

Analyse de sensibilité sur sorties multivariées classifiées

Sébastien Roux (INRA - MISTEA)
Patrice Loisel (INRA- MISTEA)
Samuel Buis (INRA- EMMAH)



Projet INRA-EA 'ASM' (2015-2017)
[R.Narci, C.Bidot, H.Monod]

Analyse de sensibilité sur sorties multi-variées classifiées

- Introduction

- Problématique
- Modèle jouet pour l'exposé
- Etat de l'art

- Clustering

- Crisp vs fuzzy clustering
- algo dans cette étude : fuzzy c-means
- Application sur Toycurves

- Proposition Algo : AS on membership functions

- principe
- ASMF : indices par classe
- ASMF : indices par paramètre
- ASMF : indices agrégés (GSI)

- Conclusion et perspectives

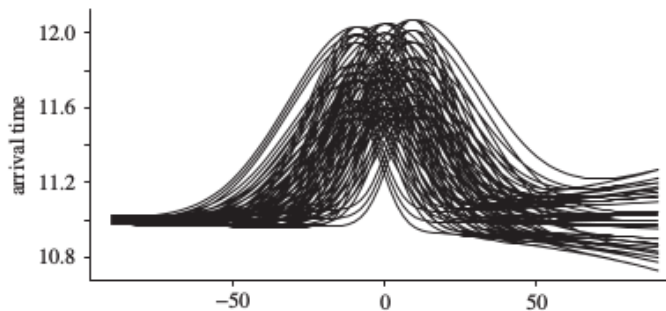
- intérêt /MSA
- intérêt /DGSA : accès aux interactions
- perspectives : intraclasse, cas réel, incertitude

Introduction

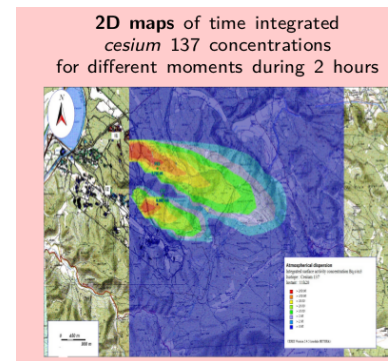
- Problématique
- Modèle jouet pour l'exposé
- Etat de l'art

Problématique

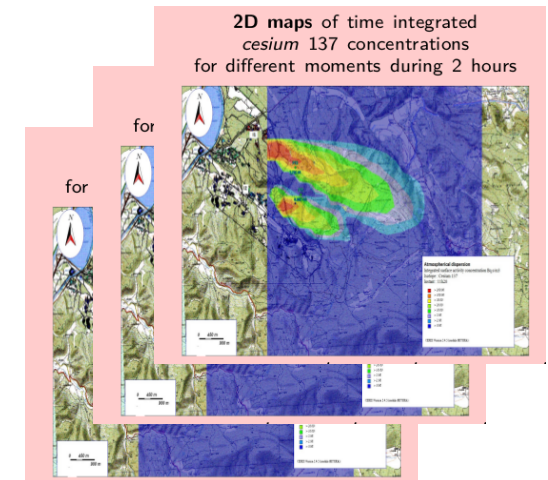
Analyse du comportement de modèles à sorties dynamiques / spatiales par rapport aux variations de leurs paramètres d'entrées



Sorties temporelles
(Campbell et al 2006)



Sorties spatialisées
(Marrel et al 2017)



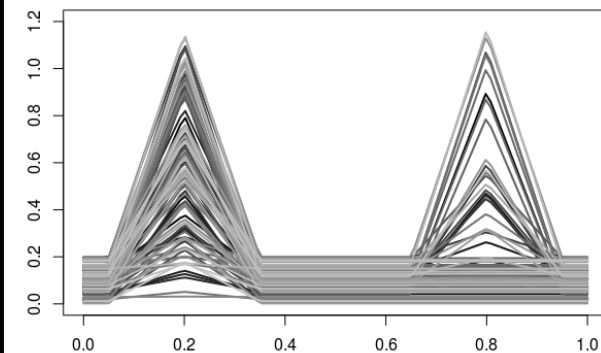
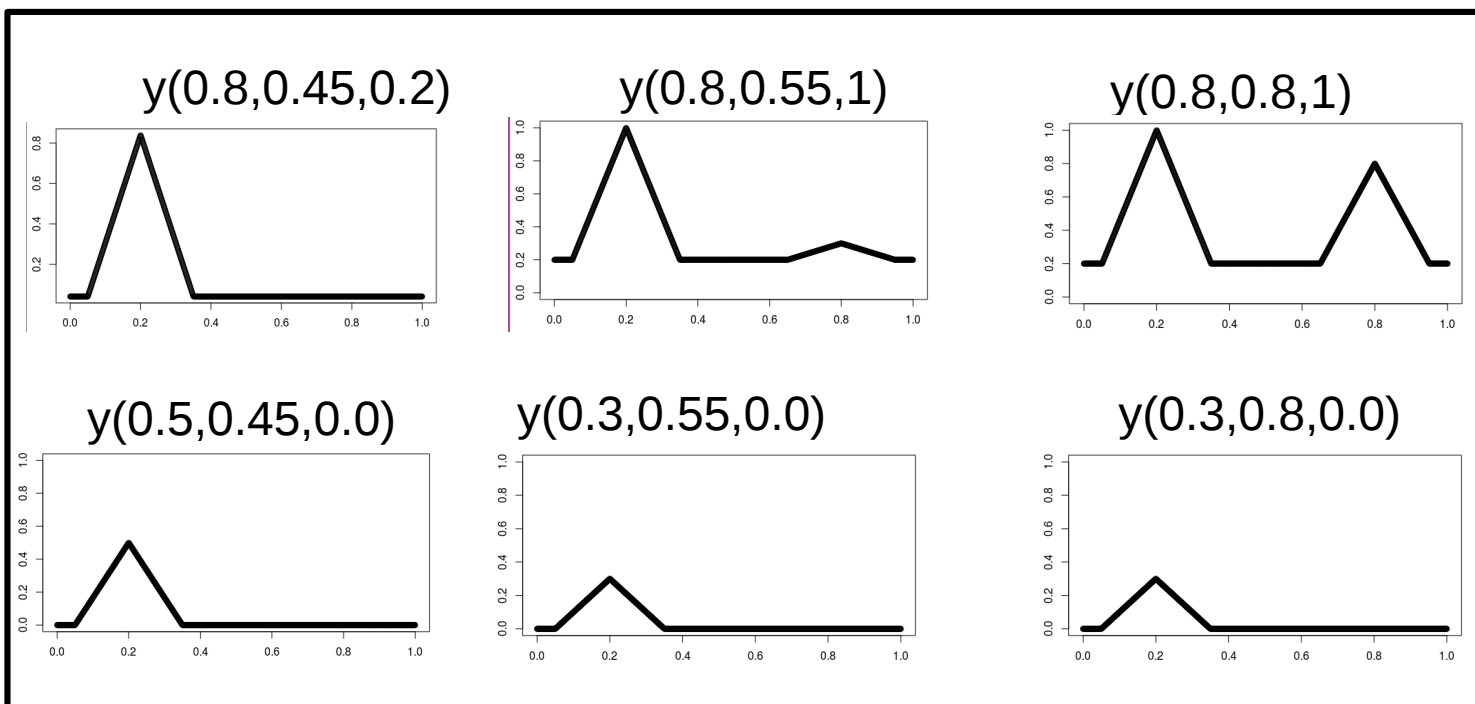
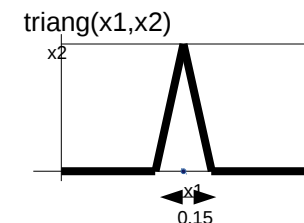
K. Campbell, et al. (2006). Sensitivity analysis when model outputs are functions. Reliability Engineering & System Safety, 91(10-11), 1468-1472

A. Marrel et al. (2017) Sensitivity Analysis of Spatial and/or Temporal Phenomena, Handbook of uncertainty Quantification, 2017

Toy model pour l'étude : 'Toycurves'

MODELE JOUET : $y(t)=\text{ToyCurve}(p_1,p_2,p_3,t)$

$$y(p_1, p_2, p_3, t) = \frac{p_3}{5} + \text{triang}(0.2, p_1)(t) + 1_{((p_2 > 0.5) \cdot (p_1 > 0.5))}(t) \cdot \text{triang}(0.8, 2 \cdot (p_2 - 0.5))(t)$$



- 1 ou 2 triangles pilotés par p_1 et p_2 :
 - p_1 pilote l'amplitude du 1^{er} (toujours present)
 - p_2 pilote l'amplitude du second, qui n'est present que si $(p_1 > 0.5)$ et $(p_2 > 0.5)$
- p_3 : petit déplacement vertical de la courbe (max 0.2)

État de l'art

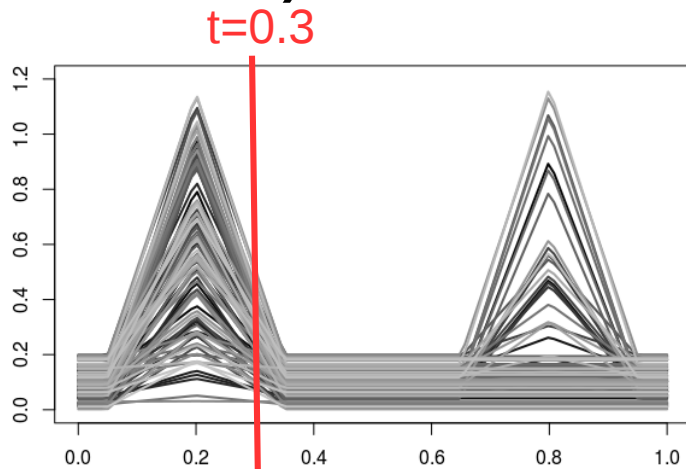
Multivariate Sensitivity Analysis (MSA)

Regional Sensitivity Analysis (RSA)

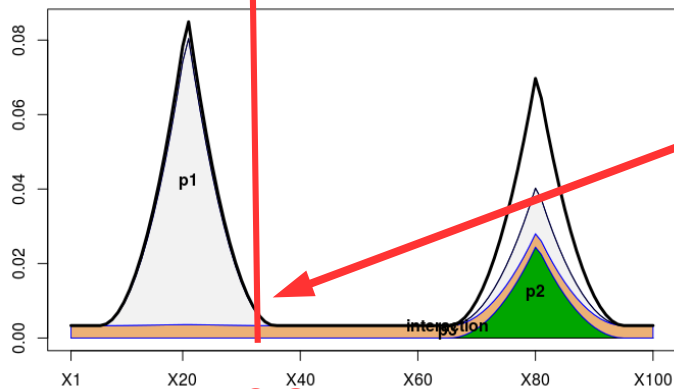
- **MSA** : Analyses répétées
- **MSA** : Analyses par composantes via projections dans une base
- **MSA** : Analyses agrégées : GSIs
- **RSA** : Extension de l'analyse de sensibilité régionale au cas multivarié

