélange de colza

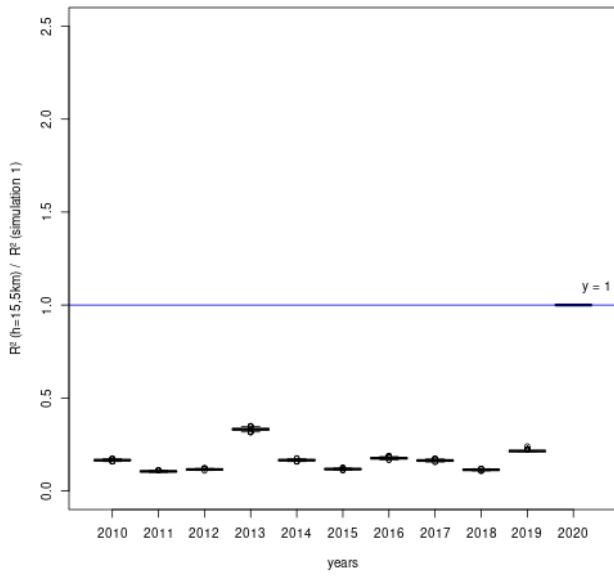
Rapport entre  $R^2$  ( $h=15,5\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2$  ( $k=11$ ) obtenu avec la méthode des  $k$  plus proches voisins



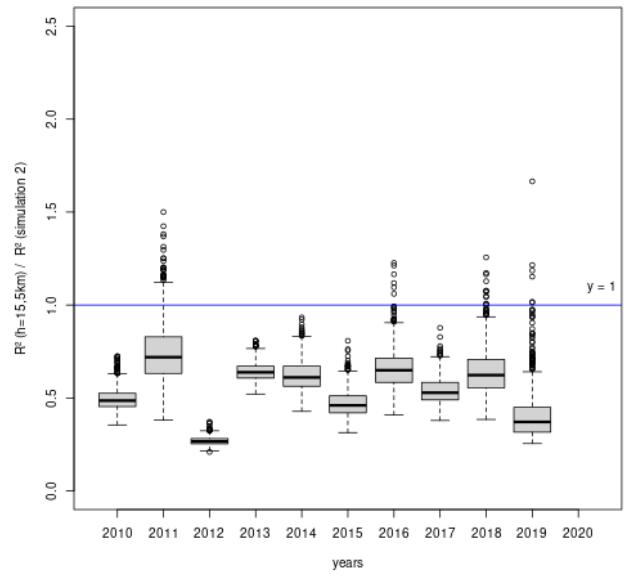
Rapport entre  $R^2$  ( $h=15,5\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2_{\text{svr}}$  par année ajusté sur toutes années avec variable année



Rapport entre  $R^2$  ( $h=15,5\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2$  obtenu avec la simulation 1

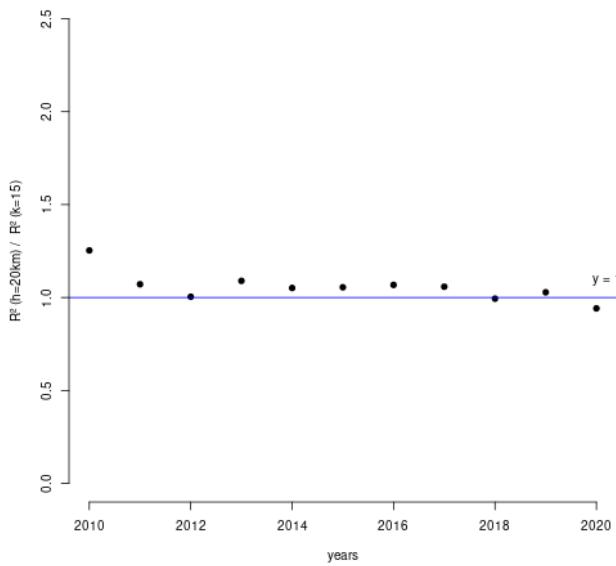


Rapport entre  $R^2$  ( $h=15,5\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2$  obtenu avec la simulation 2

