

2015

Paris

Évaluation de règles de décision par analyse ROC

David Makowski

UMR 211 INRA AgroParisTech
Thiverval-Grignon

Règle de décision

« Règle permettant de **prendre une décision** en fonction **d'informations disponibles** au moment de la **décision** ».

Pourquoi évaluer ?

- **Plusieurs règles** souvent disponibles **pour un usage donné** (exemple: raisonnement de la fertilisation azotée, raisonnement du traitement contre le sclérotinia du colza...).
- **Toutes les règles font des erreurs.**
- La performance d' une règle **ne peut être déduite** de son niveau de **complexité.**
- L' évaluation des performances des règles de décision est utile pour les utilisateurs mais **aussi pour les concepteurs.**

Différents types de règle de décision existent

Les règles se différencient selon:

- Le nombre de variables décisionnelles (une, deux, ...).
- La nature des variables décisionnelles (binaires, continues...).
- Le type d'information utilisé (mesures, prédictions de modèles).
- Le nombre et la valeur des paramètres (seuils de décision...).

Règle de décision binaire

« **Règle** permettant de choisir parmi **deux décisions** ».

Exemples de décisions binaires

- **Traiter / Ne pas traiter contre une maladie.**
- **Traiter / Ne pas traiter contre des mauvaises herbes.**

Mais aussi

- **Semer la variété 1 / Semer la variété 2.**
- **Troisième apport d'engrais / Pas de troisième apport.**

Règles de décision « non binaires »

- **Exemples: règles pour calculer des doses d'engrais**
 - méthodes du bilan,
 - modèles statiques,
 - modèles dynamiques.
- **Des méthodes existent aussi pour évaluer ces règles.**

Règle de décision binaire basée sur un indicateur et un seuil de décision

« Je traite si mon **indicateur** est supérieur/inférieur à un **seuil de décision** ».

« Je traite si $I \geq S$. Je ne traite pas sinon ».

- Indicateur I = mesure ou prédiction d'un modèle (ex: % plantes malades).
- Seuil S = Valeur numérique (ex: 20%).

Évaluation de

« Je traite si $I \geq S$. Je ne traite pas sinon ».

Les deux aspects du problème:

- Choisir le meilleur seuil S pour un indicateur I donné.
- Choisir le meilleur indicateur parmi plusieurs indicateurs disponibles.

Étapes de l'évaluation

1. Définir une série d'indicateurs (mesures et/ou modèles).
2. Définir la gamme de variation possible pour le seuil S associé à chaque indicateur (e.g 0-100 % de fleurs malades).
3. Définir un ou plusieurs critères pour évaluer les règles de décision (i.e les combinaisons entre I et S).
4. Estimer les valeurs de ces critères pour chaque règle.
5. Choisir la « meilleure » règle.

Etape 1. Définir une série d'indicateurs (mesures et/ou modèles)

- **Exemple 1: Traitement du piétin échaudage du blé d'hiver.**
- **Exemple 2: Traitement du sclérotinia du colza.**
- **Exemple 3: Traitement des mauvaises herbes sur colza.**

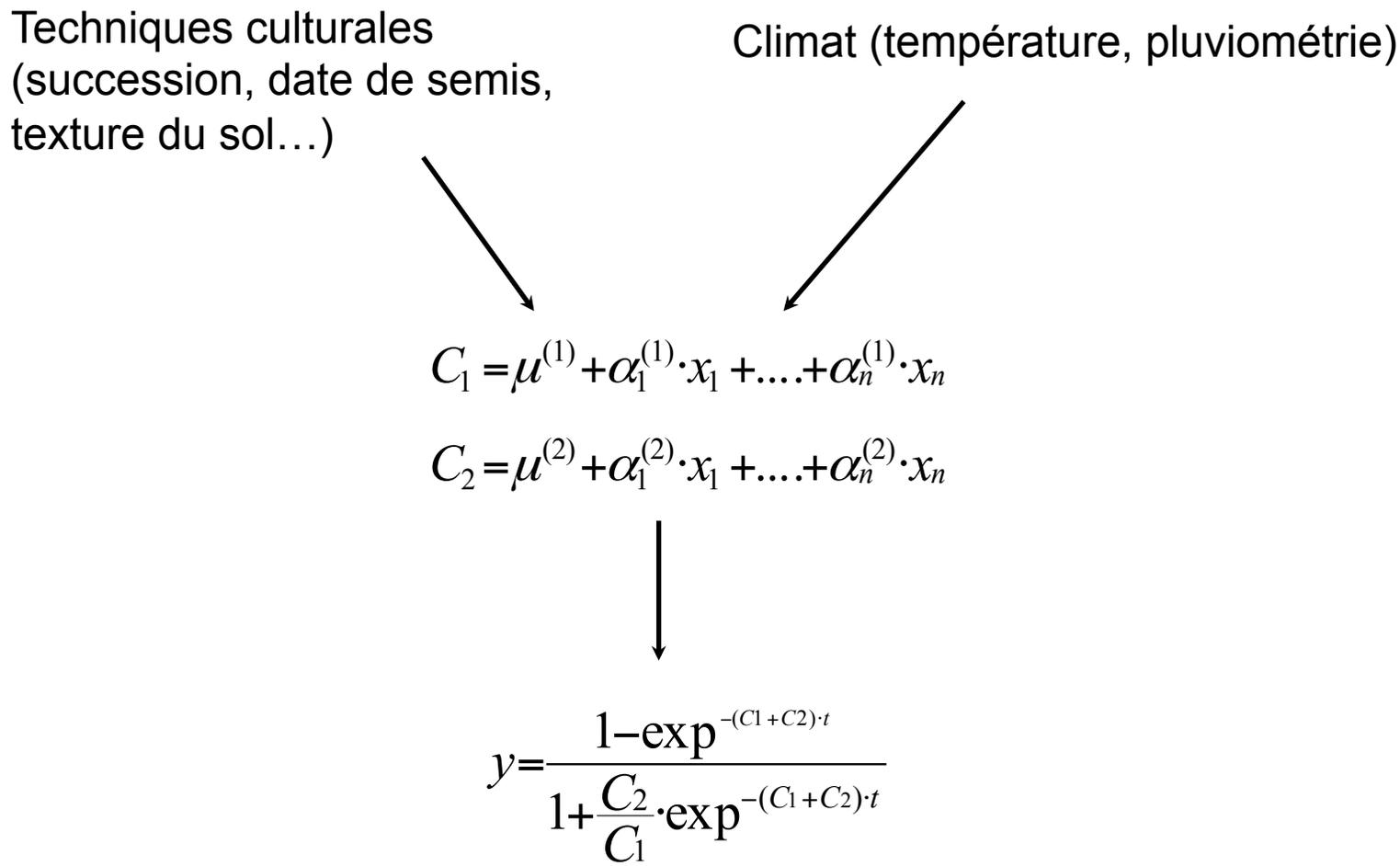
Seize modèles simulant le % plantes attaquées par le piétin échaudage (Ennaïfar et al., 2007)

