

# MIXMOD<sup>®</sup>

## Un ensemble logiciel de classification des données par modèles de mélange

Florent Langrognat



**(Lm<sup>B</sup>)**  
laboratoire de mathématiques de besançon  
UNIVERSITÉ DE FRANCHE-COMTÉ • CNRS • UMR 6623



# PLAN

## 1 Pourquoi les modèles de mélanges en classification des données

- La classification des données
- Intérêts des modèles de mélanges

## 2 Mixmod

- Principales fonctionnalités
- Illustrations
- Quelques outils

## 3 Diffusion et valorisation

- Prototype ou logiciel ?
- Un long fleuve tranquille ?
- C'est en marchant qu'on apprend à marcher
- Diffusion, valorisation : des difficultés et des actions concrètes
- La diffusion, la valorisation se mesurent-elles ?

## 4 Perspectives

# MIXMOD

## 1 Pourquoi les modèles de mélanges en classification des données

- La classification des données
- Intérêts des modèles de mélanges

## 2 Mixmod

- Principales fonctionnalités
- Illustrations
- Quelques outils

## 3 Diffusion et valorisation

- Prototype ou logiciel ?
- Un long fleuve tranquille ?
- C'est en marchant qu'on apprend à marcher
- Diffusion, valorisation : des difficultés et des actions concrètes
- La diffusion, la valorisation se mesurent-elles ?

## 4 Perspectives

# MIXMOD

## 1 Pourquoi les modèles de mélanges en classification des données

- La classification des données
- Intérêts des modèles de mélanges

## 2 Mixmod

- Principales fonctionnalités
- Illustrations
- Quelques outils

## 3 Diffusion et valorisation

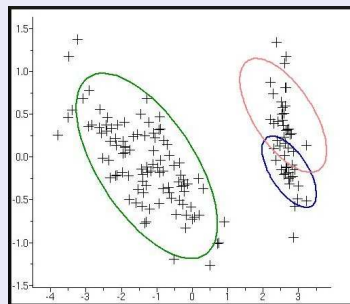
- Prototype ou logiciel ?
- Un long fleuve tranquille ?
- C'est en marchant qu'on apprend à marcher
- Diffusion, valorisation : des difficultés et des actions concrètes
- La diffusion, la valorisation se mesurent-elles ?

## 4 Perspectives

# Classification des données

## Problématiques - Objectifs

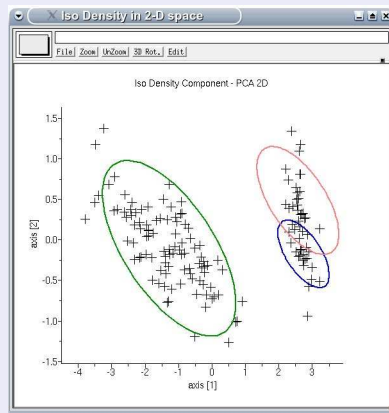
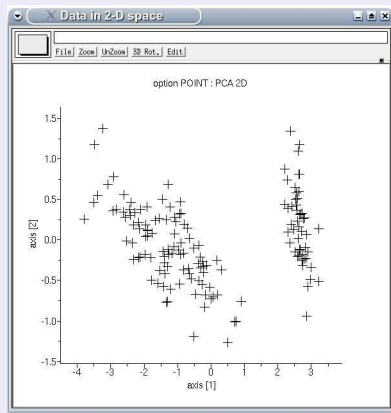
- Trouver une **partition** dans un jeu de données
  - ▶ Trouver le 'bon' **nombre de classes**
  - ▶ **Classer** les individus
- **Caractériser** les classes
- Définir une **règle de classement** à partir d'un échantillon d'apprentissage



# Classification des données

## La question du nombre de classes

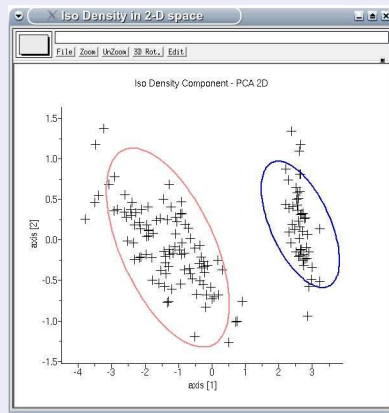
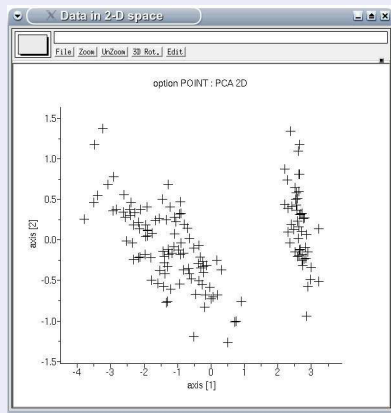
3 classes



