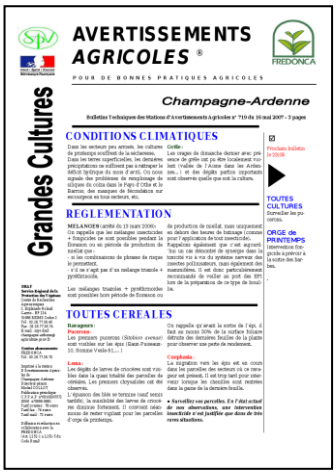
Financement : GIS GC-HP2E

Etude de faisabilité du développement et de la valorisation d'une base de données sur l'évolution des pressions biotiques dans les parcelles agricoles

Coordonnateur : Vincent Cellier (INRA)
Mail : vincent.cellier@epoisses.inra.fr

Autres personnes impliquées :

- Agerberg Julia (INRA),
- Cariolle Michel (ITB),
- Gouache David (Arvalis – Institut du végétal)
- Aubertot Jean-Noël (INRA)


AVERTISSEMENTS AGRICOLES ®
POUR DE BONNES PRATIQUES AGRICOLES
Grandes Cultures
Champagne-Ardenne
CONDITIONS CLIMATIQUES
RÉGLEMENTATION
TOUTES CEREALES

Mots clefs : pressions biotiques, protection des végétaux, avertissements agricoles, base de données documentaire, outil de recherche sémantique.

Résumé

De 1943 à 2008, les Services Régionaux de la Protection des Végétaux (SRPV) ont élaboré des documents destinés à informer les agriculteurs des pressions biotiques s'exerçant sur les cultures. Une étude financée par le GIS GC-HP2E a permis de montrer l'intérêt de collecter et d'analyser les avertissements agricoles, les bilans de campagne nationaux des SRPV et certains documents disponibles dans les instituts techniques agricoles. Le corpus documentaire ainsi constitué sera mis à disposition des utilisateurs avec des outils permettant la fouille et l'extraction d'informations. Des exemples d'études réalisables avec ces informations ont été identifiés. L'étude a également montré l'intérêt de relier les observations sur les pressions biotiques à des données connexes permettant de mieux les exploiter. Des préconisations sont faites afin que ces données connexes soient recensées dans le cadre du dispositif actuel d'épidémiosurveillance.

Contexte et objectif

Les avertissements agricoles et bilans de campagne des Services Régionaux de la Protection des Végétaux représentent une source d'informations extrêmement précieuse, car elles recouvrent une très large gamme de situations de production, sur des séries temporelles longues et sur l'ensemble du territoire national. Ce projet a étudié la faisabilité de la collecte et de la structuration de ces données historiques sur les pressions biotiques et les modalités de leur mobilisation par les acteurs de la recherche-développement. D'autres données des instituts techniques et des chambres d'agriculture, pourraient s'y ajouter. Ces données pourront servir à réaliser des études sur un plus long terme en analysant par exemple l'évolution spatio-temporelle de la répartition des bioagresseurs suivis et en étudiant l'influence de différents facteurs sur cette répartition. Les connaissances ainsi générées pourront avoir des retombées sur les stratégies de protection intégrée des cultures. Elles pourront également apporter des pistes de réflexion sur la manière de recueillir des données dans le cadre de la surveillance biologique du territoire réalisée actuellement.

Méthodes

Le travail d'inventaire des documents a été réalisé à partir d'enquêtes dans plusieurs structures : Services Régionaux de l'Alimentation (SRAl) qui sont dépositaires des documents des SRPV, Bibliothèque nationale de France (BnF), archives nationales, Instituts et Centres Techniques Agricoles (ICTA), Chambres d'Agriculture. Un travail d'inventaire plus poussé a été réalisé dans 3 SRAl (Bourgogne, Champagne-Ardenne et Midi-Pyrénées), ainsi qu'à l'ITB et au Cetiom, afin de recueillir des informations sur la volumétrie des documents, leur contenu, et les problèmes techniques pouvant être rencontrés lors de la numérisation des documents papier ou la récupération des documents électroniques. Le choix des documents a été fait en prenant en compte leur utilité potentielle, leur disponibilité, leur facilité d'inventaire, leur centralisation et leur homogénéité dans le temps afin de pouvoir constituer des séries temporelles.

Un questionnaire en ligne abordant les domaines de valorisation des données et les types de données souhaitées a été diffusé à un public potentiellement intéressé par une telle base documentaire. Des entretiens téléphoniques plus détaillés ont permis de compléter les réponses au questionnaire. Dans certains cas, si un projet précis pouvait être décrit, des fiches-projet ont été élaborées et pourront servir de base pour une première valorisation de la base documentaire.

Ces entretiens ont également permis de mettre en avant les fonctionnalités de la future base de données jugées intéressantes. La réflexion sur la mise à disposition des données ainsi engagée a été complétée par un travail sur la fouille et l'extraction de données avec deux équipes de l'INRA.