| | Name | Type | Equation | Fsem30 ^z | Selection |
|----|--------------------|----------------|-----------------------|---------------------|-----------------------------|
| 1 | DynL0 ⁻ | Dynamic | Linear | Absent | Bibliographic |
| 2 | DynL0 ⁺ | Dynamic | Linear | Absent | Statistical + Bibliographic |
| 3 | DynL1 ⁻ | Dynamic | Linear | Present | Bibliographic |
| 4 | DynL1 ⁺ | Dynamic | Linear | Present | Statistical + Bibliographic |
| 5 | DynM0 ⁻ | Dynamic | Multiplicative | Absent | Bibliographic |
| 6 | DynM0 ⁺ | Dynamic | Multiplicative | Absent | Statistical + Bibliographic |
| 7 | DynM1 ⁻ | Dynamic | Multiplicative | Present | Bibliographic |
| 8 | DynM1 ⁺ | Dynamic | Multiplicative | Present | Statistical + Bibliographic |
| 9 | StaL0 ⁻ | Static | Linear | Absent | Bibliographic |
| 10 | StaL0 ⁺ | Static | Linear | Absent | Statistical + Bibliographic |
| 11 | StaL1 ⁻ | Static | Linear | Present | Bibliographic |
| 12 | StaL1 ⁺ | Static | Linear | Present | Statistical + Bibliographic |
| 13 | StaM0 ⁻ | Static | Multiplicative | Absent | Bibliographic |
| 14 | StaM0 ⁺ | Static | Multiplicative | Absent | Statistical + Bibliographic |
| 15 | StaM1 ⁻ | Static | Multiplicative | Present | Bibliographic |
| 16 | StaM1 ⁺ | Static | Multiplicative | Present | Statistical + Bibliographic |

Un des 16 modèles simulant le % plantes attaquées par le piétin échaudage (Ennaïfar et al., 2007)

Techniques culturales
(succession, date de semis,
texture du sol...)

Climat (température, pluviométrie)



$C_1 = \mu^{(1)} + \alpha_1^{(1)} \cdot x_1 + \dots + \alpha_n^{(1)} \cdot x_n$

$C_2 = \mu^{(2)} + \alpha_1^{(2)} \cdot x_1 + \dots + \alpha_n^{(2)} \cdot x_n$

$$y = \frac{1 - \exp^{-(C_1 + C_2) \cdot t}}{1 + \frac{C_2}{C_1} \cdot \exp^{-(C_1 + C_2) \cdot t}}$$

Trois indicateurs pour raisonner le traitement du sclérotinia du colza

- **Pourcentage de fleurs malades parmi N fleurs prélevées dans une parcelle.**
- **Niveau de risque prédit par un modèle qui ne tient pas compte du climat.**
- **Niveau de risque prédit par un modèle qui tient compte du climat.**

Niveau de risque de sclérotinia prédit par un modèle qui ne tient pas compte du climat

| Risk factor | Level | Points |
|--|----------|--------|
| Number of oil-seed crops during the last ten years | >5 | 30 |
| | 3-5 | 20 |
| | 2-3 | 10 |
| | 1 | 0 |
| Other host crops during the last five years | Yes | 15 |
| | No | 0 |
| Level of infection in the last crop | High | 15 |
| | Moderate | 5 |
| | Low | 0 |
| Type of field | Wet | 10 |
| | Dry | 0 |
| Plant density | High | 10 |
| | Normal | 5 |
| | Low | 0 |

Vingt modèles simulant le risque d'infestation du colza par les mauvaises herbes (Primot *et al.*, 2006)

| Model | Equation | Variables | Barralis grid | Threshold of weed biomass at winter (t ha ⁻¹) |
|-----------------------|----------|------------|---------------|---|
| <i>LIN.weed</i> | Linear | Weed | No | - |
| <i>LIN.tech</i> | Linear | Techniques | No | - |
| <i>LIN.all</i> | Linear | All | No | - |
| <i>LIN.selec</i> | Linear | Selected | No | - |
| <i>LIN.selec.b</i> | Lineal | Selected | Yes | - |
| <i>LOG.weed.10</i> | Logistic | Weed | No | 0.10 |
| <i>LOG.tech.10</i> | Logistic | Techniques | No | 0.10 |
| <i>LOG.all.10</i> | Logistic | All | No | 0.10 |
| <i>LOG.selec.10</i> | Logistic | Selected | No | 0.10 |
| <i>LOG.selec.10.b</i> | Logistic | Selected | Yes | 0.10 |
| <i>LOG.weed.15</i> | Logistic | Weed | No | 0.15 |
| <i>LOG.tech.15</i> | Logistic | Techniques | No | 0.15 |
| <i>LOG.all.15</i> | Logistic | All | No | 0.15 |
| <i>LOG.selec.15</i> | Logistic | Selected | No | 0.15 |
| <i>LOG.selec.15.b</i> | Logistic | Selected | Yes | 0.15 |
| <i>LOG.weed.20</i> | Logistic | Weed | No | 0.20 |
| <i>LOG.tech.20</i> | Logistic | Techniques | No | 0.20 |
| <i>LOG.all.20</i> | Logistic | All | No | 0.20 |
| <i>LOG.selec.20</i> | Logistic | Selected | No | 0.20 |
| <i>LOG.selec.20.b</i> | Logistic | Selected | Yes | 0.20 |

Modèle logistique pour prédire le risque d'infestation par des mauvaises herbes

$$P(y > D_{thresh}) = \frac{\exp\left(\alpha_0 + \sum_{i=1}^p \alpha_i x_i\right)}{1 + \exp\left(\alpha_0 + \sum_{i=1}^p \alpha_i x_i\right)}$$

y : biomasse de mauvaise herbe en hiver.

D_{thresh} : Seuil de nuisibilité.

x_1, \dots, x_p : variables explicatives.

$\alpha_0, \dots, \alpha_p$: paramètres à estimer.

Modèle linéaire pour prédire le risque d'infestation par des mauvaises herbes

$$y = \alpha_0 + \sum_{i=1}^p \alpha_i x_i + \varepsilon$$

y : biomasse de mauvaise herbe en hiver.

x_1, \dots, x_p : variables explicatives.

$\alpha_0, \dots, \alpha_p$: paramètres à estimer.