MSA : Analyses répétées

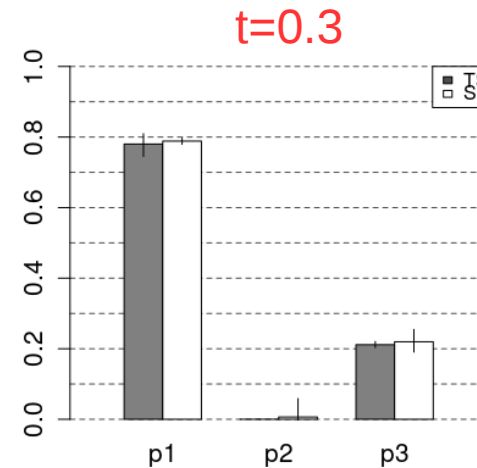
- Répétition d'analyses de sensibilité globales scalaires (e.g. basées sur la variance) de la sortie à chaque instant



Courbes du plan d'expérience



Indices non normalisés par $\text{Var}(Y(t))$

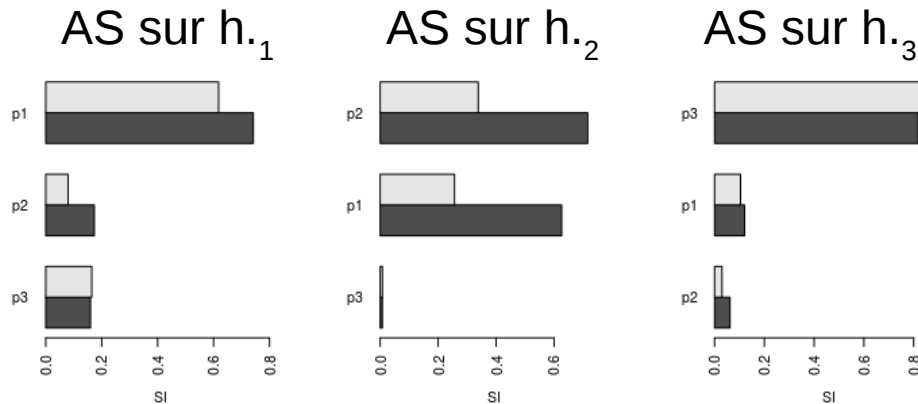
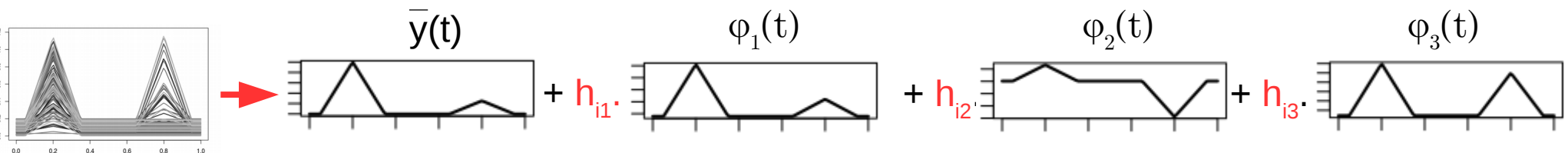


SI à $t=0.3$

=> quel impact des paramètres sur la variation de la réponse à chaque instant ?

MSA : Analyses par composantes via projections dans une base

- $y_i(t) - \bar{y}(t) \approx \sum_{k=1}^K h_{ik} \varphi_k(t)$ for $1 \leq i \leq N$
- GSA sur les coordonnées (h_{ik}) (projections sur $\varphi_k(t)$)
- Bases fixes (eg polynômes) ou bases adaptatives (e.g. ACP)

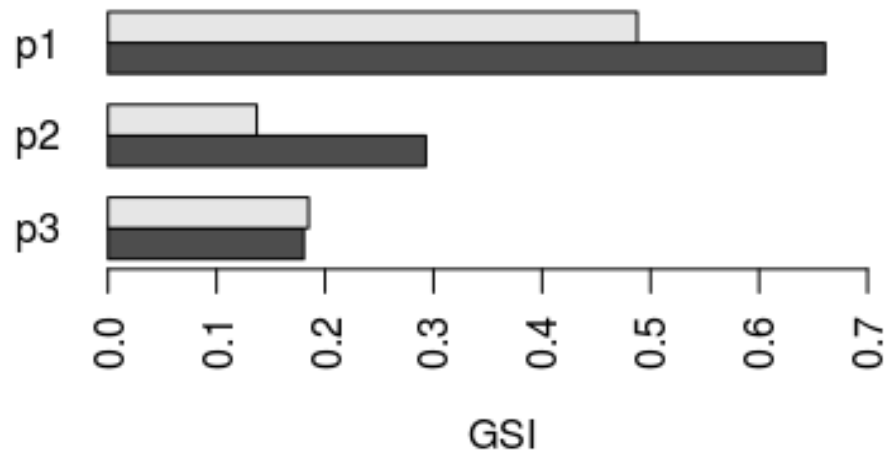
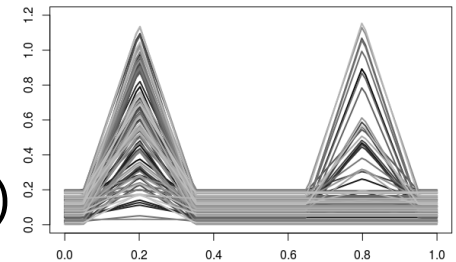


=> quel impact des paramètres sur la variation des composantes de la réponse ?

MSA : Analyses agrégées

- Explication globale de la variabilité de la sortie par extraction d'indices issus de la décomposition de la covariance

GSI : basé sur la trace de la matrice de covariance
 Décomposition de la somme des variances des y_i
 (lamboni11, Gamboa15, Xiao2018, Marell2017, lamboni18)

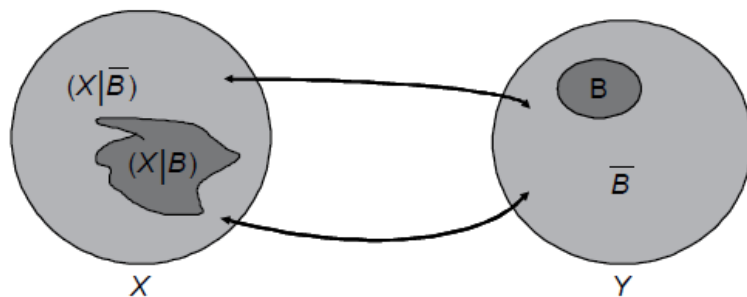


=> quel impact global des paramètres sur la variation de la réponse ?

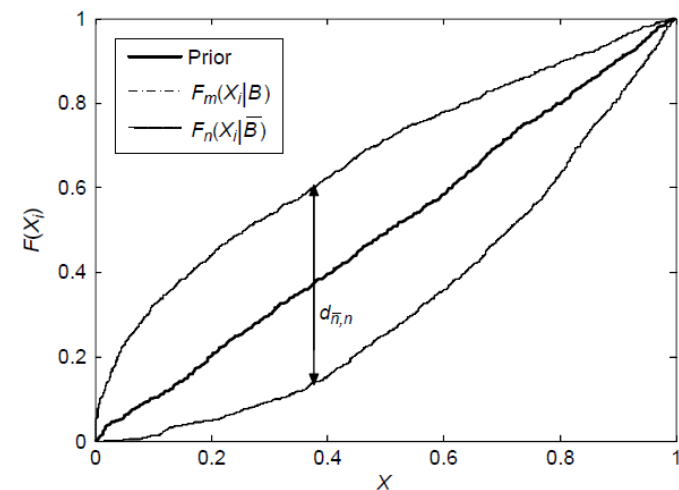
RSA : Extension de l'analyse de sensibilité régionale au cas multivarié

RSA original : Young et al. (1978) et Spear et Hornberger (1980)

1. Calculer la sortie du modèle Y pour un échantillon des X
2. Classer les valeurs de Y en 2 catégories : behavioural (B) et non-behavioural (\bar{B})
3. Comparer les distributions de $(X|B)$ et $(X|\bar{B})$ avec un test de Kolmogorov-Smirnov



Source : Saltelli et al (2008)

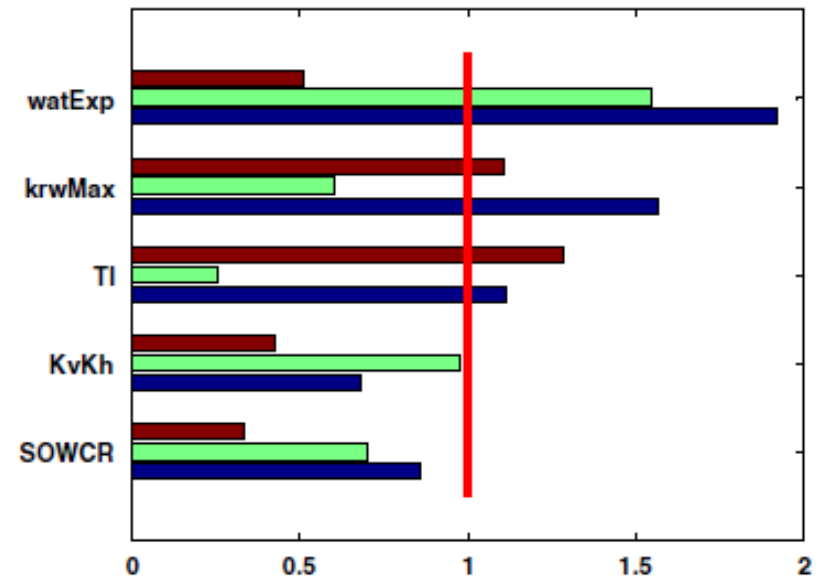
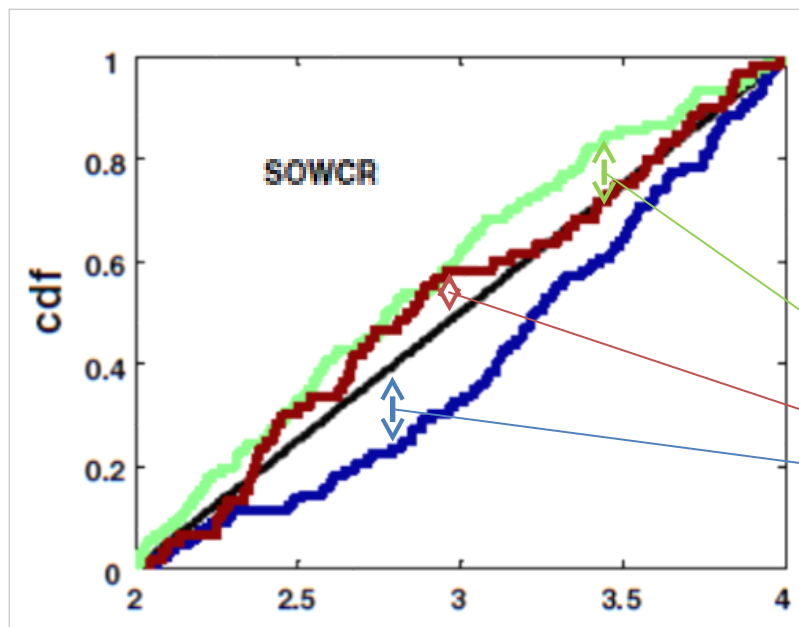


=> Quels paramètres mènent le modèle vers un comportement donné (e.g. vers des réponses qui simulent correctement le système réel)

RSA : Extension de l'analyse de sensibilité régionale au cas multivarié

DGSA (extension RSA) : Fenwick et al (2014)

- Prise en compte de sorties multivariées
- Quantification des interactions d'ordre 2
- Test de significativité par bootstrap



CDF distances

=> Quels paramètres mènent le modèle vers un comportement donné (caractérisé par une classe de courbes)