# Classification des données

## La question du nombre de classes

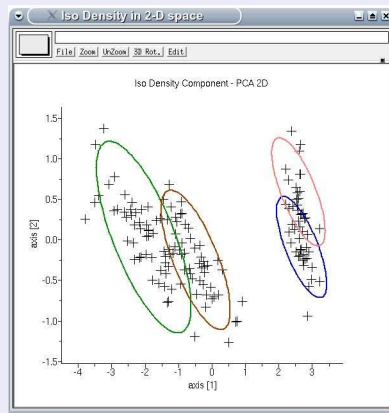
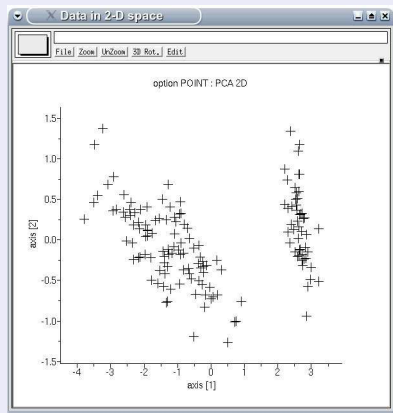
2 classes ?



# Classification des données

## La question du nombre de classes

4 classes ?





# Les différentes problématiques

- Classification non supervisée

- ▶ Choisir une (/ la meilleure) **partition** dans un jeu de données et l'**interpréter**
- ▶ **Classer** les individus

- Classification supervisée

- ▶ **Construire/choisir** une règle de classement à partir d'une **partition connue** et l'**évaluer**
- ▶ **Appliquer** cette règle à de nouveaux individus

- Classement semi supervisée

Entre supervisée et non supervisée !

- ▶ Classer les individus non étiquetés en utilisant toute l'information disponible
- ▶ **Se servir des individus non étiquetés** (en plus de ceux étiquetés) pour construire un règle de classement

# Classification des données

## Pour qui ?

Large spectre de domaines concernés

Génome, reconnaissance de formes, biologie, climatologie, marketing, finance, banque, ...

Exemples :

- Classification d'actifs financiers
- Segmentation de clientèle
- Définition de zones climatiques



# MIXMOD

## 1 Pourquoi les modèles de mélanges en classification des données

- La classification des données
- Intérêts des modèles de mélanges

## 2 Mixmod

- Principales fonctionnalités
- Illustrations
- Quelques outils

## 3 Diffusion et valorisation

- Prototype ou logiciel ?
- Un long fleuve tranquille ?
- C'est en marchant qu'on apprend à marcher
- Diffusion, valorisation : des difficultés et des actions concrètes
- La diffusion, la valorisation se mesurent-elles ?

## 4 Perspectives

## Notations

- **Individus** :  $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$   $n$  vecteurs de :
  - ▶  $\mathbb{R}^d$  (données quantitatives)
  - ▶  $\{1, \dots, m_1\} \times \dots \times \{1, \dots, m_d\}$  (données qualitatives)
- **Partition** :  $\mathbf{z} = (\mathbf{z}_1, \dots, \mathbf{z}_n)$  une matrice binaire où

$$\mathbf{x}_i \in G_k \iff z_{ih} = \mathbb{I}_{\{h=k\}}$$

## Illustrations

Individus $\mathbf{x}$			Partition $\mathbf{z}$		$\Leftrightarrow$	Classe
0.2	0.5	0.8	0	1	$\Leftrightarrow$	$G_2$
0.3	0.1	0.2	1	0	$\Leftrightarrow$	$G_1$
0.9	2.0	-1.2	0	1	$\Leftrightarrow$	$G_2$
0.4	1.2	-0.2	0	1	$\Leftrightarrow$	$G_2$

# Minimisation de l'inertie intra-classe

Sélectionner la partition  $\mathbf{z}$  **minimisant**

$$W_{\mathbf{M}}(\mathbf{z}) = \sum_{i=1}^n \sum_{k=1}^K z_{ik} \|\mathbf{x}_i - \bar{\mathbf{x}}_k\|_{\mathbf{M}}^2$$