Variables d'entrée

| Type of the variable | Variables | Minimum | Maximum | Unit |
|----------------------|---|------------------|-------------------------|------------------------------------|
| Techniques | Sowing Date (SD) | 27 July (207) | 7 September (249) | Days since 1 st January |
| Techniques | Soil mineral Nitrogen at emergence (N) | 26 | 317 | kg ha ⁻¹ |
| Techniques | Oilseed rape Density at emergence (OD) | 14 | 232 | Plants.m ⁻² |
| Techniques | Soil Tillage (ST) | 0 (deep tillage) | 1 (superficial tillage) | |
| Weeds | Weed Density at emergence (WD) | 1 | 532 | Plants.m ⁻² |
| Weeds | Type of the most abundant weed at emergence | Type 1 | Type 7 | See weed typology in Table 2 |

Variables d'entrée (suite)

| Type | Group | Time of emergence of the weed | Height of weed relative to rapeseed |
|------|------------------|-------------------------------|-------------------------------------|
| 1 | Monocotyledonous | Summer | Equal or lower |
| 2 | Monocotyledonous | Indifferent | Equal or lower |
| 3 | Dicotyledonous | Summer | Equal or lower |
| 4 | Dicotyledonous | Summer | Higher |
| 5 | Dicotyledonous | Autumn | Equal or lower |
| 6 | Dicotyledonous | Indifferent | Equal or lower |
| 7 | Dicotyledonous | Indifferent | Higher |

Étapes de l'évaluation

1. Définir une série d'indicateurs (mesures et/ou modèles).
2. Définir la gamme de variation possible pour le seuil S associé à chaque indicateur (0, 0.1, ..., 0.9, 1).
3. Définir un ou plusieurs critères pour évaluer les règles de décision (*i.e* les combinaisons entre I et S).
4. Estimer les valeurs de ces critères pour chaque règle.
5. Choisir la « meilleure » règle.

Étape 3. Définir un ou plusieurs critères pour évaluer les règles de décision (*i.e* les combinaisons entre *I* et *S*).

- **Évaluer la qualité prédictive des modèles (MSEP).**
- **Évaluer le résultat de l'application de la règle de décision.**
 - **Marge brute moyenne obtenue en appliquant la règle.**
 - **Analyse ROC (sensibilité, spécificité).**

Étape 4. Estimer les valeurs des critères pour chaque règle.

Estimation de l'erreur moyenne de prédiction

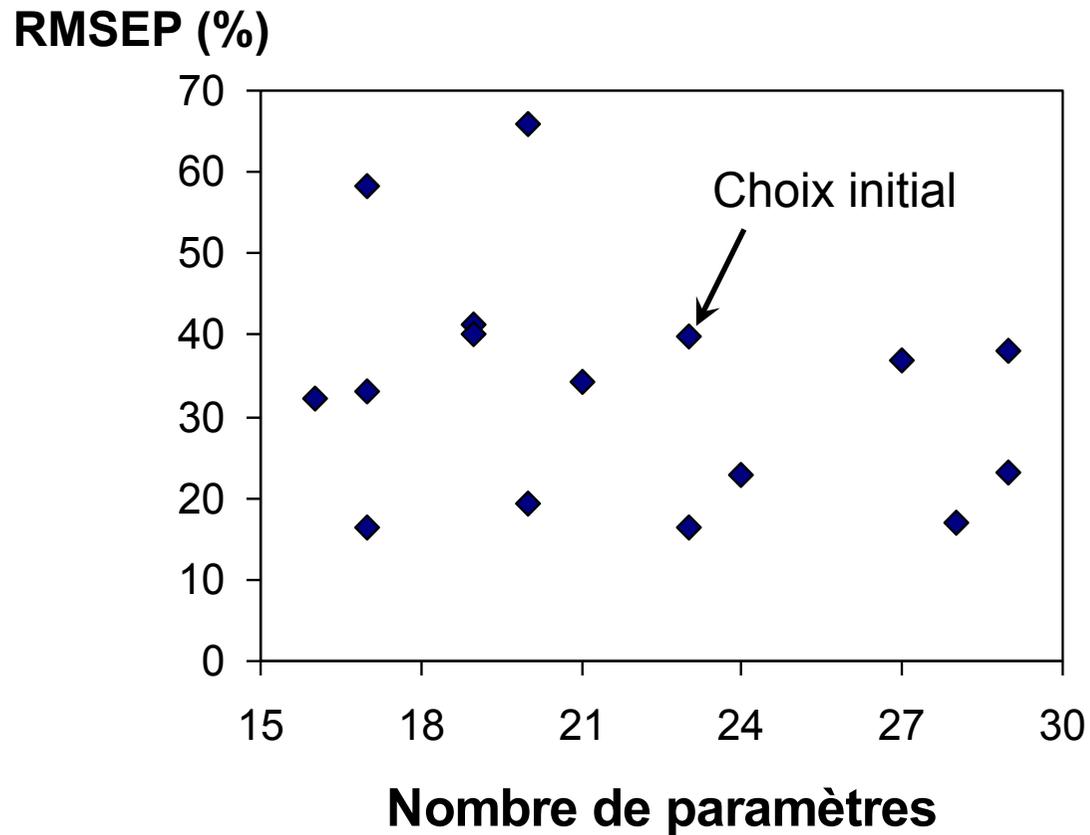
- **Utilisation de données indépendantes.**
- **Validation croisée...**

Application aux seize modèles simulant le % plantes attaquées par le piétin échaudage (Ennaïfar et al., 2007)

| | Name | Type | Equation | Fsem30 ^z | Selection |
|----|--------------------|----------------|-----------------------|---------------------|-----------------------------|
| 1 | DynL0 ⁻ | Dynamic | Linear | Absent | Bibliographic |
| 2 | DynL0 ⁺ | Dynamic | Linear | Absent | Statistical + Bibliographic |
| 3 | DynL1 ⁻ | Dynamic | Linear | Present | Bibliographic |
| 4 | DynL1 ⁺ | Dynamic | Linear | Present | Statistical + Bibliographic |
| 5 | DynM0 ⁻ | Dynamic | Multiplicative | Absent | Bibliographic |
| 6 | DynM0 ⁺ | Dynamic | Multiplicative | Absent | Statistical + Bibliographic |
| 7 | DynM1 ⁻ | Dynamic | Multiplicative | Present | Bibliographic |
| 8 | DynM1 ⁺ | Dynamic | Multiplicative | Present | Statistical + Bibliographic |
| 9 | StaL0 ⁻ | Static | Linear | Absent | Bibliographic |
| 10 | StaL0 ⁺ | Static | Linear | Absent | Statistical + Bibliographic |
| 11 | StaL1 ⁻ | Static | Linear | Present | Bibliographic |
| 12 | StaL1 ⁺ | Static | Linear | Present | Statistical + Bibliographic |
| 13 | StaM0 ⁻ | Static | Multiplicative | Absent | Bibliographic |
| 14 | StaM0 ⁺ | Static | Multiplicative | Absent | Statistical + Bibliographic |
| 15 | StaM1 ⁻ | Static | Multiplicative | Present | Bibliographic |
| 16 | StaM1 ⁺ | Static | Multiplicative | Present | Statistical + Bibliographic |