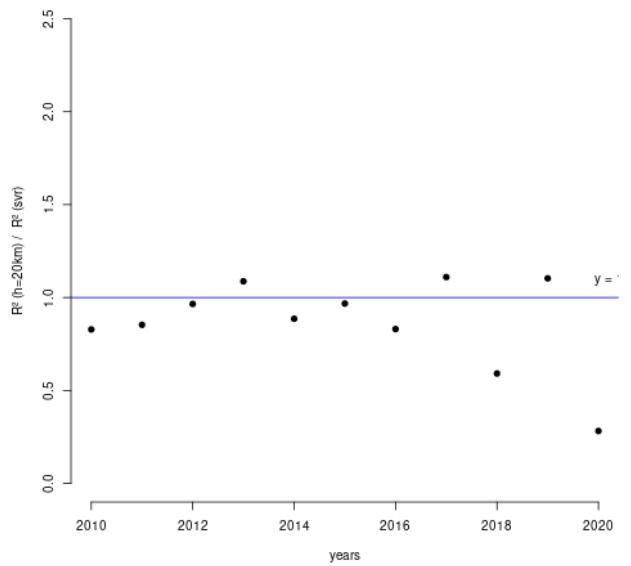


## 12.5 Comparaison des résultats obtenus sur le charançon du bourgeon terminal

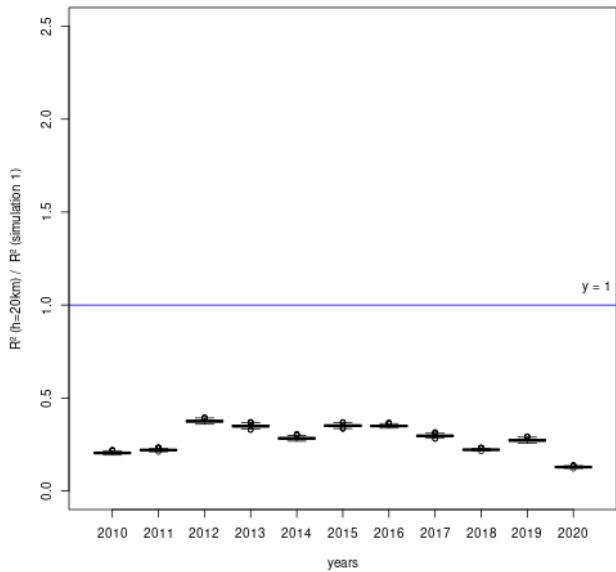
Rapport entre  $R^2$  ( $h=20\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2$  ( $k=15$ ) obtenu avec la méthode des  $k$  plus proches voisins



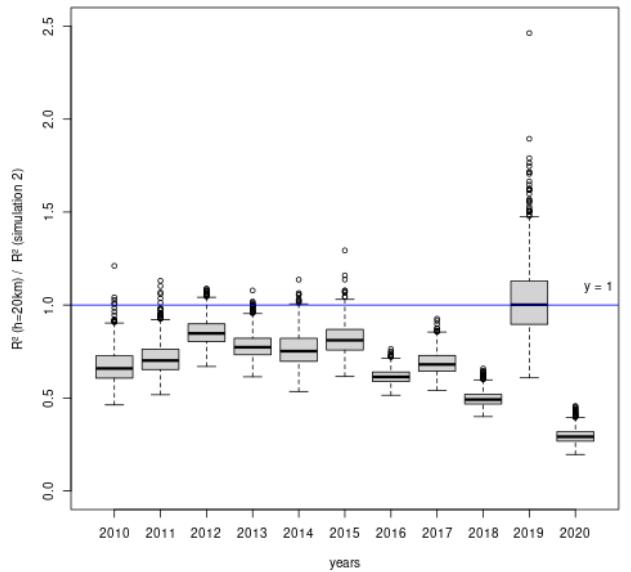
Rapport entre  $R^2$  ( $h=20\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2_{svr}$  par année ajusté sur toutes années avec variable année