Le travail sur les données connexes a été réalisé en identifiant et hiérarchisant les principaux facteurs physiques et biotiques influençant la dynamique de développement des bioagresseurs, grâce à des entretiens avec des experts. Cet inventaire a permis de sélectionner les

Financement : GIS GC-HP2E
éléments principaux à mettre en relation avec les informations sur les pressions biotiques et d'identifier des observations pertinentes qui pourraient être faites à l'avenir sur le terrain par les acteurs de l'épidémiosurveillance du territoire.

Principaux résultats obtenus et applications envisageables, lien au Plan Ecophyto

Choix des documents

L'étude préconise la collecte des avertissements agricoles, bulletins d'alerte de 2 à 3 pages destinés aux agriculteurs et rédigés par les Services Régionaux de la Protection des Végétaux (SRPV), des bilans nationaux rédigés par un rapporteur chargé d'un type de culture et d'un groupe de bioagresseurs donné et de documents du Cetiom (rapports d'activité et bulletins) et de l'ITB (comptes-rendus de travaux et comptes-rendus d'essais).

Collecte et mise à disposition

L'étude préconise une démarche chronologique partant des documents les plus récents et remontant progressivement dans le temps, afin de ne pas avoir de coupure avec la période actuelle. Cela permet également de commencer par des documents informatisés (à partir de 1995), en théorie plus faciles à collecter, avant de s'atteler à la numérisation de documents papier, suivie obligatoirement par une opération de reconnaissance de caractères afin de pouvoir exploiter le contenu des documents.

Valorisations envisagées

On constate un très fort intérêt pour les jeux de données sur les pressions biotiques dans le domaine de la conception et la validation de modèles et d'outils d'aide à la décision. Les personnes enquêtées citent aussi leur volonté de mieux décrire et expliquer les pressions biotiques ou d'explorer les liens entre pressions biotiques et d'autres facteurs, les pratiques agricoles notamment. Au cours de l'étude, des projets pilotes ont été identifiés et pourraient démarrer dès la mise à disposition des documents.

Mise à disposition des documents

La proposition est de mettre à disposition le fonds documentaire avec des outils permettant la fouille des documents afin que chaque utilisateur puisse rechercher les informations dont il a précisément besoin. La plateforme CorTexT, soutenue par l'unité de recherche SenS 1326 de l'Institut Francilien « Recherche, Innovation et Société » (IFRIS) et la technologie Alvis, utilisée par l'équipe Bibliome de l'INRA (Claire Nédellec), pourraient être utilisés.

Les données connexes

Beaucoup d'informations permettant de mieux utiliser les données sur les pressions biotiques ne sont pas accessibles dans les documents. On peut envisager d'utiliser des bases de données existantes, susceptibles d'apporter les informations manquantes mais la localisation étant rarement très précise, seules des données connexes assez générales pourront être obtenues. Ce degré d'information est néanmoins suffisant dans de nombreux cas d'exploitation des documents évoqués par des futurs utilisateurs potentiels.

Perspectives et conclusions

L'étude a montré la pertinence du développement d'une base de données documentaire sur les pressions biotiques dans les parcelles agricoles pour répondre à une large diversité de questions scientifiques et techniques exprimées par les acteurs de la recherche et du développement agricoles. Il serait également intéressant que les propositions concernant les données connexes permettant de mieux exploiter ces sources d'information soient étudiées par les structures en charge de la surveillance biologique du territoire et de l'élaboration des Bulletins de Santé du Végétal. Enfin, cette étude a été réalisée dans le cadre du GIS GC-HP2E et ses résultats ne concernent donc que les grandes cultures, même si certains documents sont communs à d'autres filières. Il semble évident que la démarche devra être étendue aux autres filières.

Il ne sera sans doute pas possible de collecter l'ensemble des fonds documentaires mentionnés dans cette étude au cours d'un seul projet et il semble judicieux de commencer par les documents des SRPV qui représentent un volume important, assez homogène et bien connu des utilisateurs potentiels. Une collecte, même partielle, de ce fonds permettrait de réaliser une première version de la base documentaire et de ses outils d'extraction d'informations et de montrer l'intérêt de poursuivre le travail de collecte documentaire

Références bibliographiques

Fussler L., Kobes N., Bertrand F., Maumy M., Grosman J., and Savary S. 2008. A Characterization of Grapevine Trunk Diseases in France from Data Generated by the National Grapevine Wood Diseases Survey. Phytopathology 98:571-579

Jeger M.J. et Pautasso M. 2008. Plant disease and global change – the importance of long-term data sets. New Phytologist 177: 8–11

Zwankhuizen M. J., Zadoks J. C. 2002. *Phytophthora infestans*'s 10-year truce with Holland: a long-term analysis of potato late-blight epidemicsin the Netherlands. *Plant Pathology* 51 : 413–423.