Avantages / Limites

	Avantages	Limites
Analyses répétées	Facile à mettre en œuvre et interpréter	Dynamique pas prise en compte dans sa globalité
Analyse par composantes	Information concernant l'impact des paramètres sur la forme de la dynamique	<ul style="list-style-type: none"> • Interprétation complexe • Contraintes sur la définition des composantes
Analyses agrégées	Hiérarchie globale des paramètres	Pas d'information sur l'impact des paramètres sur la forme de la réponse
DGSA	<ul style="list-style-type: none"> • Facile à interpréter • Identification des paramètres qui mènent à un comportement donné • Souplesse de caractérisation des comportements 	Méthode peu puissante (interactions limitées à l'ordre 2)

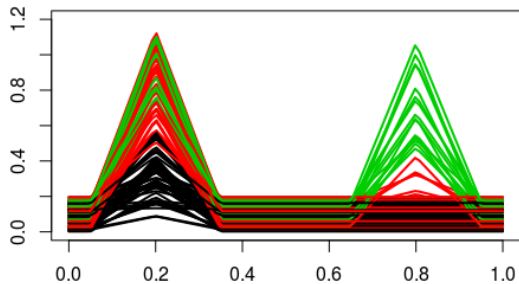
Idée

- **Combiner classification des dynamiques simulées pour caractériser des comportements distincts de la réponse du modèle**
 - sans contraintes d'orthogonalité (!= méthodes de projection)
 - en optimisant la séparation de l'information (!=PCA)
- **et méthodes basées sur la variance**
 - pour identifier de façon robuste les paramètres (ou les interactions de paramètres) qui influent sur un comportement donné

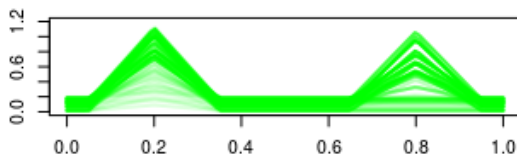
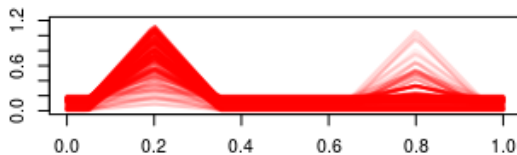
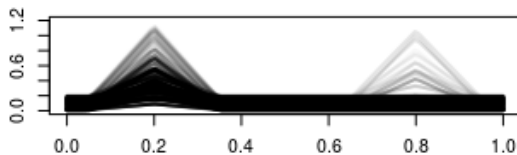
Clustering

- Crisp vs fuzzy clustering
- algo dans cette étude : fuzzy c-means
- Application sur Toycurves

Crisp vs fuzzy clustering



Crisp clustering :
 $y(t) \Rightarrow$ indice de classe $\in [1..K]$



fuzzy clustering : membership (proba)
 $y(t) \Rightarrow (u_1, \dots, u_K)$

Crisp(hard) versus Fuzzy(soft)

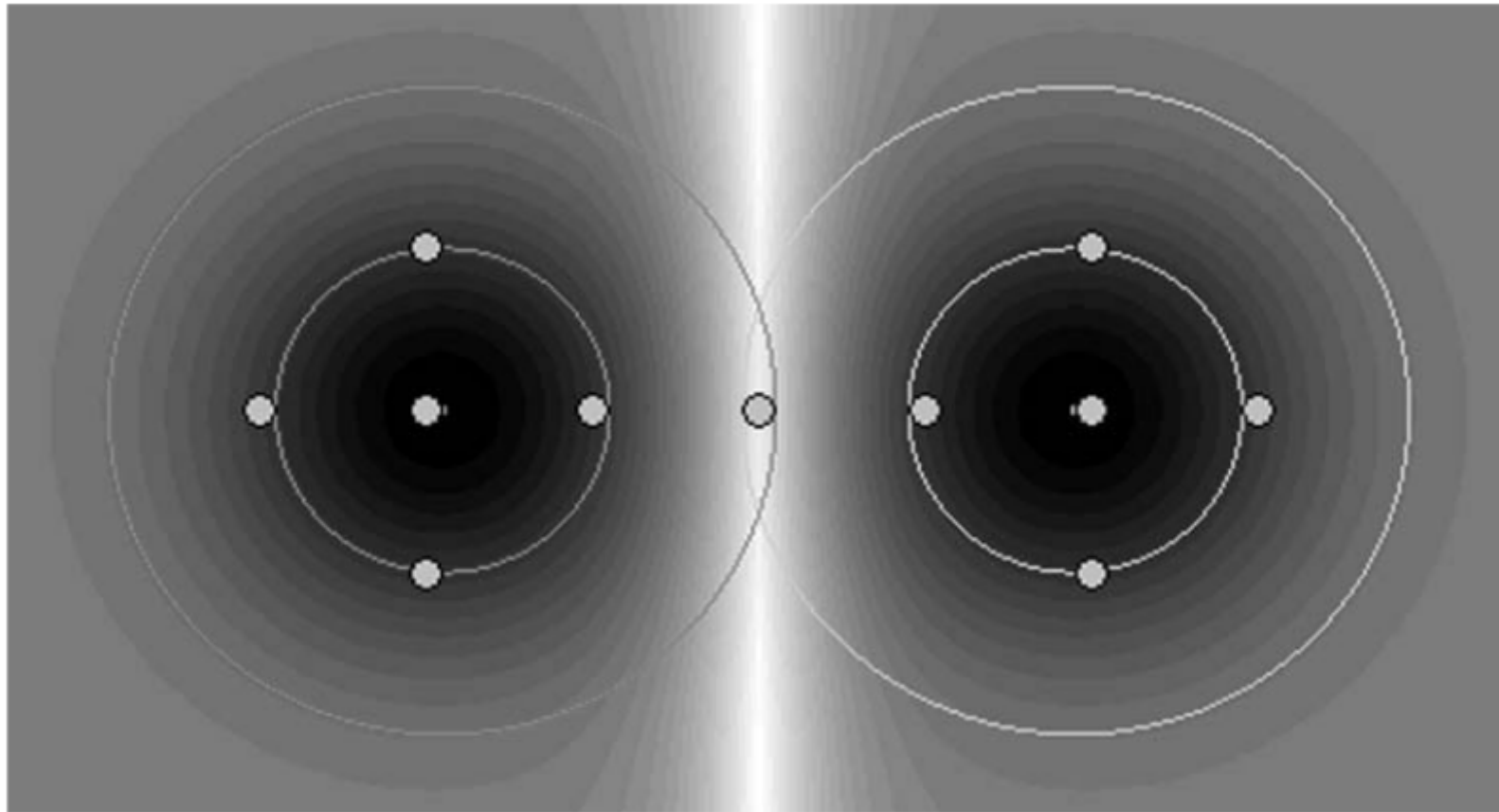


Figure 1.1 A symmetric data-set with two clusters.

Algorithme standard : fuzzy c-means (fcm)

Principe algo de base

$$J_f(\mathbf{X}, \mathbf{U}_f, \mathbf{C}) = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m d_{ij}^2,$$

$$\sum_{j=1}^n u_{ij} > 0, \quad \forall i \in \{1, \dots, c\}, \quad \text{and}$$

$$\sum_{i=1}^c u_{ij} = 1, \quad \forall j \in \{1, \dots, n\}$$

d_{ij} is the distance between x_j
and cluster center c_i ,

Alternate optimization scheme :

The objective function J_f is alternately optimized, i.e., first the membership degrees are optimized for fixed cluster parameters, then the cluster prototypes are optimized for fixed membership degrees:

$$\mathbf{U}_\tau = j_U(\mathbf{C}_{\tau-1}), \quad \tau > 0 \quad \text{and} \quad (1.11)$$

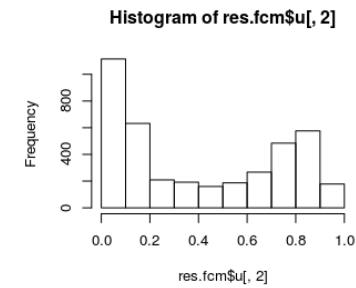
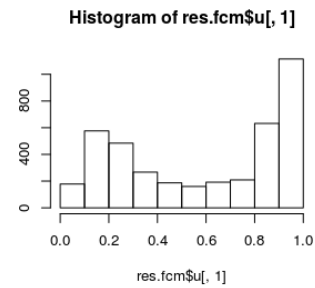
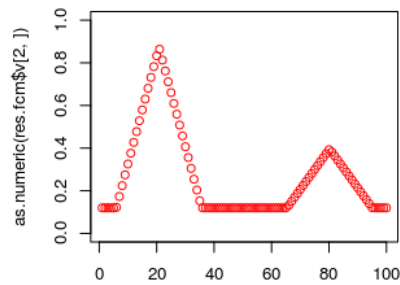
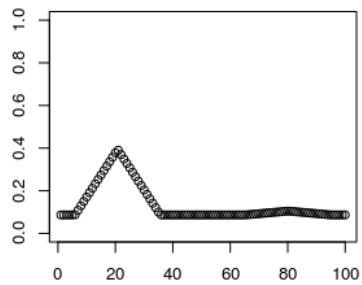
$$\mathbf{C}_\tau = j_C(\mathbf{U}_\tau). \quad (1.12)$$

$$u_{ij} = \frac{1}{\sum_{l=1}^c \left(\frac{d_{ij}^2}{d_{lj}^2} \right)^{\frac{1}{m-1}}} = \frac{d_{ij}^{\frac{2}{m-1}}}{\sum_{l=1}^c d_{lj}^{\frac{2}{m-1}}}. \quad \mathbf{c}_i = \frac{\sum_{j=1}^n u_{ij}^m \mathbf{x}_j}{\sum_{j=1}^n u_{ij}^m}.$$

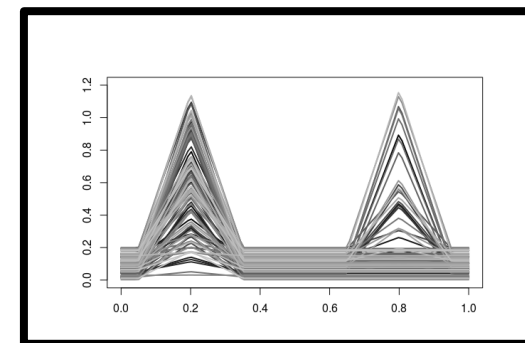
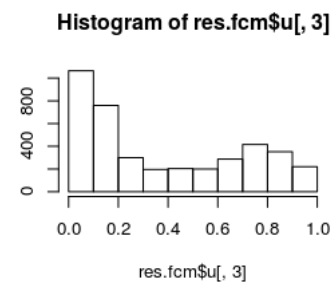
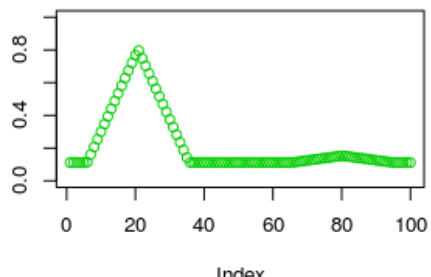
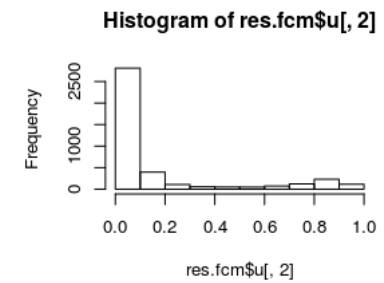
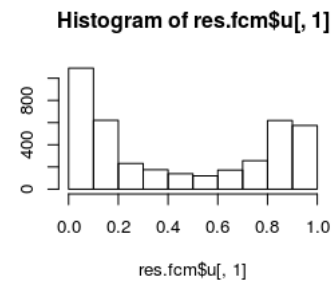
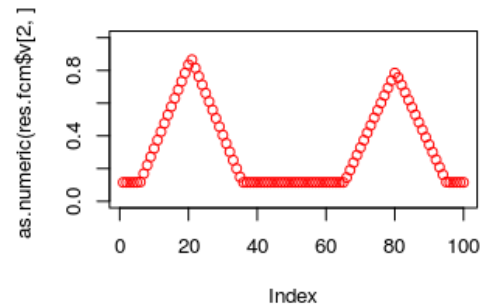
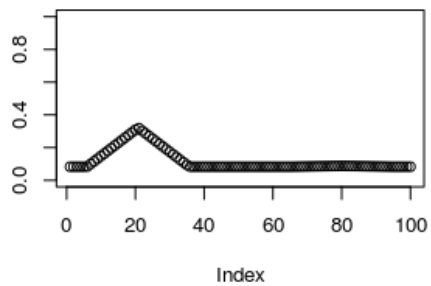
On the theoretical side it has been proven that either the iteration sequence itself or any convergent subsequence of the probabilistic FCM converges in a saddle point or a minimum – but not in a maximum – of the objective function (Bezdek, 1981).