Rapport entre  $R^2$  ( $h=20\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2$  obtenu avec la simulation 1

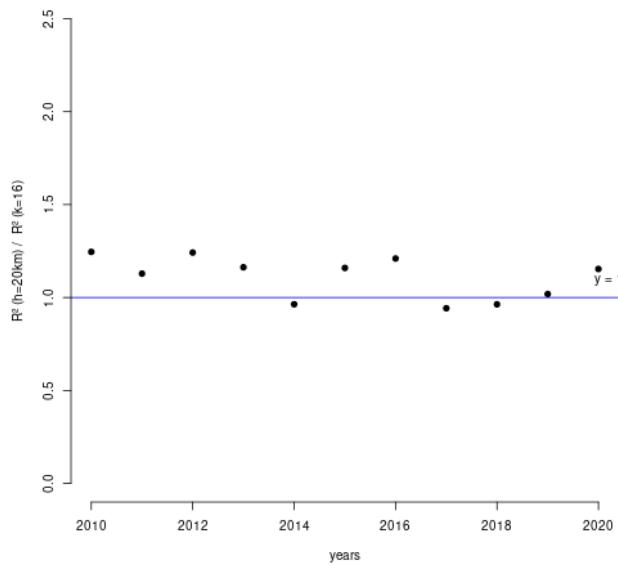


Rapport entre  $R^2$  ( $h=20\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2$  obtenu avec la simulation 2

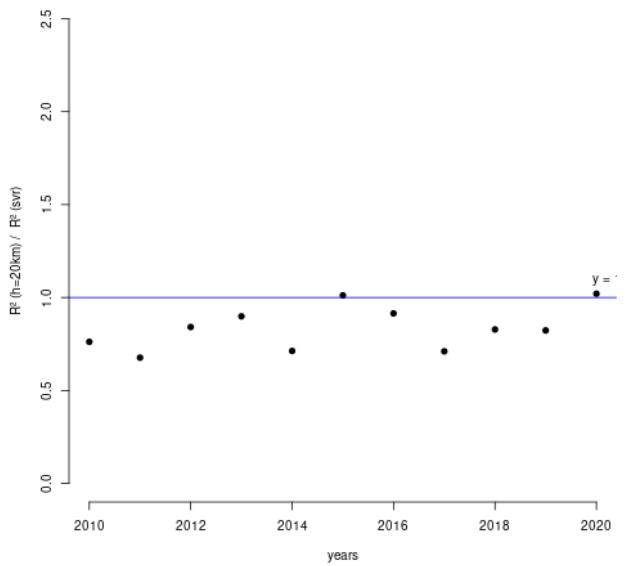


## 12.6 Comparaison des résultats obtenus sur le charançon de la tige du colza

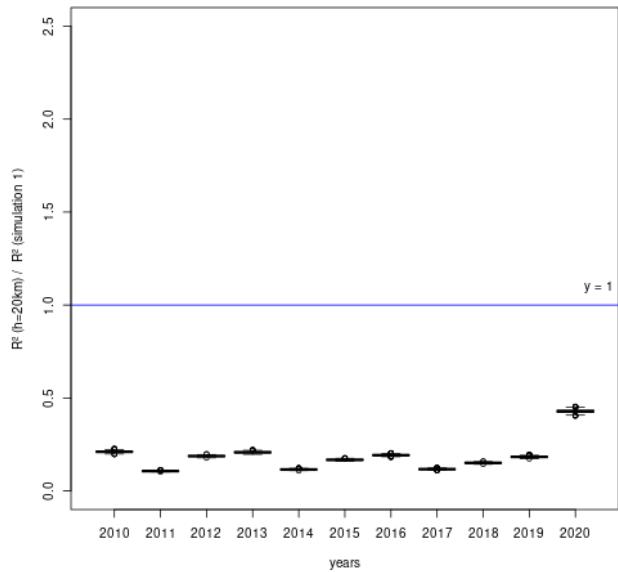
Rapport entre  $R^2$  ( $h=20\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2$  ( $k=16$ ) obtenu avec la méthode des  $k$  plus proches voisins