| | Nom du modèle | Nombre de paramètres | RMSEP (%) | Biais (%) |
|----|--------------------------|-----------------------|--------------|-----------|
| 1 | DynL0 ⁻ | 18+11=29 ^z | 38.1 | 1.17 |
| 2 | DynL0 ⁺ | 14+7=21 | 34.3 | 1.54 |
| 3 | DynM0 ⁻ | 17+10=27 | 36.98 | 5.44 |
| 4 | DynM0 ⁺ | 15+8=23 | 39.68 | 5.96 |
| 5 | StaL0⁻ | 19 | 41.3 | 0.41 |
| 6 | StaL0 ⁺ | 16 | 32.1 | 0.64 |
| 7 | StaM0⁻ | 19 | 40.10 | 7.08 |
| 8 | StaM0 ⁺ | 17 | 33.18 | 7.15 |
| 9 | DynL1⁻ | 17+11=28 | 16.9 | 0.54 |
| 10 | DynL1⁺ | 15+8=23 | 16.3 | 0.48 |
| 11 | DynM1 ⁻ | 17+12=29 | 23.23 | 4.53 |
| 12 | DynM1 ⁺ | 16+8=24 | 22.83 | 4.87 |
| 13 | StaL1⁻ | 20 | 19.4 | 0.24 |
| 14 | StaL1⁺ | 17 | 16.4 | 0.20 |
| 15 | StaM1⁻ | 20 | 65.93 | 2.96 |
| 16 | StaM1⁺ | 17 | 58.22 | 2.82 |

Evaluation des erreurs de 16 modèles prédisant l' incidence du piétin échaudage du blé au stade GS33.



Analyse ROC

Notations

Y : une variable aléatoire prenant les valeurs 0 ou 1 pour une réponse négative ou positive respectivement.

I : une variable continue correspondant à un indicateur.

S : un seuil de décision.

Exemples

$Y = 0$ si perte de rendement due à la maladie = 0, $Y=1$ sinon.

$Y = 0$ si pourcentage de plante malade à la récolte < 10%, $Y=1$ sinon.

$Y = 0$ si biomasse de mauvaise herbe < 0.15 t/ha, $Y=1$ sinon.

Deux types d'erreurs de classification

Faux positif: $I > S$ mais $Y=0$

Faux négatif: $I < S$ mais $Y=1$

| | | | |
|------------------------------------|------------|-------------------|--------------|
| | | Décision optimale | |
| | | $Y = 0$ | $Y = 1$ |
| Décision basée sur le modèle | $I < S$ | Vrai négatif | Faux négatif |
| | $I \geq S$ | Faux positif | Vrai positif |

Le but de l'analyse ROC est d'estimer les fréquences de faux négatif et de faux positif pour tous les seuils de décision S .

Analyse ROC

n parcelles avec $Y=0$ (e.g. biomasse mauv. herbes $< 0.15 \text{ t.ha}^{-1}$).

m parcelles avec $Y=1$ (e.g. biomasse mauv. herbes $> 0.15 \text{ t.ha}^{-1}$).

(i). Calcul de l'indicateur I pour chaque parcelle.

(ii). Définition d'un seuil de décision, I_T .

(iii). **Sensibilité** = $Prob(I \geq I_T | Y=1)$

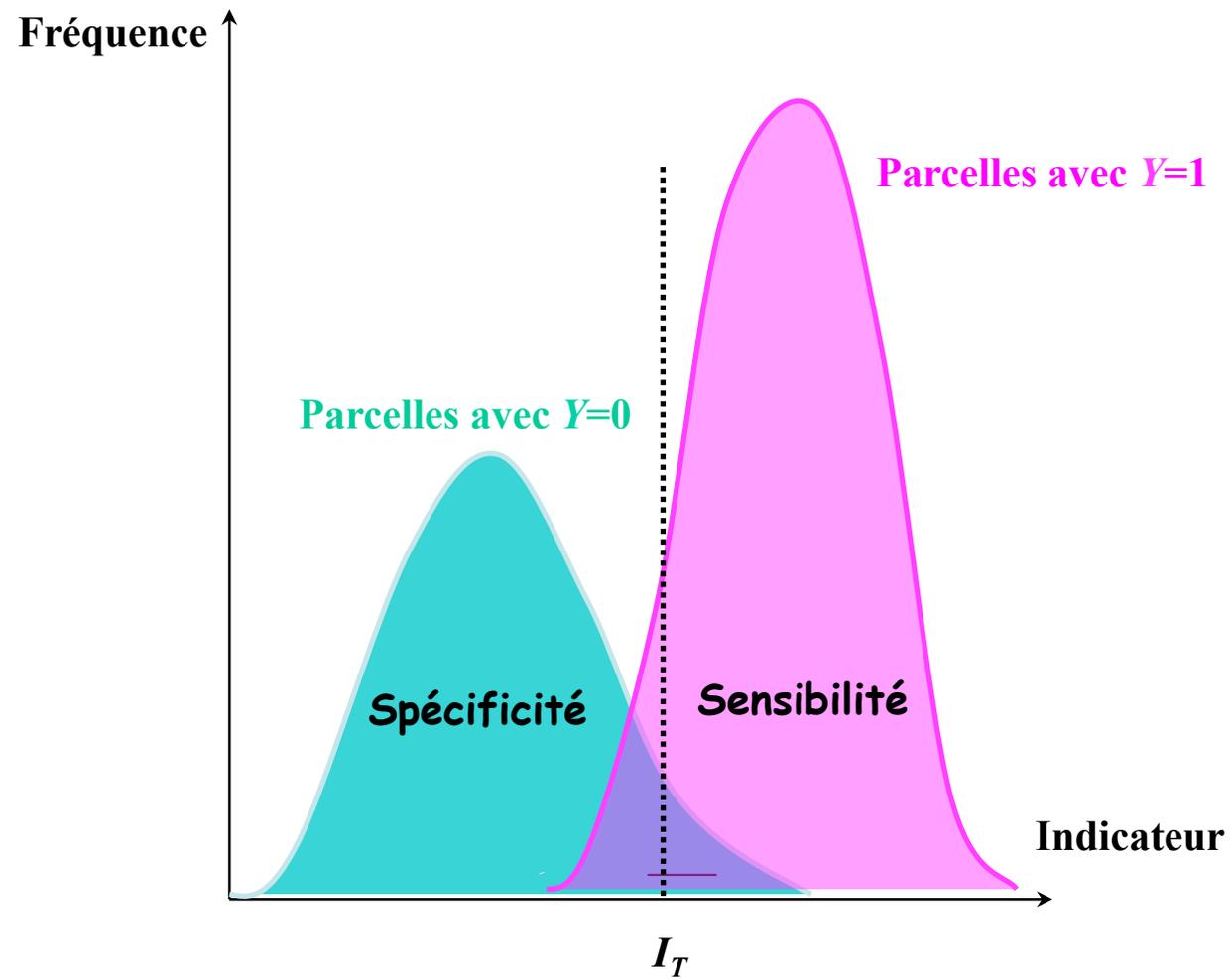
(iv). **Spécificité** = $Prob(I < I_T | Y=0)$

(v). **Courbe ROC**: Sensibilité (I_T) versus $1 - \text{Spécificité}$ (I_T)

(vi). Estimation la surface sous la courbe (**AUC**) pour chaque indicateur I .