Application sur Toycurves

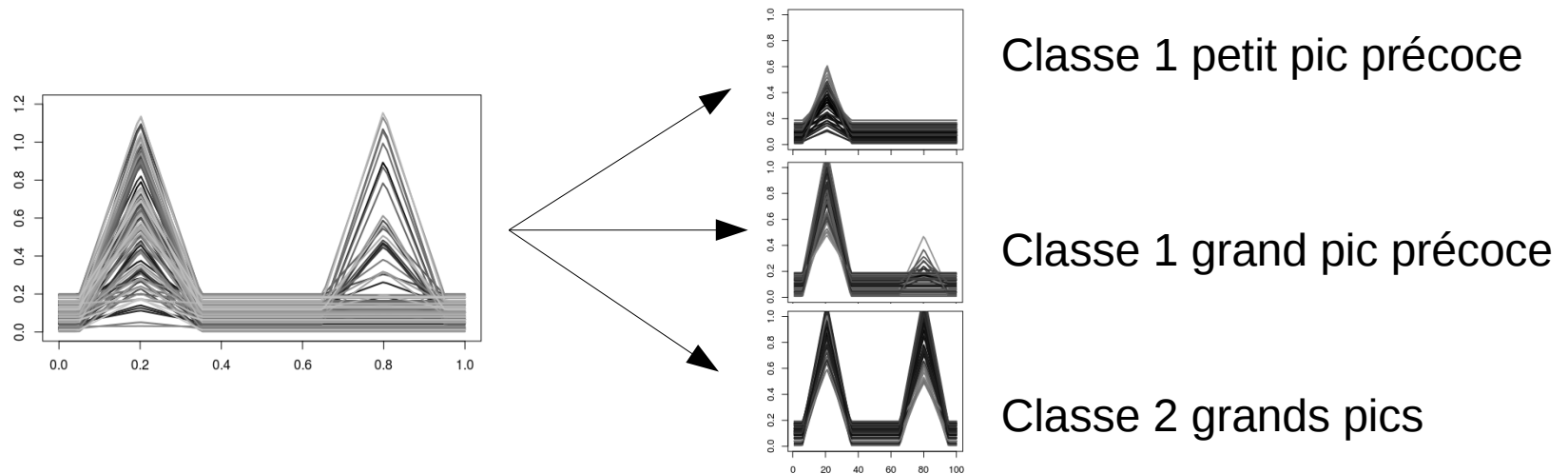
2 classes



3 classes



Questions d'intérêt si les sorties «sont bien classifiées »



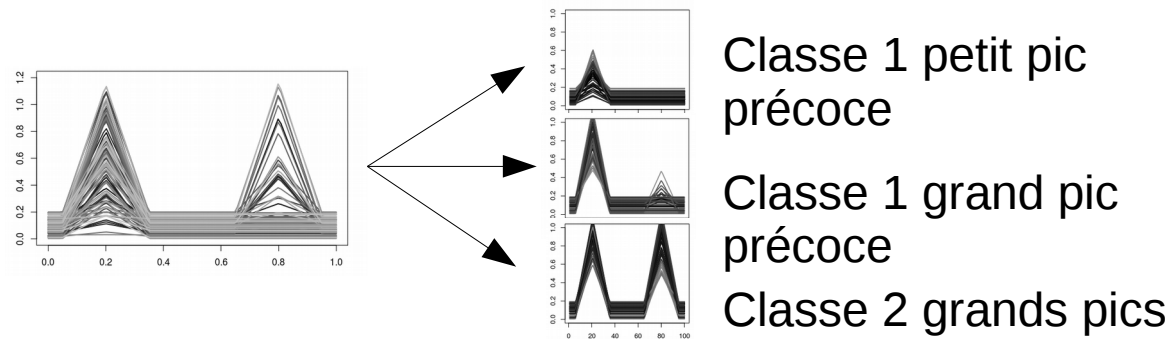
Sorties multivariées => classes de comportement

- Quels paramètres influent sur l'appartenance à une classe cible ?
- Quelle influence d'un paramètre sur les différents comportements ?
- Quelle hiérarchie globale d'importance des paramètres sur le changement de comportement ?

AS on Membership functions

- Principe
- Illustration sur Toycurves avec apprentissage (fcm)

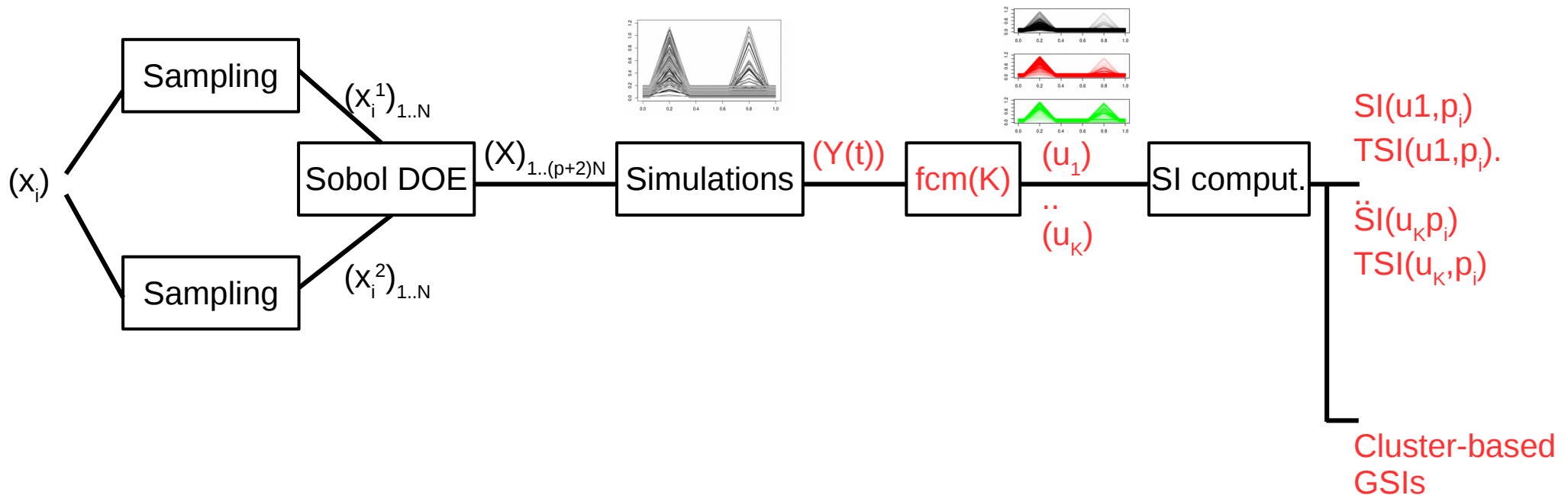
AS on Membership functions



- Quels paramètres influent sur l'appartenance à une classe cible ?
 - => quels paramètres influent sur la variabilité de la MF d'une classe
 - => Variance-based AS sur les MF
- Quelle influence d'un paramètre sur les différents comportements ?
 - => comparaison des indices de sensibilité obtenus sur les MF
- Quelle hiérarchie globale d'importance des paramètres sur le changement de comportement ?
 - => extension de la notion de GSI au vecteur de MF