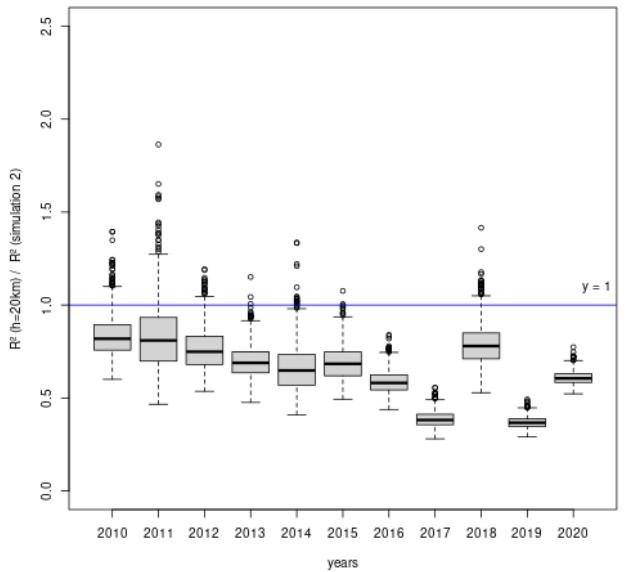
Rapport entre  $R^2$  ( $h=20\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2_{\text{svr}}$  par année ajusté sur toutes années avec variable année



Rapport entre  $R^2$  ( $h=20\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2$  obtenu avec la simulation 1

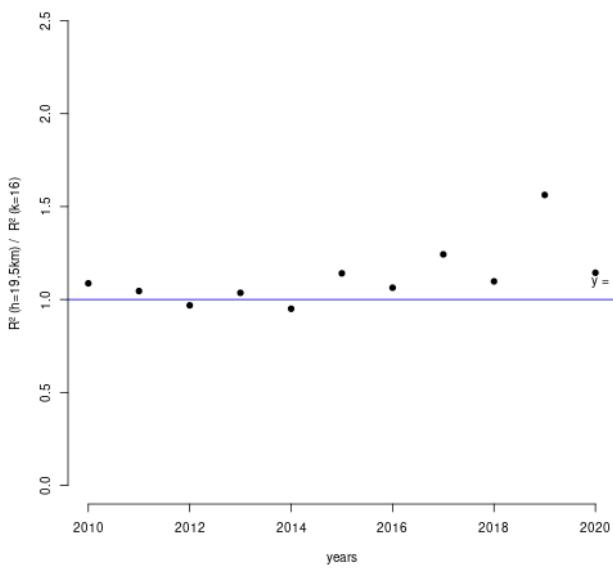


Rapport entre  $R^2$  ( $h=20\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2$  obtenu avec la simulation 2

