

CATI-CODEX groupe *Math Lab*

Pré-projet « Guns N Roses »

Contexte : Modèles biologiques pour la santé des plantes

Réseau bayésien

Un réseau bayésien B peut se définir comme un graphe G et un ensemble de probabilités θ pour chaque nœud du graphe conditionnellement à l'état de ses parents dans ce graphe.

On associe chaque élément du graphe à une variable aléatoire.

$B = (G, \theta)$ est défini par :

- $G = (X, E)$, graphe acyclique orienté dont les nœuds sont associés à un ensemble de variables aléatoires $X = \{X_1, \dots, X_n\}$.

- $\theta = \{P(X_i | P a(X_i))\}$, ensemble des probabilités de chaque nœud X_i conditionnellement à l'état de ses parents $P a(X_i)$ dans G .



$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}.$$

Exemple simple utilisant l'inférence



Œillet Sinensis
Œillet Barbatus



Rouge
Jaune
Blanc

Connaissant la fleur, on connaît la probabilité de la couleur *

	Fleur Sinensis	Barbatus
Couleur Rouge	0.7	0.05
Couleur Jaune	0.29	0.1
Couleur Blanc	0.01	0.85

$$\forall j, \sum_{i=0}^{i=2} P_{ij} = 1$$

Ensuite, sachant qu'il y a dans un panier autant d'œillets Sinensis que d'œillets Barbatus, déterminer la variété d'une fleur tirée au hasard connaissant sa couleur.

Exemple avec les valeurs simples ci-dessus :

Si couleur jaune $P(\text{Sinensis}) = 0,744$ et $P(\text{Barbatus})=0.256$

Si couleur rouge $P(\text{Sinensis})= 0.933$ et $P(\text{Barbatus})=0.067$

Si couleur blanche $P(\text{Sinensis}) = 0.012$ $P(\text{Barbatus})=0.988$

La protection biologique intégrée est un métier où il faut ...



Diagnostiquer et suivre les infestations



Evaluer les risques sur la production



Inventorier la faune auxiliaire



Envisager un moyen d'action

Où ? (local, généralisé)

Quand ? (immédiat, attendre)

Comment ? (produit, auxiliaires, autres)



puis ... DECIDER

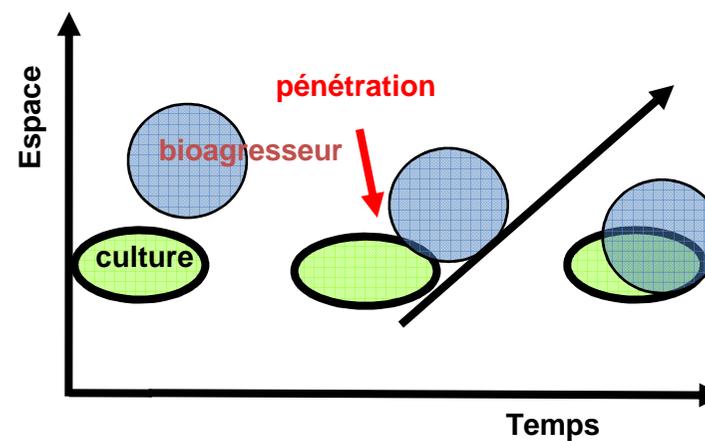
Le professionnel fait face à un contexte de contraintes multifactorielles