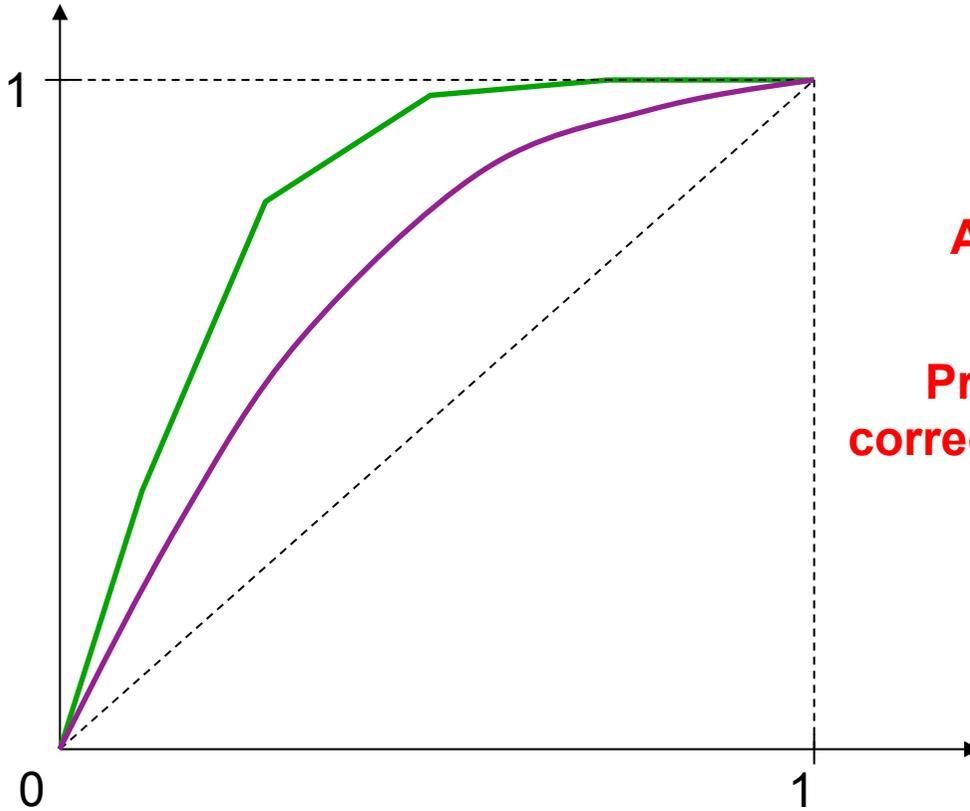
Si $AUC \sim 0.5$, l'indicateur n'est pas mieux qu'une décision aléatoire.

Analyse ROC



Analyse ROC

Sensibilité



Aire sous la courbe

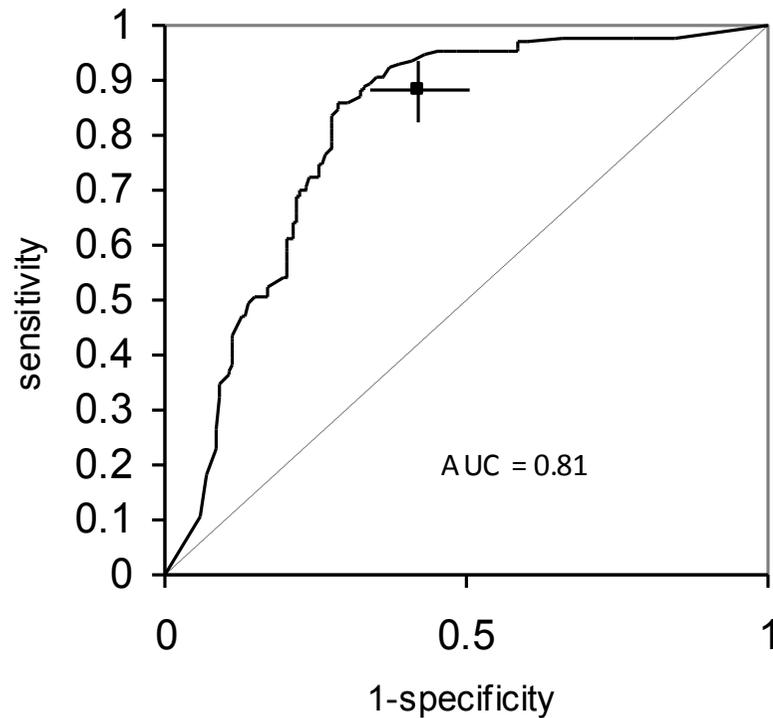
=

**Probabilité de classer
correctement deux parcelles**

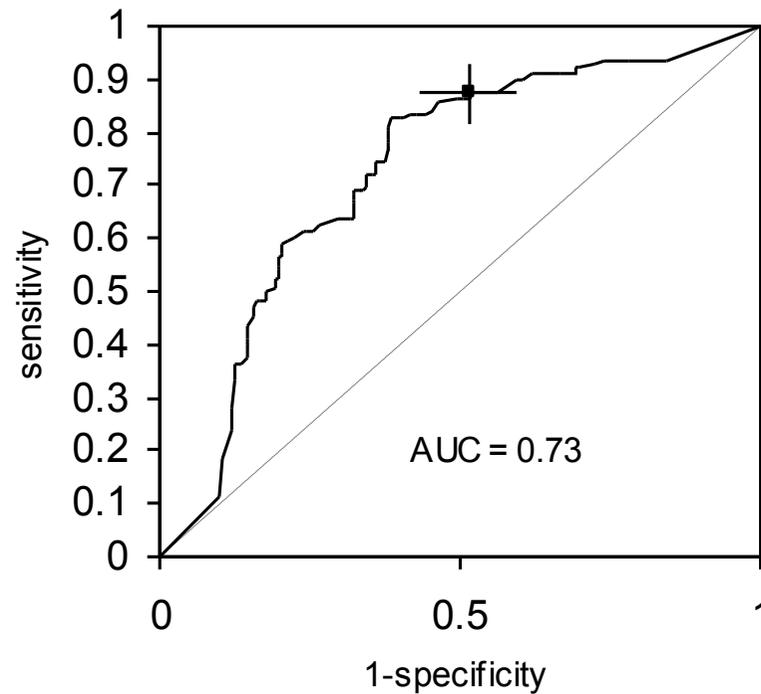
1 - Spécificité

Modèles simulant le risque d'infestation du colza par les mauvaises herbes (Primot *et al.*, 2006)