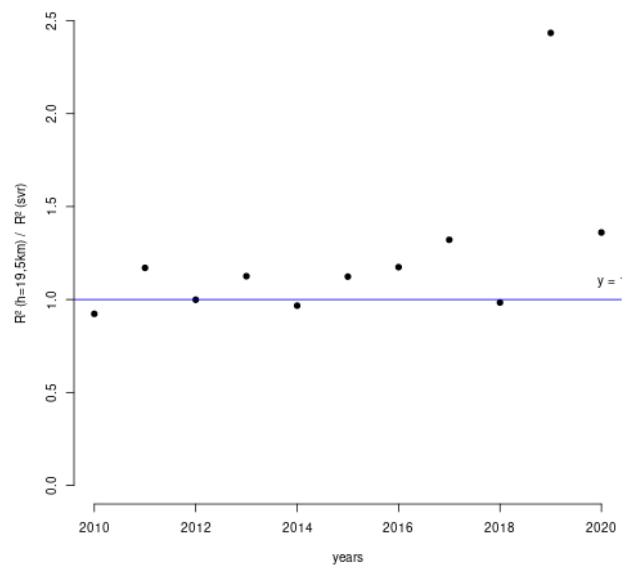


## 12.7 Comparaison des résultats obtenus sur le charançon de la tige du chou

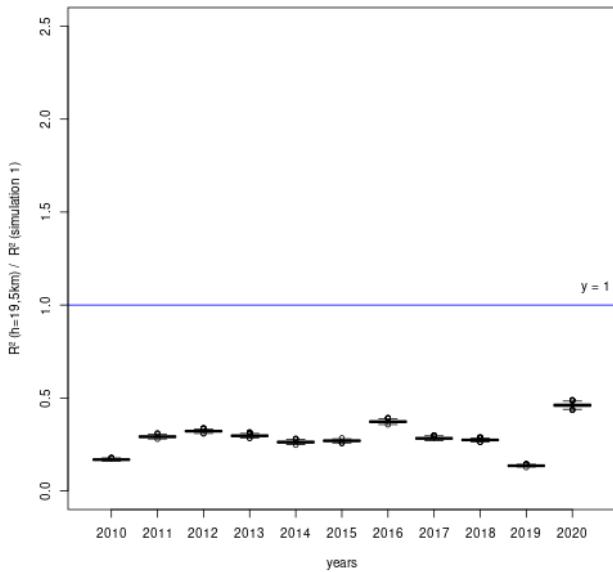
Rapport entre  $R^2$  ( $h=19,5\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2$  ( $k=16$ ) obtenu avec la méthode des  $k$  plus proches voisins



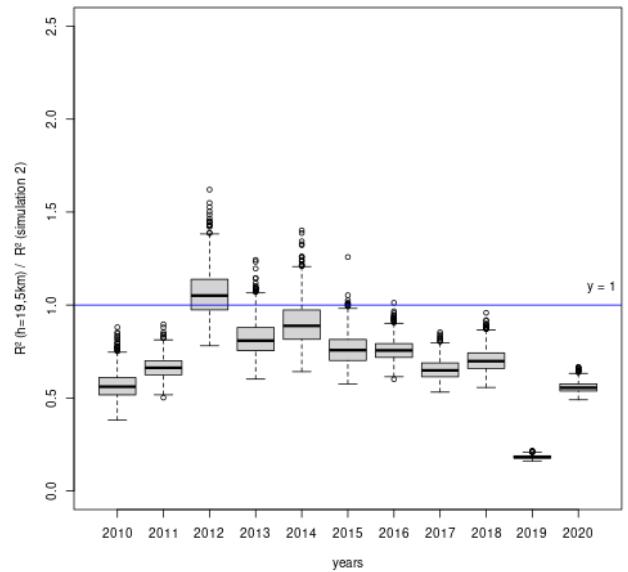
Rapport entre  $R^2$  ( $h=19,5\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2,svr$  par année ajusté sur toutes années avec variable année