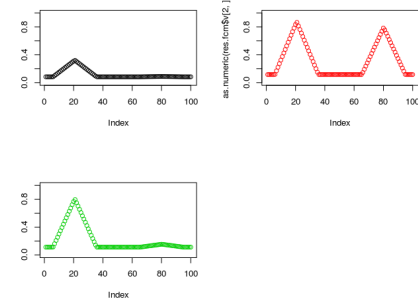
AS on Membership Functions

GSA scheme on clustered outputs with a Sobol approach



AS on Membership functions

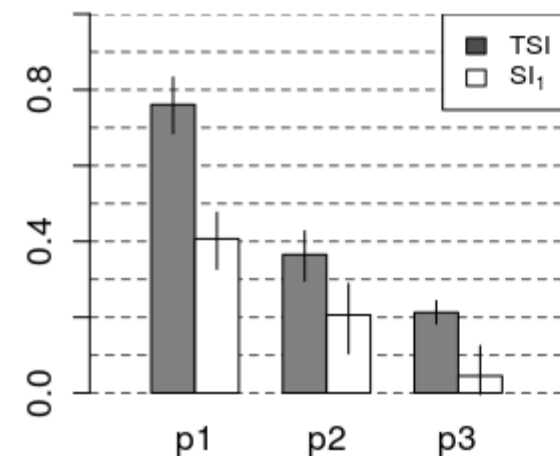
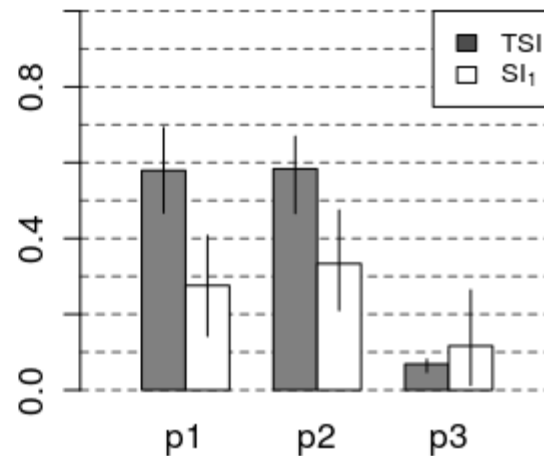
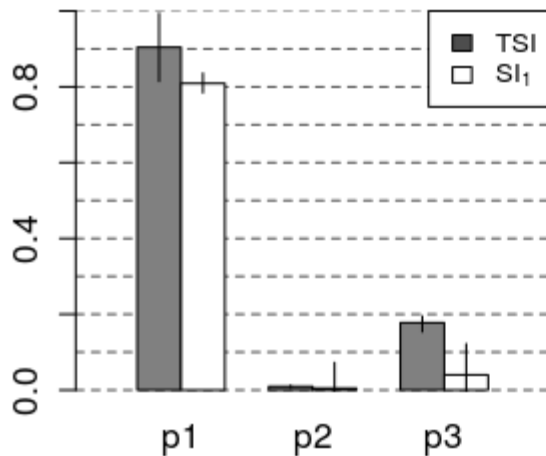
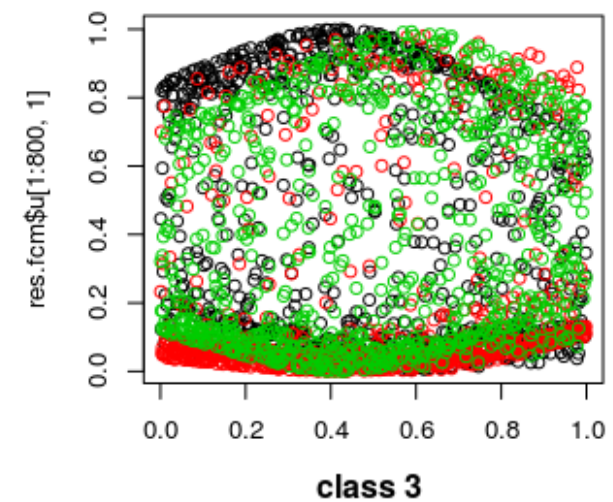
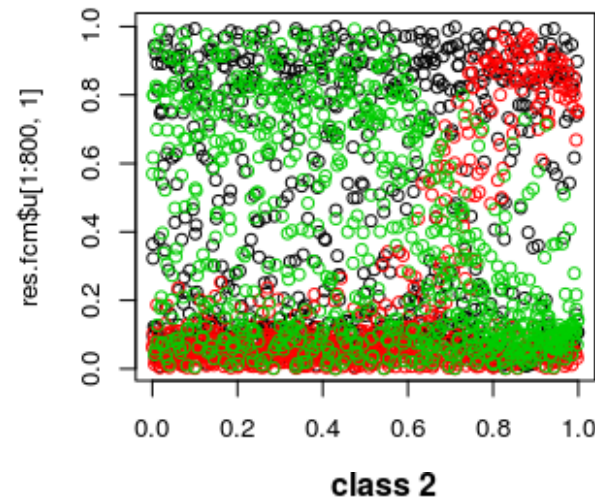
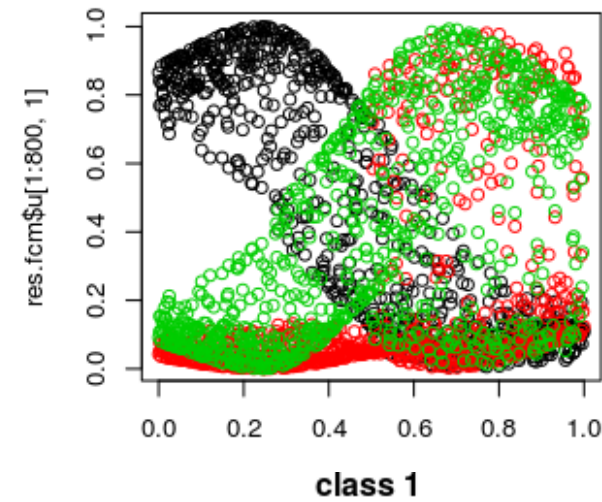
ToyCurve avec clustering (fcm(3)): by class



$u(p1)$

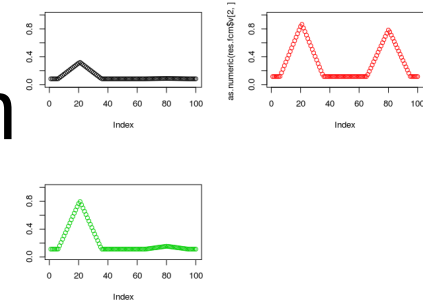
$u(p2)$

$u(p3)$

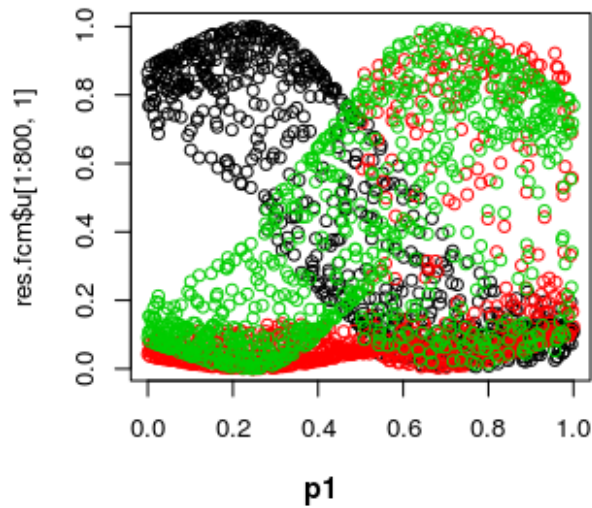


AS on Membership functions

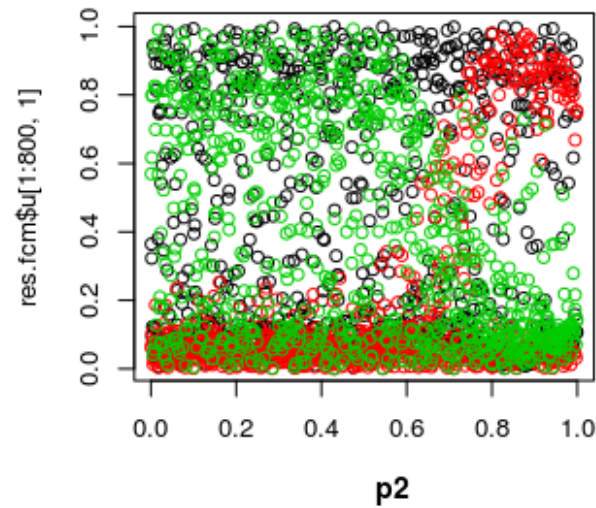
ToyCurve avec clustering (fcm(3)): by param



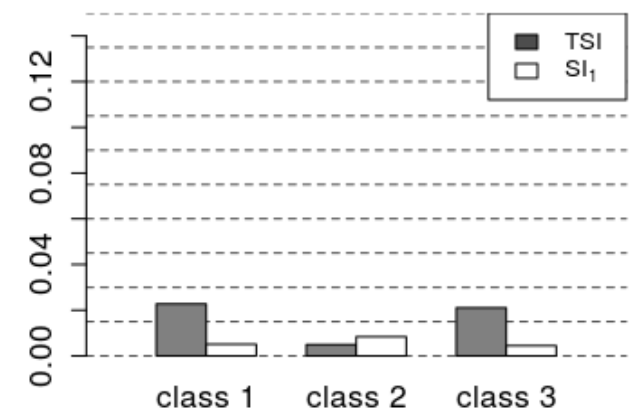
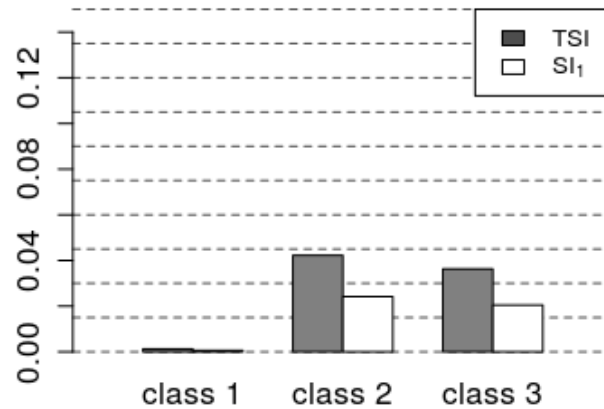
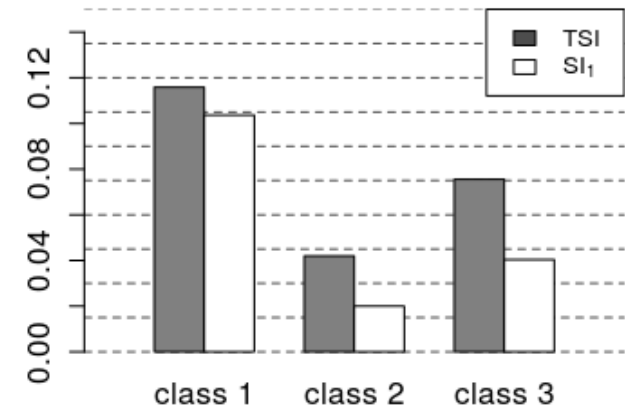
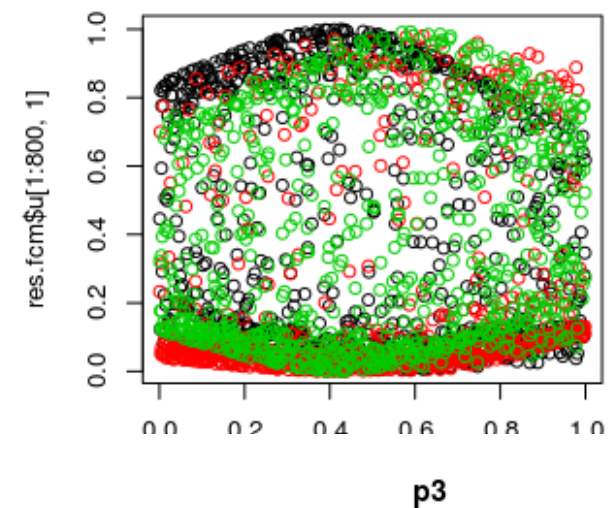
$u(p1)$