a. Variables sélectionnées par stepwise



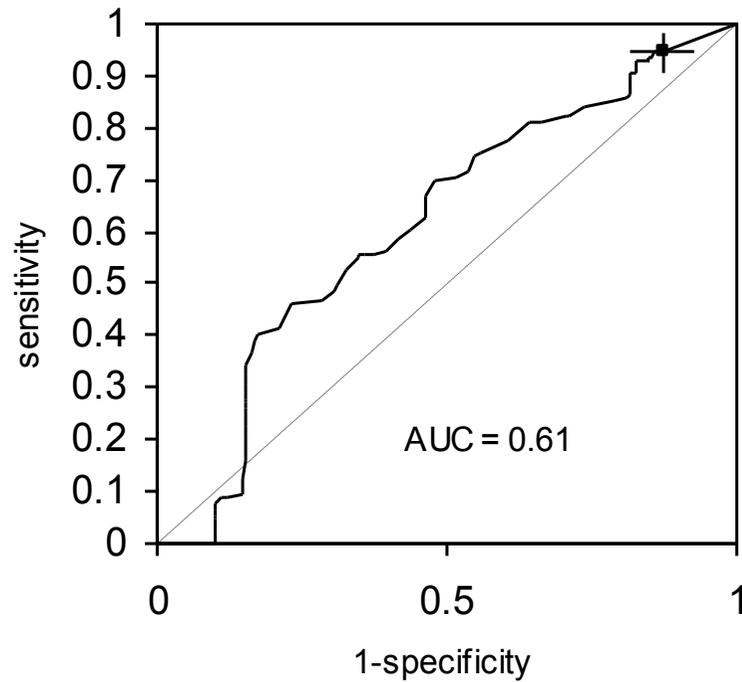
b. Toutes les variables



Modèles simulant le risque d'infestation du colza par les mauvaises herbes (Primot *et al.*, 2006)

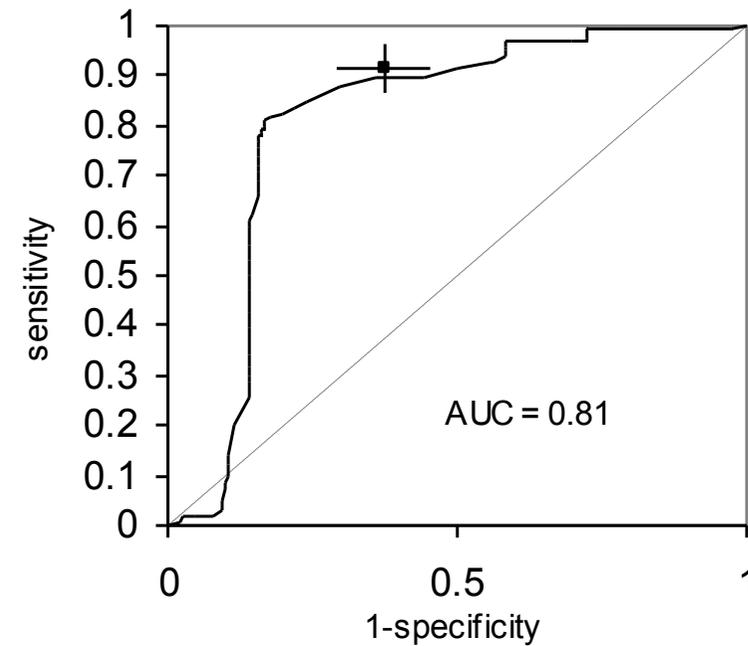
Variables « Techniques »

c.



Variables « Plantes »

d.



Sclerotinia sclerotiorum, Lib., de Bary, in oilseed rape crops

- High variability of disease incidence across sites and years.
- High yield losses if disease incidence at harvest > 10%.
- Efficient chemical treatments exist, but are **not always** required.



Scoring systems

Table 1. Risk factors with points, and number of plots in each category.

| Risk factor | Level | Points | Number of plots |
|--|-------------------|--------|-----------------|
| Number of oil-seed crops during the last ten years | >5 | 30 | 0 |
| | 3-5 | 20 | 12 |
| | 2-3 | 10 | 48 |
| | 1 | 0 | 25 |
| Other host crops during the last five years | Yes | 15 | 47 |
| | No | 0 | 38 |
| Level of infection in the last crop | High | 15 | 19 |
| | Moderate | 5 | 35 |
| | Low | 0 | 31 |
| Type of field | Wet | 10 | 26 |
| | Dry | 0 | 59 |
| Plant density | High | 10 | 14 |
| | Normal | 5 | 57 |
| | Low | 0 | 14 |
| Rain in the last month before flowering | More than normal | 10 | 5 |
| | Normal (50-60 mm) | 5 | 8 |
| | Less than normal | 0 | 72 |
| Number of days with rain (1 to 10 mm) during the last two weeks before flowering | 10-14 | 30 | 0 |
| | 9-5 | 20 | 5 |
| | <5 | 10 | 55 |
| | 0 | 0 | 25 |
| Average temperature during the last 5 days before flowering | >15 °C | 15 | 5 |
| | 10-15 | 10 | 41 |
| | <10 | 0 | 39 |
| Weather forecast | Rain | 30 | 2 |
| | Variable | 15 | 32 |
| | Dry | 0 | 51 |
| Soil | Very wet | 15 | 0 |
| | Wet | 10 | 12 |
| | Dry | 0 | 73 |

A test for the control of *Sclerotinia sclerotiorum*



n collected flowers.