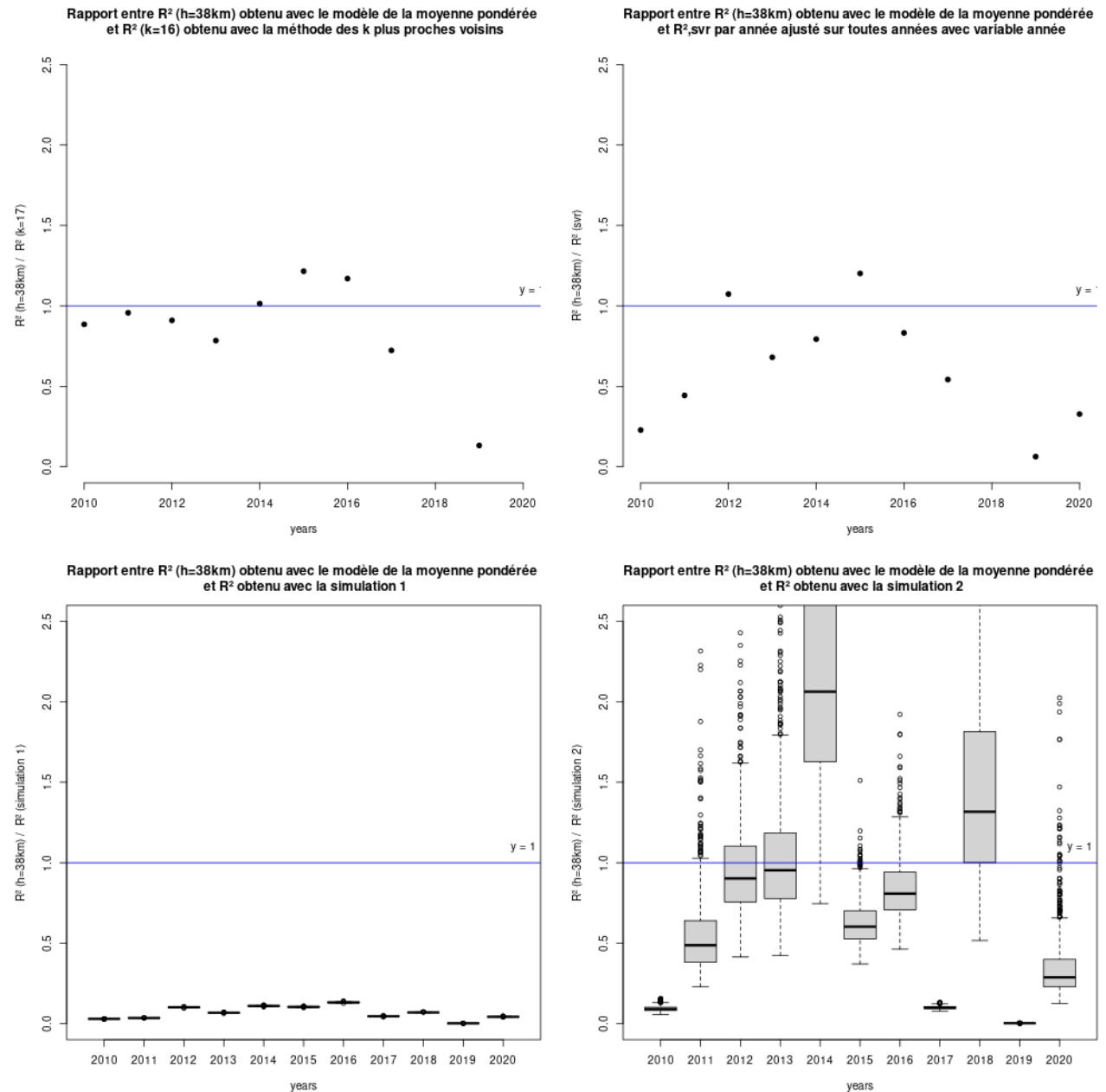
Rapport entre  $R^2$  ( $h=19,5\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2$  obtenu avec la simulation 1



Rapport entre  $R^2$  ( $h=19,5\text{km}$ ) obtenu avec le modèle de la moyenne pondérée et  $R^2$  obtenu avec la simulation 2



## 12.8 Comparaison des résultats obtenus sur l'altise petite des crucifères



### Commentaire sur les résultats

Pour rappel on a essayé de comparer le modèle de la moyenne pondérée avec les autres modèles en faisant le rapport des  $R^2$  correspondant aux observations de chaque année (2010 à 2020). La comparaison du modèle de référence avec la méthode des  $k$  plus proches voisins donnent des valeurs qui sont proches de 1 pour la plupart des années. De même avec le modèle SVR. Cela signifie que les différentes méthodes d'interpolation sont équivalents, avec un léger avantage éventuel à la méthode des moyennes pondérées (par l'exponentiel de l'opposé de la distance).

Pour une année donnée, le modèle à utiliser dépend du rapport des  $R^2$ . Si le rapport est supérieur à 1, il faudra choisir le modèle qui est au dénominateur. Si résultat donne l'inverse, il faut prendre celui qui est au numérateur.