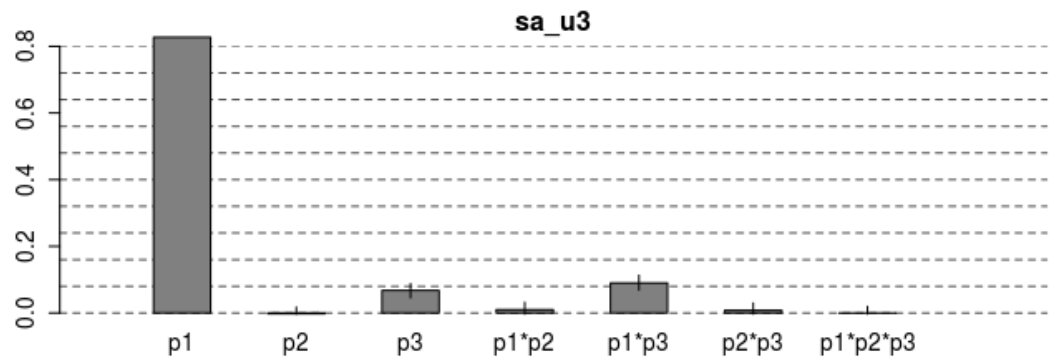
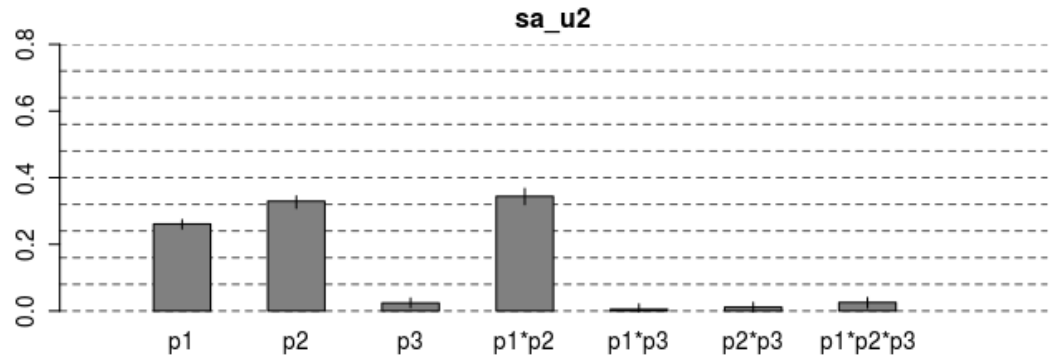
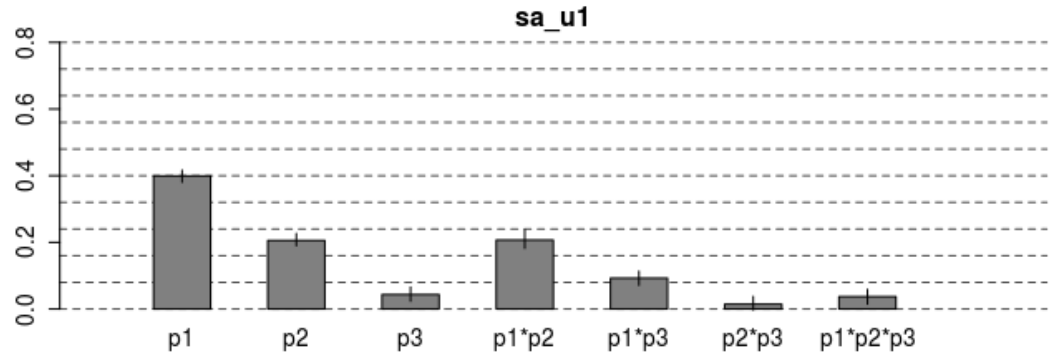
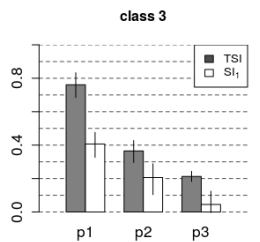
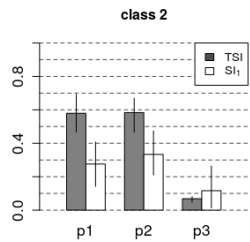
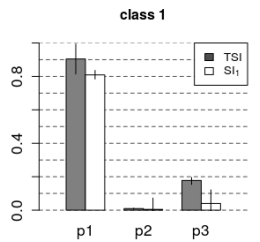
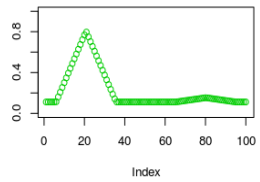
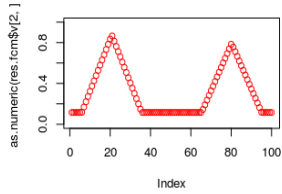
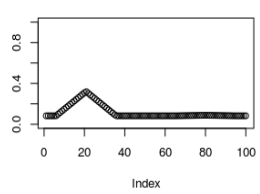
$u(p2)$

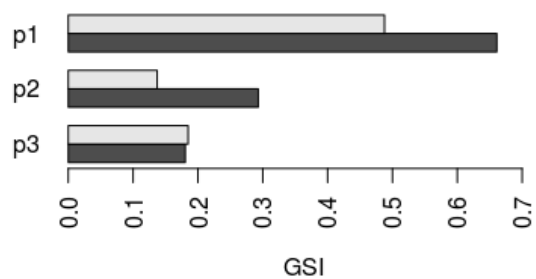
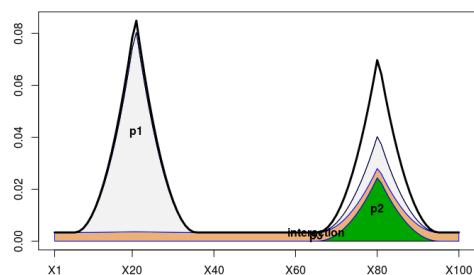
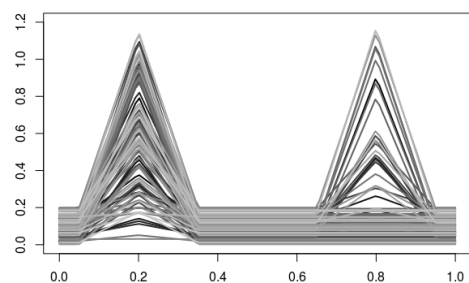
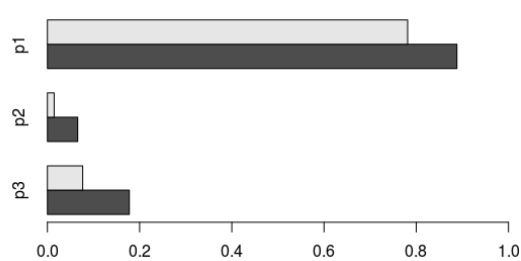
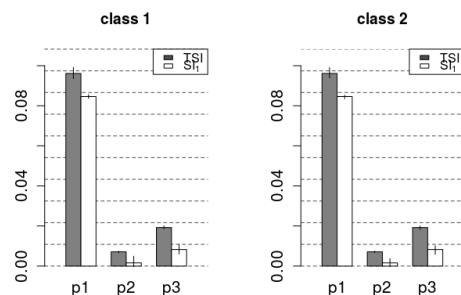
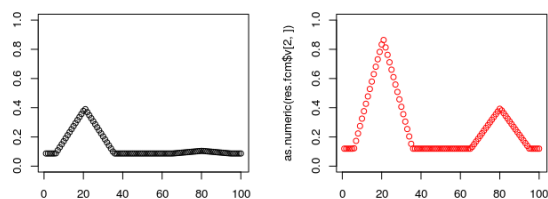
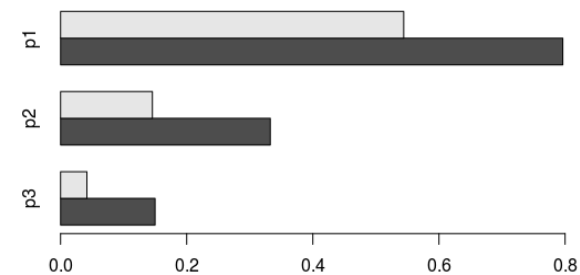
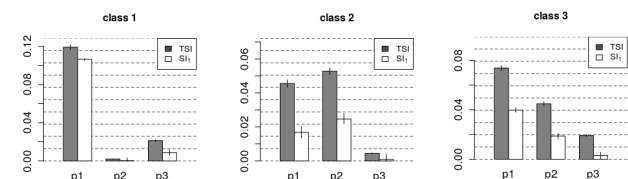
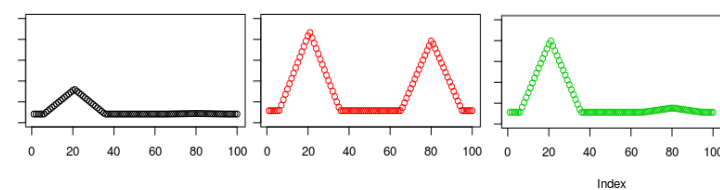


$u(p3)$



AS-MF : access to all interactions



GSI on $[u_1, \dots, u_K]$ *GSI standard sur $y(t)$* *GSI sur fcm(2)**GSI sur fcm(3)*

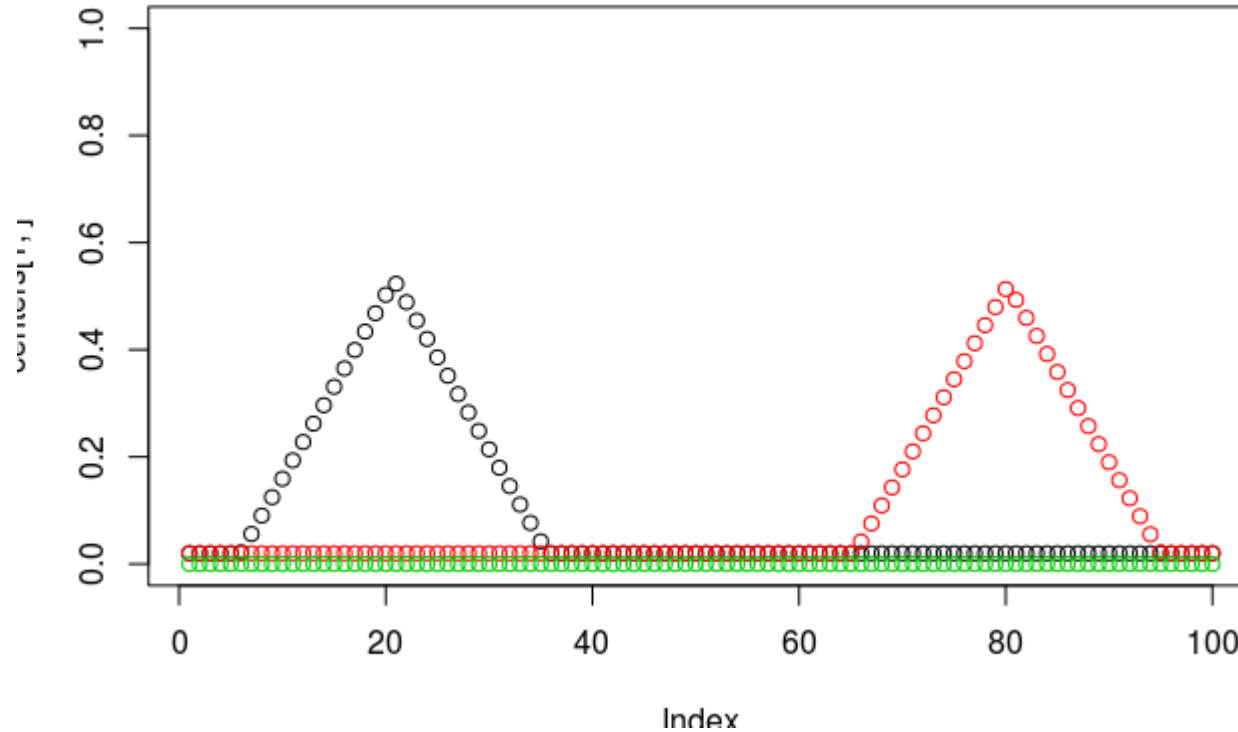
Conclusion et perspectives

- **AS on MF versus MSA :**
 - Extension de MSA sur composantes
 - Extension de GSI
- **AS on MF versus DGSA :**
 - indices interclasses basés sur la variance
 - accès aux interactions de tout ordre (si la méthode d'AS le permet)
- **Autres possibilités offertes par AS on MF :**
 - AS sur classification experte
 - AS sur les directions
- **Perspectives :**
 - AS intra-classes (vs inter-classes)
 - Cas d'étude réaliste

Annexes

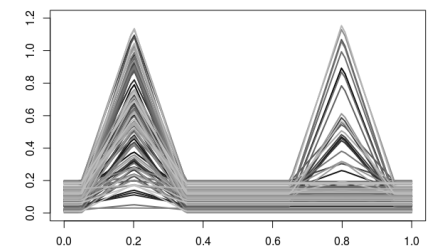
- AS on prior clusters
- AS on MF deltas

ASMF with prior clusters



- 1 early triangle
- 1 late triangle
- A flat signal close to 0

NB : single late triangle **never observed in the data set generated with the model**



ASMF with prior clusters

MF computation

From the fcm algorithm, once centers are specified, membership functions defined by

$$u_{ij} = \frac{1}{\sum_{l=1}^c \left(\frac{d_{ij}^2}{d_{lj}^2} \right)^{\frac{1}{m-1}}} = \frac{d_{ij}^{-\frac{2}{m-1}}}{\sum_{l=1}^c d_{lj}^{-\frac{2}{m-1}}}.$$