Incubation in Petri dishes



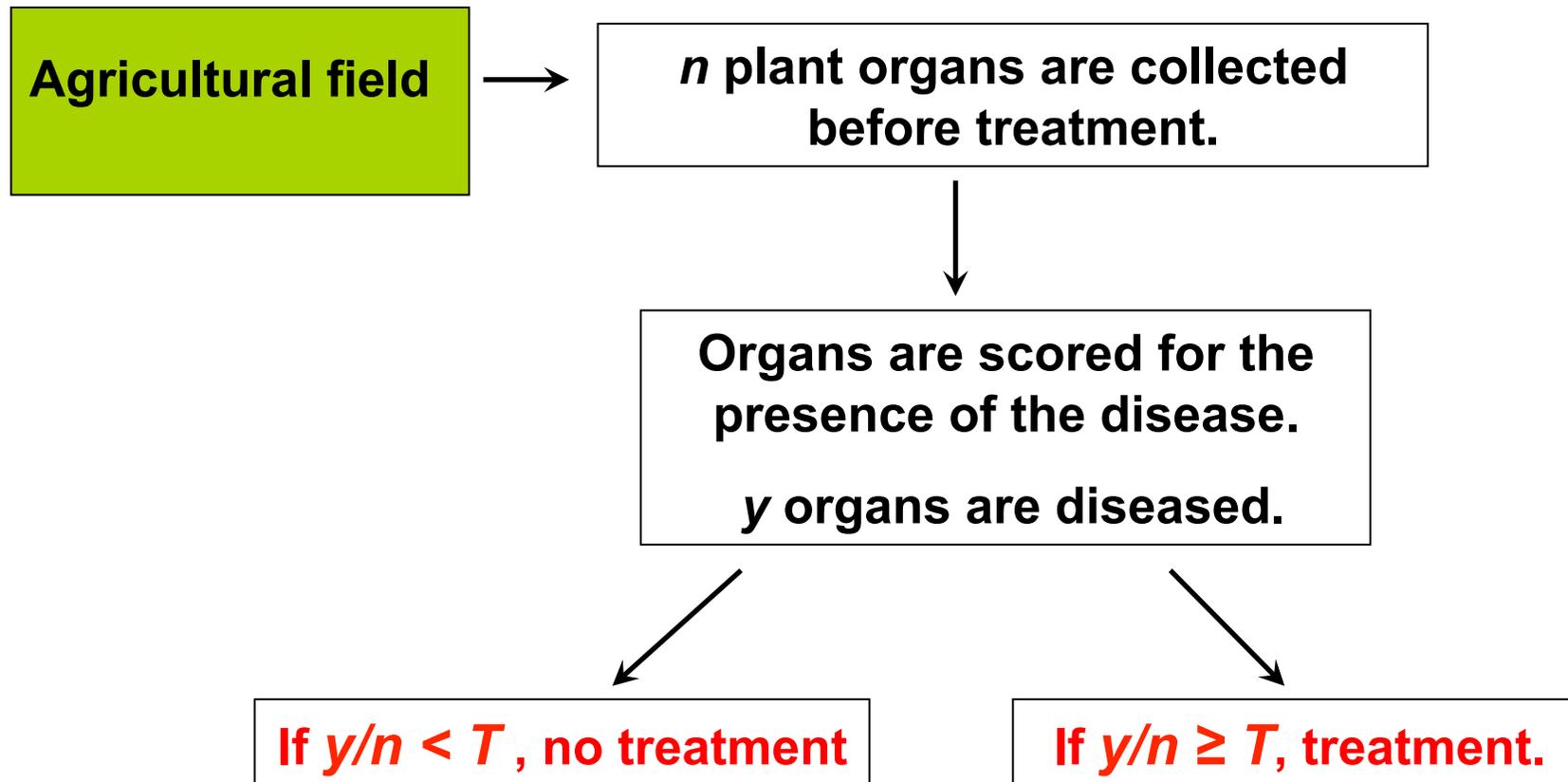
y diseased flowers.



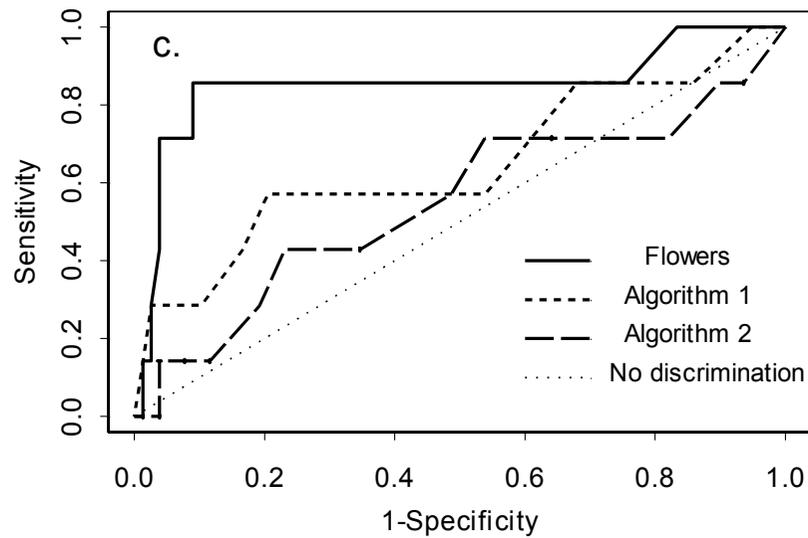
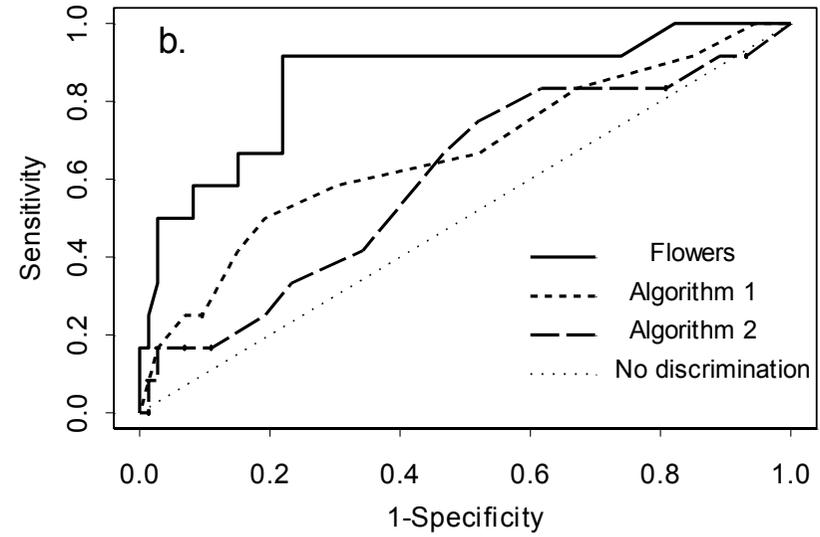
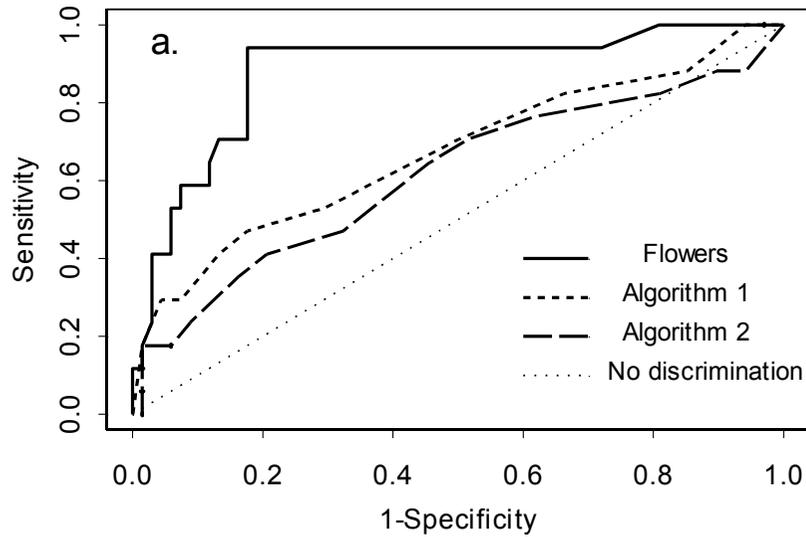
If $y/n \geq T$ treatment

else no treatment.