En ce qui concerne la comparaison aux résultats de simulations, on peut voir que les  $R^2$  observés sont autour de 25 % des  $R^2$  obtenus avec les simulations à partir des rapports observés. Le simple fait que le rapport observé soit généré par un faible nombre d'observations diminue de manière conséquente le  $R^2$  possible par rapport à 1. Si l'on prend en compte que les observations sont elles-même le fruit de ce tirage comme on le fait dans la simulation 2, on observe que les  $R^2$  observés sont de l'ordre de 50 % des  $R^2$  simulés. Cela suggère que nos interpolations expliquent de l'ordre de 50 % de la variance non lié au caractère discret des observations. C'est très cohérent avec l'*a priori* agronomique qui considère qu'en général 50 % de la pression de bioagresseurs est déterminé par le rôle du paysage, la météo locale et les pratiques culturales dans le champ.

### 13 Conclusion et perspectives

Des articles scientifiques ont montré que la présence des bioagresseurs dépend de plusieurs facteurs comme :

- facteurs climatiques ;
- degré de fertilité du sol ;
- variétés résistantes ou sensibles : le choix de variétés à semer a un impact sur l'évolution des bioagresseurs. Une variété résistante empêche le développement des bioagresseurs grâce à ses gènes de résistance. Une variété sensible est incapable de stopper le développement de bioagresseurs ;
- position de la parcelle sur la toposéquence ;
- etc...

La pression des bioagresseurs varie fortement en fonction des facteurs que nous venons de citer. Réduire le calcul de la pression de bioagresseurs dans une parcelle en se basant uniquement sur sa proximité distancielle par rapport aux autres parcelles pourrait être insuffisant. Pour améliorer ce dernier, il serait judicieux de trouver un modèle qui tient en compte les facteurs climatiques, le degré de fertilité du sol, la position de la parcelle sur la toposéquence, la variété qui a été semée, etc...

Ce stage a été très enrichissant pour moi, il m'a permis de découvrir le monde de l'agronomie, ses acteurs, ses contraintes. Il m'a permis aussi de mettre en pratique les connaissances théoriques acquises tout au long de ma formation et de découvrir le monde de la recherche. J'ai beaucoup appris de cette expérience, j'en sors grandi avec beaucoup d'espoir.

## **14 Bibliographies**

[1] Terres Inovia, *Les ravageurs du colza*

[2] [https://ecophytopic.fr/sites/default/files/actualites\\_doc/2011-08-Vigicultures\\_Observation\\_triticale\\_proc\\_SIMMONEAU.pdf](https://ecophytopic.fr/sites/default/files/actualites_doc/2011-08-Vigicultures_Observation_triticale_proc_SIMMONEAU.pdf)

[3] <https://duddal.org/files/original/645068a050efe8d00c2948ba8f481185473f3051.pdf>

[4] <http://uved-scv.cirad.fr/co/PressionBioagresseur.html>

[5] <https://www.agrifind.fr/wp-content/uploads/Guide-Agrifind-maladies-et-carences-dubl%C3%A9.pdf>

[6] <https://hal.archives-ouvertes.fr/hal-01986221/document>

[7] [https://arvalis.wedia.fr/file/galleryelement/pj/7d/1f/01/33/473\\_complement%20web\\_%20rmt\\_boite\\_outil7473107969049484787.pdf](https://arvalis.wedia.fr/file/galleryelement/pj/7d/1f/01/33/473_complement%20web_%20rmt_boite_outil7473107969049484787.pdf)

[8] <https://anr.fr/Projet-ANR-19-ECOM-0009>

[5] <https://www.agrifind.fr/wp-content/uploads/Guide-Agrifind-maladies-et-carences-dubl%C3%A9.pdf>

[6] <https://hal.archives-ouvertes.fr/hal-01986221/document>

[7] [https://arvalis.wedia.fr/file/galleryelement/pj/7d/1f/01/33/473\\_complement%20web\\_%20rmt\\_boite\\_outil7473107969049484787.pdf](https://arvalis.wedia.fr/file/galleryelement/pj/7d/1f/01/33/473_complement%20web_%20rmt_boite_outil7473107969049484787.pdf)

[8] <https://anr.fr/Projet-ANR-19-ECOM-0009>