Plusieurs bioagresseurs aux mêmes endroits en même temps

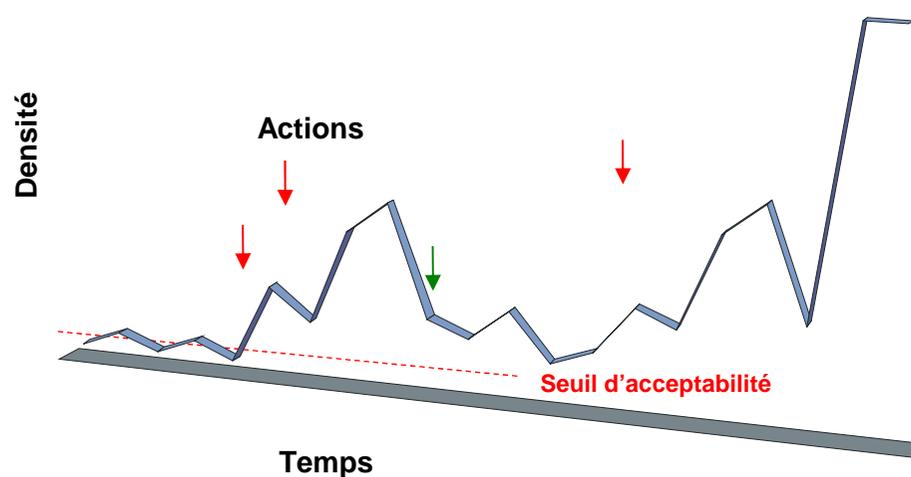


acarions aleurodes pucerons thrips oïdium botrytis

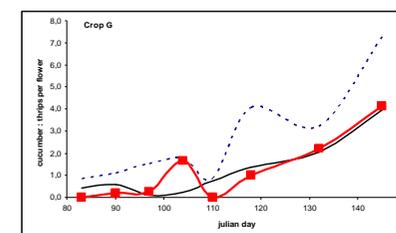
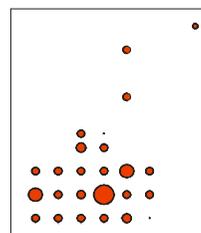
l'objectif : gérer les bioagresseurs dès leur pénétration dans la serre et pendant toute la durée de la culture



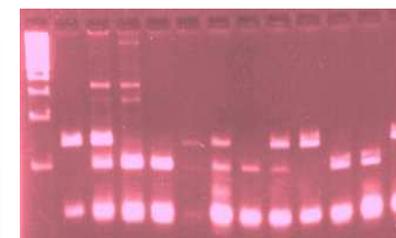
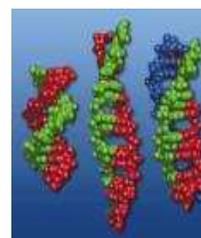
Q. De DDP : Quelle sera l'évolution des infestations ?



Q. De DDP : Quand et où traiter ?



Q. De GDP : Comment traiter ?



La représentation des classes « visuelles » d'abondance à prévoir à t+1

Classe 1 : absence



symbolisée « . » sur les cartographies, cette classe est essentielle pour la validité des cartographies prévisionnelles.

Classe 2 : présence



symbolisée ○ sur les cartographies, cette classe impacte les actions de surveillance et/ou d'intervention.

Classe 3 : beaucoup



symbolisée ● sur les cartographies, cette classe représente les foyers de contamination et de diffusion : niveau critique avec dégâts.

Classe 4 : entoilage



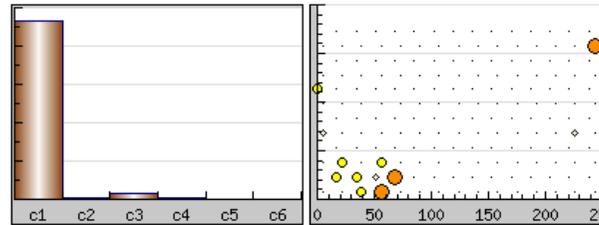
symbolisée ● sur les cartographies, cette classe traduit l'échec de la protection localement.

La prévision d'une classe d'abondance à t+7 sur chaque observation de l'échantillon enrichira considérablement l'information contribuant à la prise de décision globale ou localisée : action écophyto.

La prévision à court terme des infestations locales est inscriptible aux actions de monitoring PBI

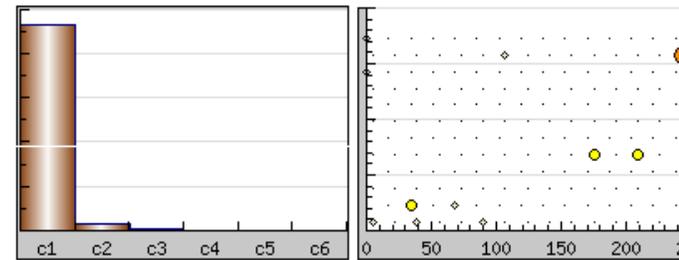
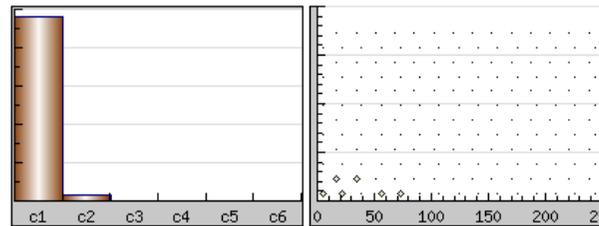
Approche déterministe : exemple l'acarien du rosier