Diagnostic test based on a measured proportion of diseased plant organs



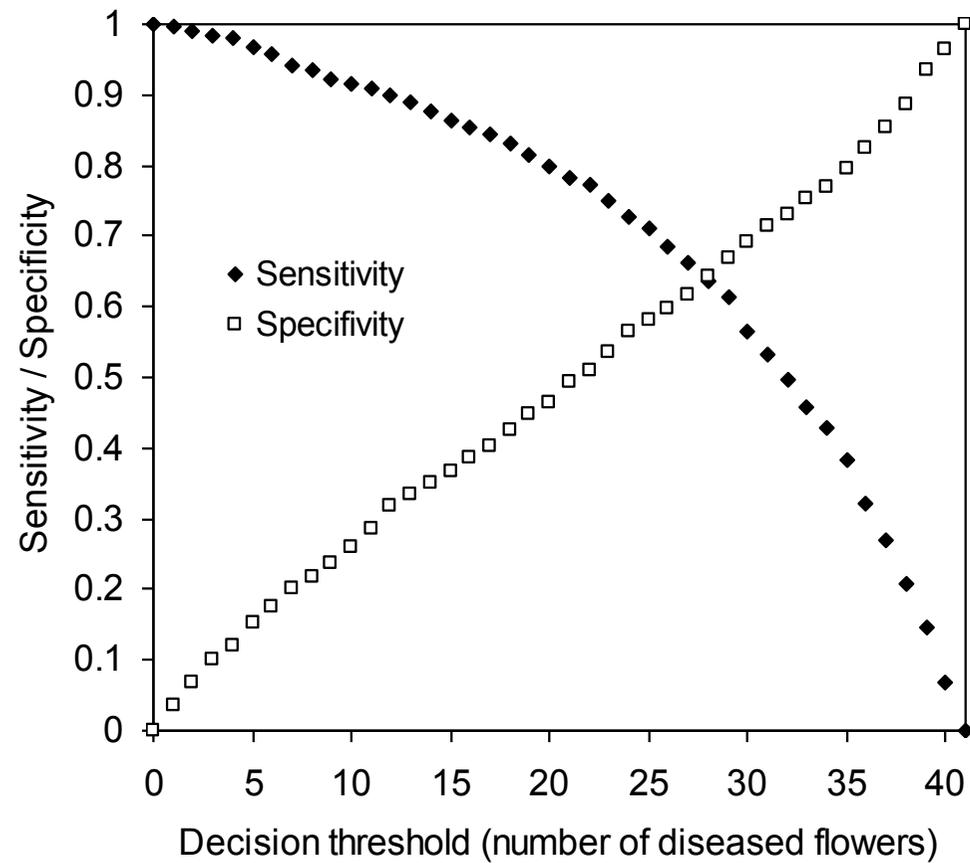
Indicateurs de risque de sclérotinia (Makowski *et al.*, 2005)



Indicateurs de risque de sclérotinia

Combien de fleurs ?

(Makowski *et al.*, 2007)



Cas des règles de décision binaire sans seuil

Exemples:

- PCR qualitative pour raisonner le traitement d'une maladie,
- Effet précédent (traitement si le précédent est un hôte...).

Principe:

- On calcule $A_0 = \text{Sensibilité} + \text{Spécificité}$.
- On estime $P(A \geq A_0)$ sous l'hypothèse nulle (règle non informative).

Parker et David, 1999.

Étape 5. Choix de la meilleure règle

- Avec une analyse ROC, on choisit un compromis entre une forte sensibilité et une forte spécificité.
- Avec la marge moyenne ou le MSEF, on choisit la règle conduisant à un niveau satisfaisant du critère.
- Dans tous les cas, on doit tenir compte du coût d' utilisation.

Bonne solution: fournir un maximum d' informations à l' utilisateur et le laisser choisir.

Code R (library ROCR)

```
TAB<-read.table("f:\\David\\Projets\\Fichier.txt",header=T,sep="\t")
```

```
#Ref=binary variables indicating situations with High vs. Low risk of yield loss  
#ROC analysis
```

```
Ind.list<-c(1,2,3,4,5)
```

```
Auc.vec<-Ind.list
```

```
for (i in 1:length(Ind.list)) {
```

```
    Ind<-TAB[,i]
```

```
    pred<-prediction(Ind,Ref)
```

```
    perf<-performance(pred,"auc")
```

```
    auc<-perf@"y.values"
```

```
    Auc.vec[i]<-as.numeric(auc)
```

```
}
```

```
print(Auc.vec)
```

Références

- Barbottin A, Makowski D, Le Bail M, Jeuffroy M-H, Bouchard Ch, Barrier C. 2008. Comparison of models and indicators for categorizing soft wheat fields according to their grain protein contents. *European Journal of Agronomy* 29, 159-183.
- Makowski D., Denis J-B., Ruck L., Penaud A. 2008. A Bayesian approach to assess the accuracy of a diagnostic test based on plant disease measurement. *Crop Protection* 27:1187-1193.
- Makowski, D., M. Taverne, J. Bolomier, M. Ducarne. 2005. Comparison of risk indicators for sclerotinia control in oilseed rape. *Crop Protection* 24:527-531
- Makowski D., Tichit M., Guichard L., van Keulen H., Beaudoin N. 2009. Measuring the accuracy of agro-environmental indicators. *Journal of Environmental Management* 90, S139-S146.
- Pepe MS. 1998. Three approaches to regression analysis of ROC curves for continuous test results. *Biometrics* 54, 124-135.
- Pepe MS. 2003. The statistical evaluation of medical tests for classification and prediction. Oxford Statistical Science series-28
- Primot, S., M. Valantin-Morison, D. Makowski. 2006. Predicting the risk of weed infestation in winter oilseed rape crops. *Weed Research* 46:22-33
- Swets, J.A., 1988. Measuring the accuracy of diagnostic systems. *Science* 240, 1285-1293.