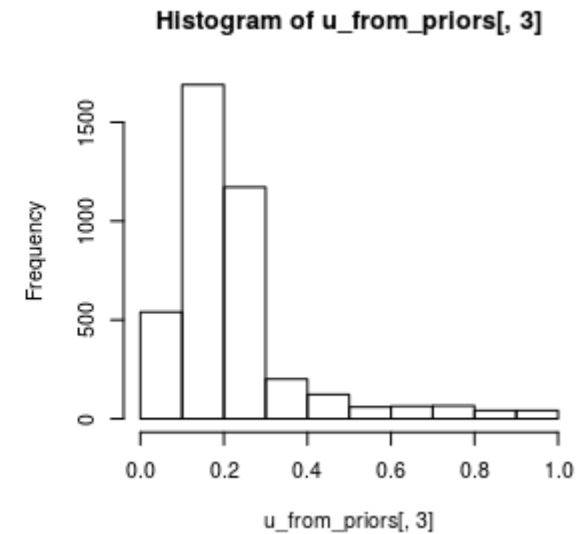
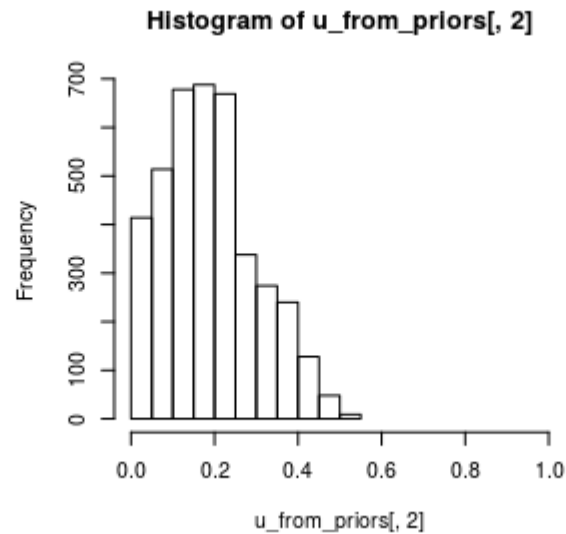
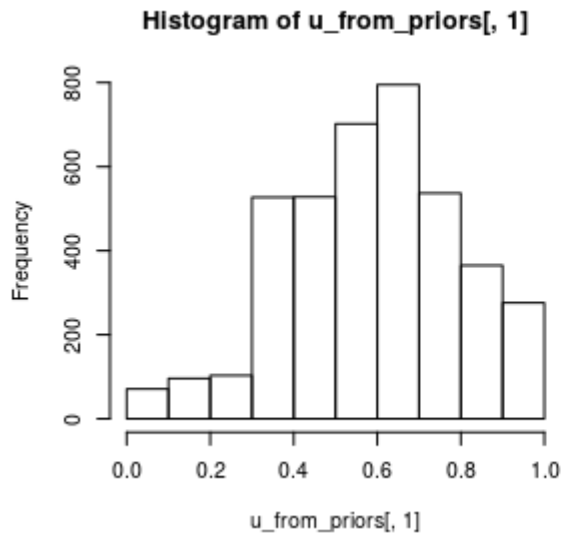
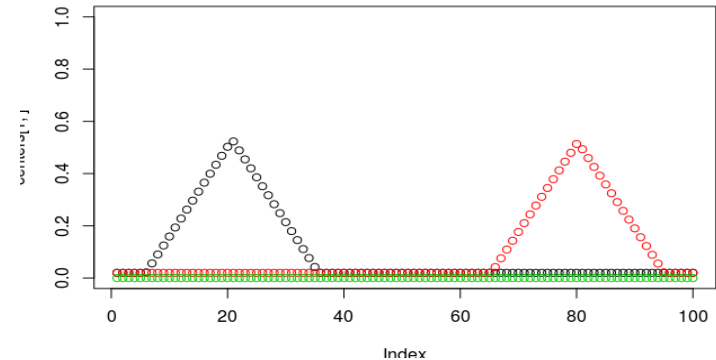
Are the one minimizing

$$J_f(\mathbf{X}, \mathbf{U}_f, \mathbf{C}) = \sum_{i=1}^c \sum_{j=1}^n u_{ij}^m d_{ij}^2,$$

when C is fixed

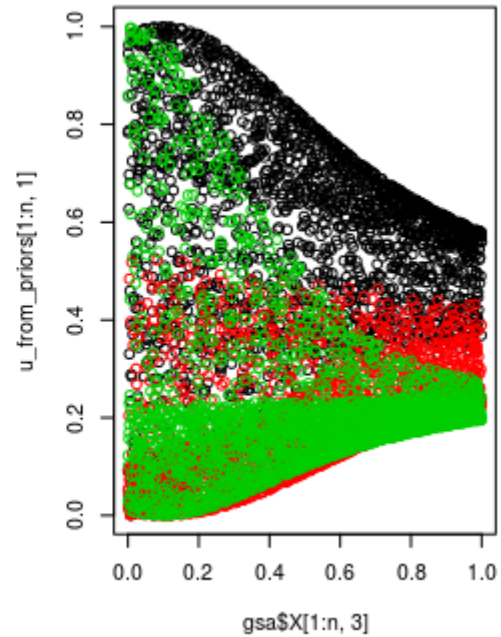
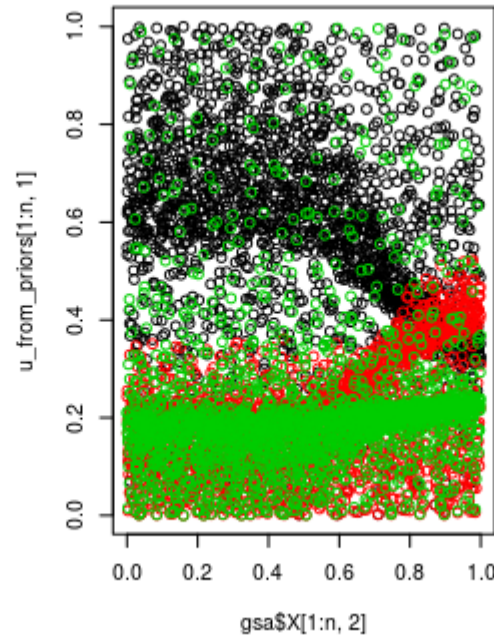
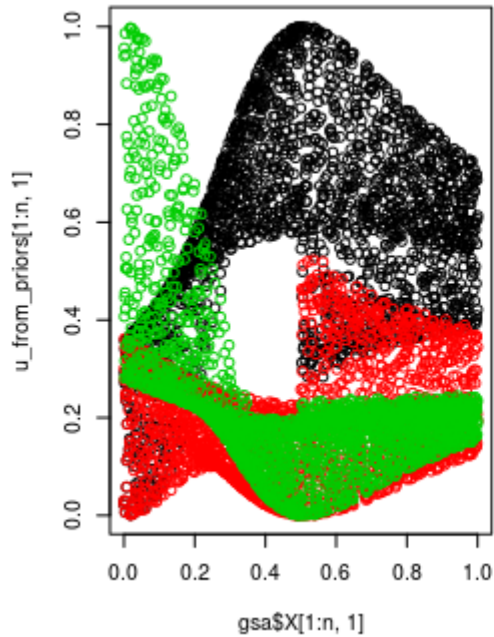
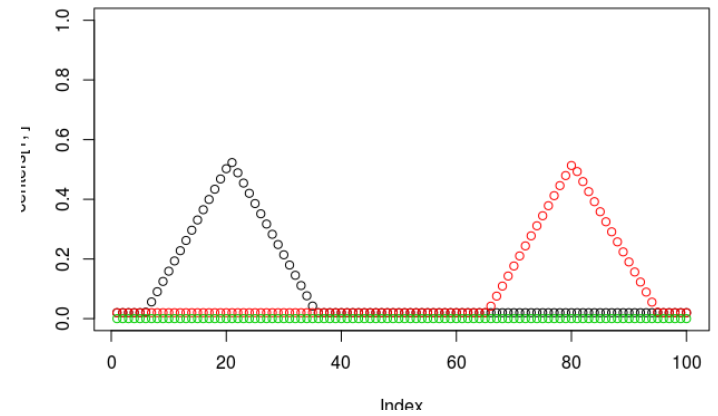
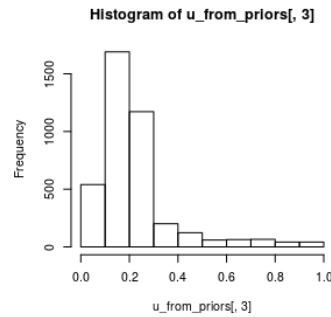
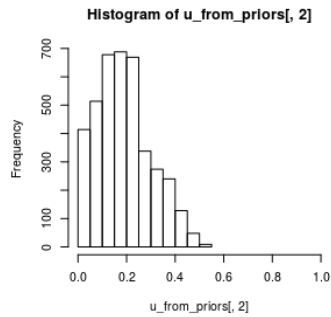
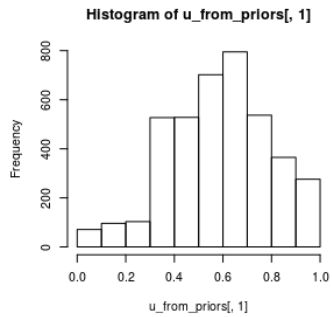
ASMF with prior clusters