OBSERVED ON DATE 2011-09-14

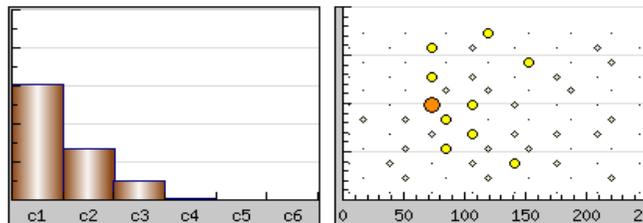


Réel du 2011-09-21

FORECASTED FOR NEXT WEEK

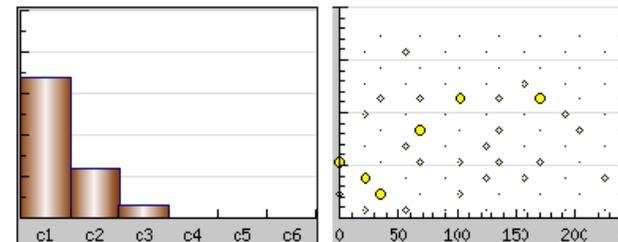
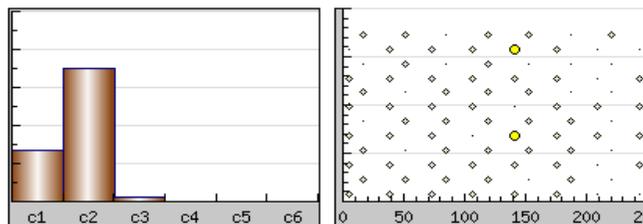


OBSERVED ON DATE 2004-05-11



Réel du 2004-05-18

FORECASTED FOR NEXT WEEK



Ce sont les jeux de données observées qui permettent d'améliorer les modèles de ce type

Approche déterministe : exemple l'acarien du rosier

Principe :

1- accumulation de données dites de « référence/entraînement »

2- construction d'un modèle prévisionnel de type « boîte noire »

Parmi près de 300 variables (biologiques, physiques et interventions), on cherche la meilleure combinaison permettant de prévoir la classe d'abondance du ravageur, SANS CHERCHER A EXPLIQUER, à cette étape, le détail des interactions entre ces variables.

3- test avec des données indépendantes dites de « supplémentaires/validation »



Pas de temps utilisé : la semaine, unité de prévision : chaque point d'échantillonnage, unité de précision : la classe d'abondance

Prédicteurs significatifs pour le modèle de régression non linéaire (PPR model)	SEI _{ref}	SEI _{supp}	Prédicteurs significatifs pour le modèle de régression logistique (POLR model)	SEI _{ref}	SEI _{supp}
P1: Taux de classes 3 <u>sur l'ensemble de l'échantillon</u> pour les acariens sur tige à la date J	0.409	0.626	P1: Classe d'abondance des acariens sur poumon de l'unité observée, à la date J-7	0.410	0.594
P2: Classe d'abondance des acariens sur poumon de l'unité observée, à la date J-7	0.396	0.550	P2: Classe d'abondance des acariens sur poumon sur les zones adjacentes de l'unité observée, à la date J	0.402	0.578
P3: Classe d'abondance des acariens sur tige sur les zones adjacentes de l'unité observée, à la date J	0.392	0.558	P3: Classe d'abondance pour l'oïdium sur tige à la date J-7	0.401	0.587
P4: Taux de classes 1 <u>sur l'ensemble de l'échantillon</u> pour les acariens sur tige à la date J-7	0.384	0.556	P4: Classe d'abondance des acariens sur les inter-zones échantillonnées, à la date J	0.400	0.588
P5: Δ pour les moyennes diurnes journalières de l'humidité relative extérieures (J-14 à J-7), codée en classe	0.379	0.550			

Bons résultats d'ensemble, mais échecs de prévision pour les situations perturbées par les conséquences des traitements phytosanitaires probablement mal pris en compte par le modèle.