Membership functions

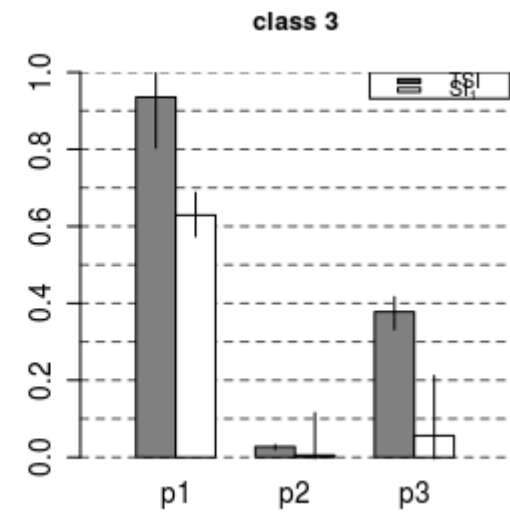
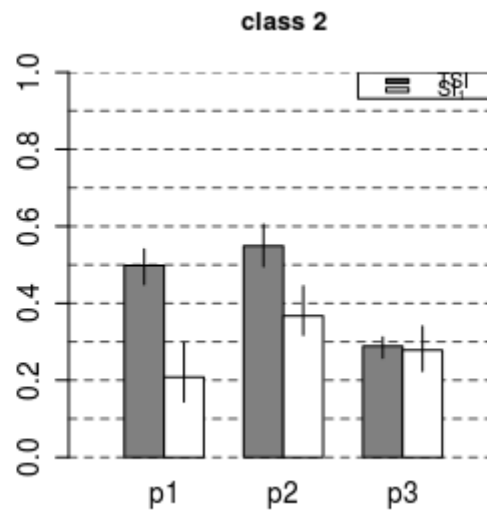
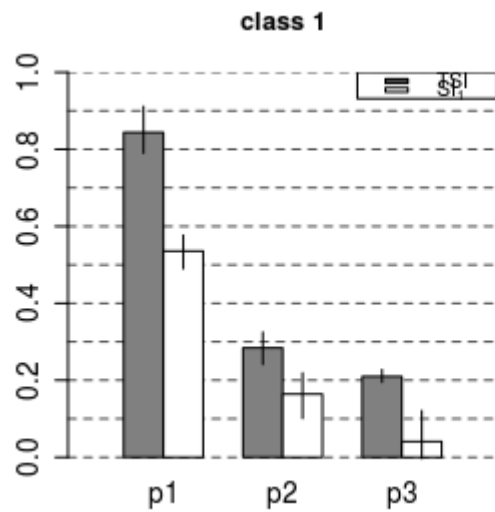
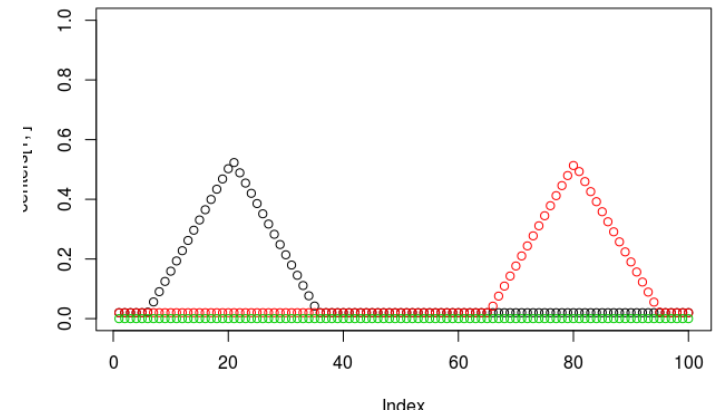
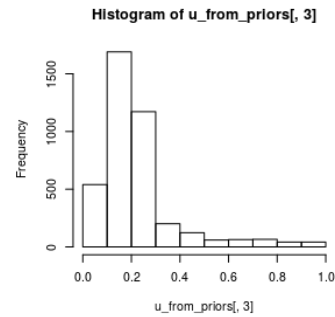
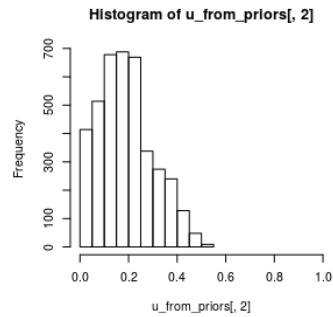
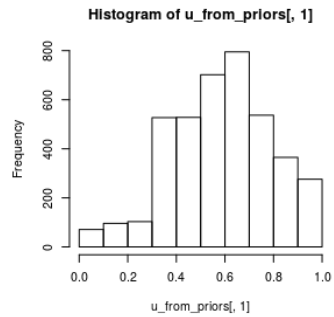


Aucune courbe avec $u_{max2} > 0.5$
(et c'est normal, il n'y a pas de courbe à un seul pic tardif)
Mais il y a des courbes tq $u_2 > u_1$ et $u_2 > u_3$ ie classifié ds le groupe 2 sur un critere de max de MF

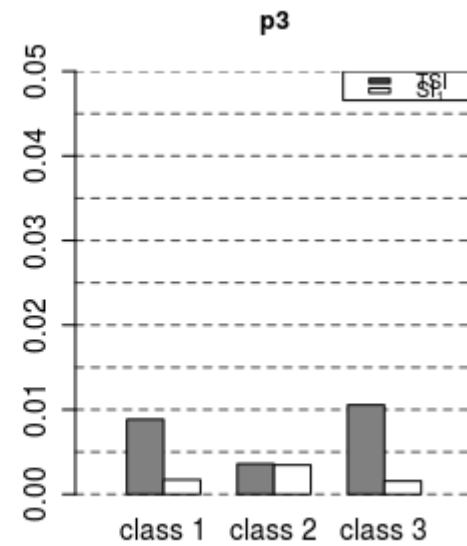
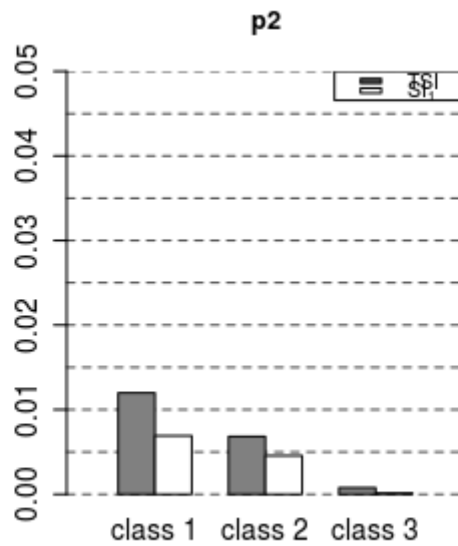
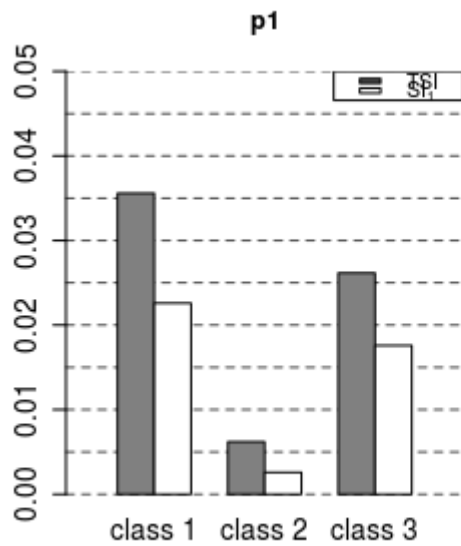
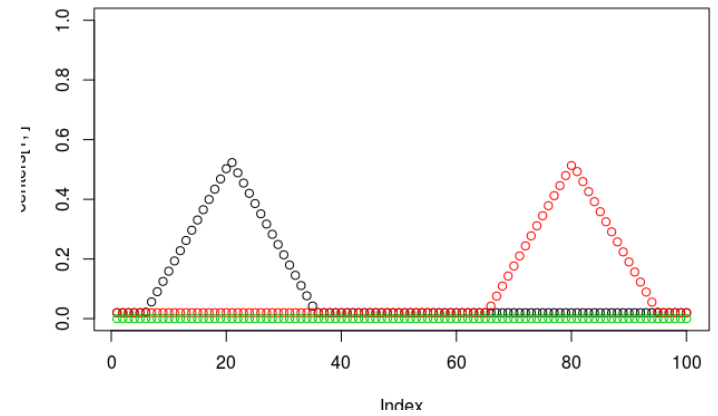
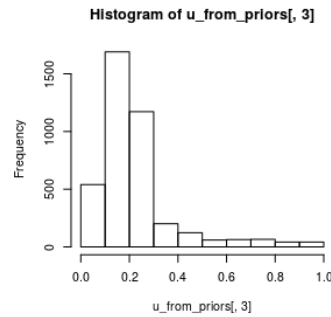
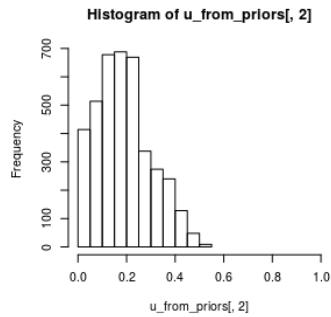
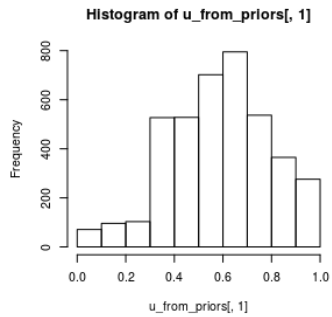
ASMF with prior clusters



ASMF with prior clusters : by class



ASMF with prior clusters : by param

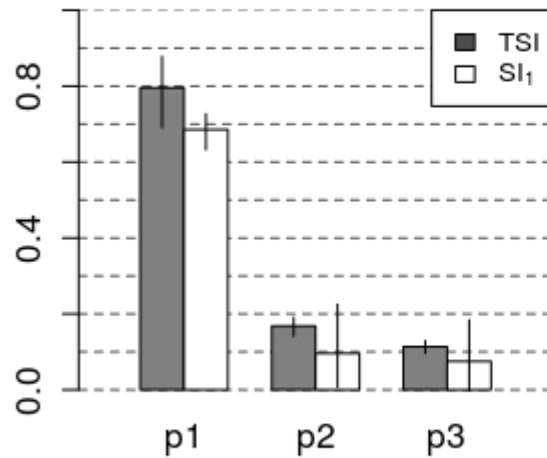


AS on Membership functions Delta

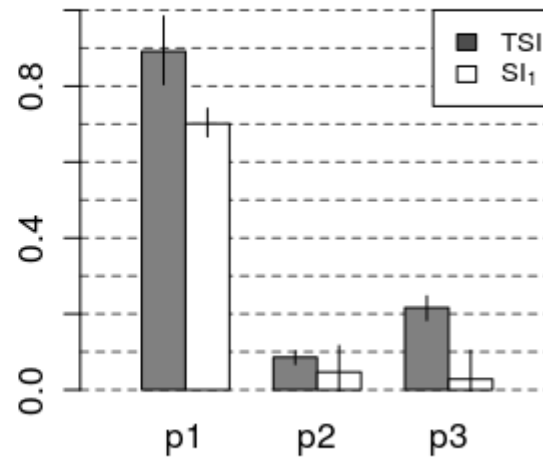
- Quels paramètres expliquent la variabilité des MF entre deux classes cibles (« direction de variation »)
- Sur quelle direction influe le plus un paramètre
- Piste $\Delta_{ij} = u_i - u_j$ (avec le signe)
 - $\Delta_{ij} \in [-1, 1]$
 - Δ_{ij} variant faiblement $\Rightarrow u_i - u_j \approx C$
 - Version crisp : $\Delta_{ij} \Rightarrow 1$ ou -1

AS on Membership functions Delta

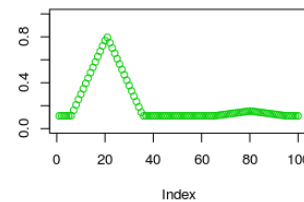
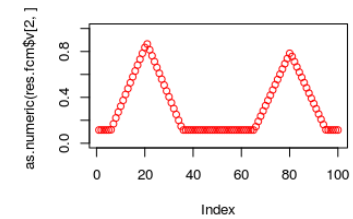
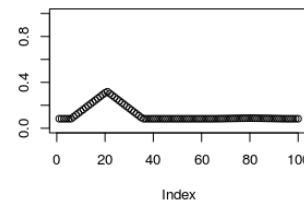
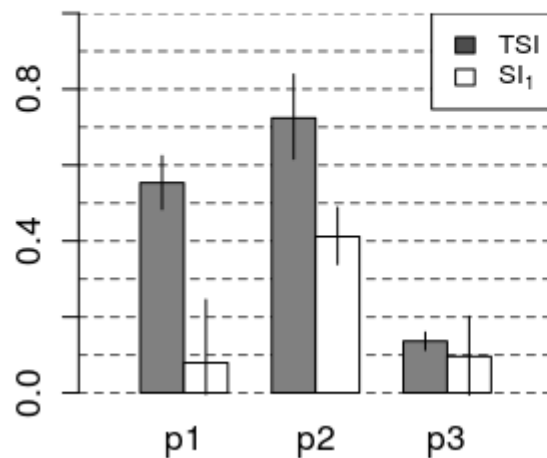
dir 12



dir 13



dir 23



AS on Membership functions Delta