Approche probabiliste : exemple l'acarien du rosier → Gun & Roses

Principes :

- 1- s'affranchir des observations pour s'appuyer sur l'expertise du domaine
- 2- conception d'un réseau bayésien pour inférer la **variable de sortie** au temps t+7 : classe d'abondance de l'acarien
Pas de temps utilisé : **la semaine**, unité de prévision : **chaque tige échantillonnée**, unité de mesure : **la classe d'abondance**
- 3- implémentation des tables de probabilité avec l'expertise
- 4- test du modèle avec des données d'exploitation commerciale
- 5- dégager une genericité pour la construction de modèles similaires sur d'autres exemples

Etapes :

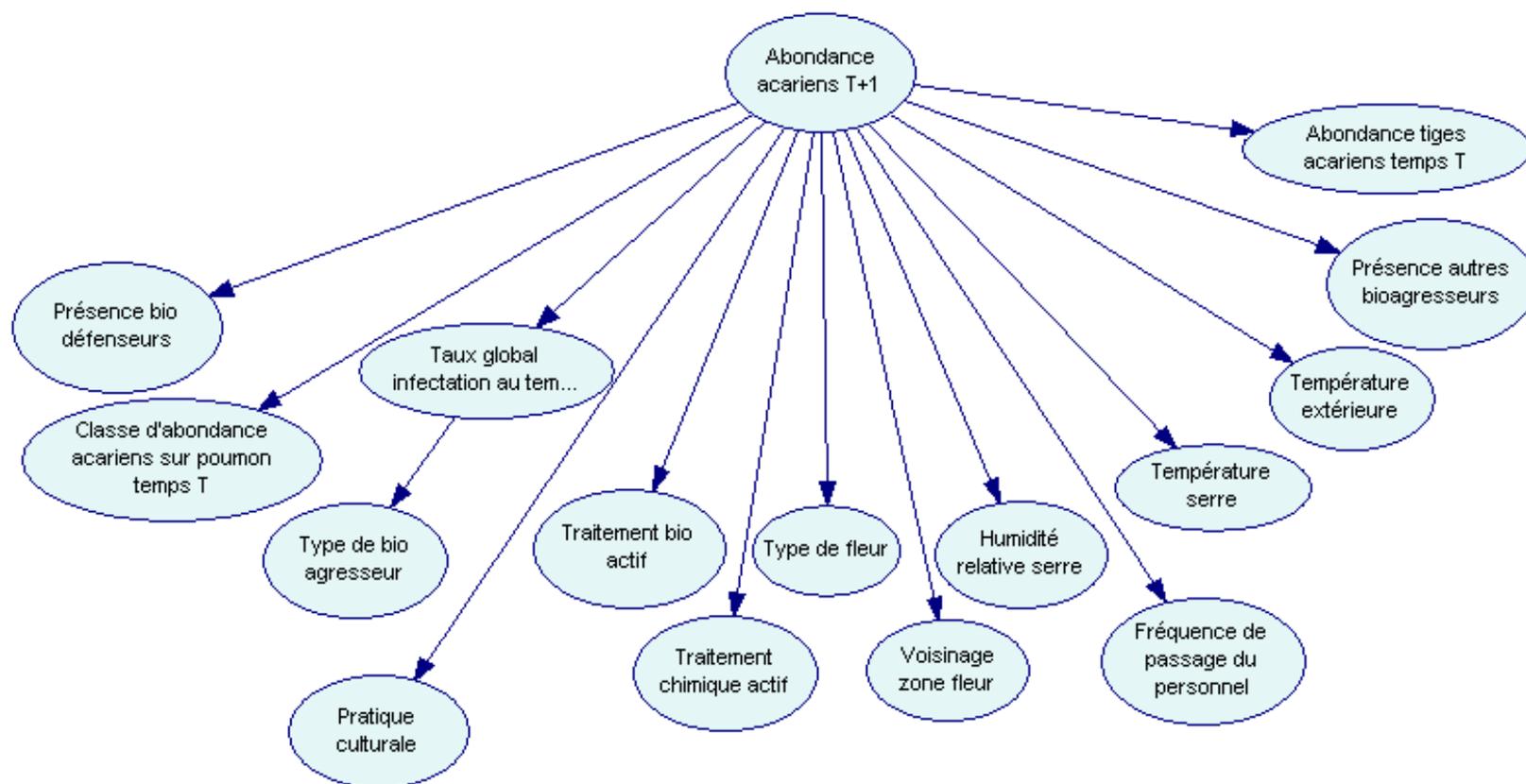
- 1- constituer un référentiel de questionnement des experts
- 2- recueillir et structurer la connaissance au niveau des tables de probabilité
- 3- paramétrer le modèle pour les outils informatiques existants
- 4- valider le modèle au moyen de d'échantillonnages de terrain

A terme, la variable de sortie devrait être la qualité commerciale de la tige observée en fonction du cumul des prévisions de tous les bioagresseurs potentiels : réseau bayésien final.

Schéma de travail du réseau bayésien « acariens sur rosier »

Guns n Roses

13/01/2014



Réseau bayésien
Ph. Abbal,
Roger Boll
Eric Latrille