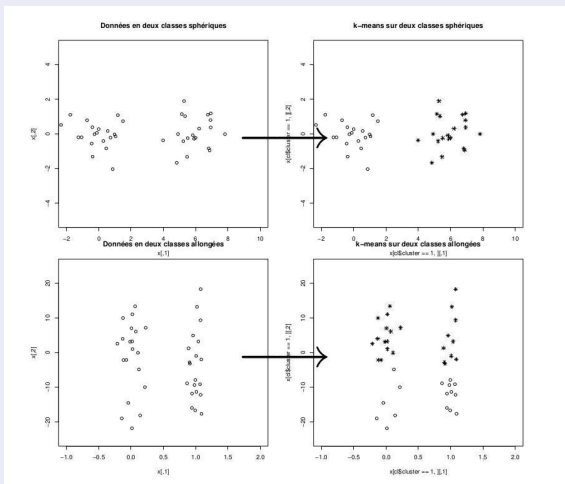
- $\|\cdot\|_{\mathbf{M}}$  est la distance euclidienne avec **métrique**  $\mathbf{M}$  dans  $\mathbb{R}^d$
- $\bar{\mathbf{x}}_k$  est la **moyenne** de la classe  $G_k$

$$\bar{\mathbf{x}}_k = \frac{1}{n_k} \sum_{i=1}^n z_{ik} \mathbf{x}_i$$

et  $n_k = \sum_{i=1}^n z_{ik}$  indique le **nombre d'individus** dans  $G_k$

# Le choix de la métrique

La métrique **M** identité : un choix courant mais risqué



# D'une démarche géométrique vers un modèle de mélange

## Questions en suspens dans l'approche géométrique

- Quelle est la **meilleure métrique**  $\mathbf{M}_{(k)}$  ?
- Choix du **nombre  $K$  de classes** :  $W_{\mathbf{M}}(\mathbf{z})$  décroît avec  $K$ ...

Se doter d'un cadre rigoureux pour répondre à ces questions...

## Idée des modèles de mélange

Individus dans une même classe  $\iff$  Il s'agit d'un échantillon de la même loi de probabilité

# L'hypothèse de mélange de lois paramétriques

- $G_k$  est modélisée par une **distribution paramétrique** :

$$\mathbf{X}_i | Z_{ik}=1 \stackrel{i.i.d.}{\sim} \phi(\cdot; \alpha_k)$$

- $G_k$  a pour probabilité  $\pi_k$  avec  $\sum_{k=1}^K \pi_k = 1$  :

$$\mathbf{Z}_i \stackrel{i.i.d.}{\sim} \text{Mult}_K(1, \pi_1, \dots, \pi_K)$$

Loi mélange de paramètre  $\theta = (\pi_1, \dots, \pi_K, \alpha_1, \dots, \alpha_K)$

$$f(\mathbf{x}_i; \theta) = \sum_{k=1}^K \pi_k \phi(\mathbf{x}_i; \alpha_k)$$



# 1<sup>er</sup> intérêt des modèles de mélanges

## Nombreux algorithmes d'estimation

### ● EM

- ▶ Algorithme de référence pour maximiser la (log)vraisemblance observée
- ▶ Croissance de la vraisemblance à chaque itération
- ▶ Convergence parfois lente, nombreux points stationnaires

### ● CEM (EM version Classification)

- ▶ Permet d'estimer simultanément les paramètres du mélange et les labels des individus par la maximisation de la vraisemblance complétée
- ▶ Convergence en un nombre fini d'itérations
- ▶ Estimateurs **biaisés** des paramètres du mélange

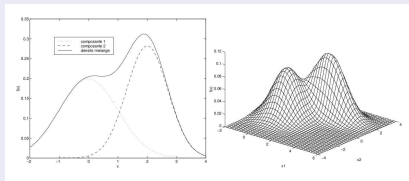
### ● SEM (EM version Stochastique)

- ▶ Permet d'estimer simultanément les paramètres du mélange et les labels des individus par la maximisation de la vraisemblance complétée
- ▶ SEM engendre une chaîne de Markov dont la **distribution stationnaire** se concentre autour de l'estimateur du maximum de vraisemblance des paramètres du mélange.

## 2<sup>e</sup> intérêt des modèles de mélange

### Des modèles gaussiens pour les données quantitatives

$$\phi(\cdot; \alpha_k) = \mathbf{N}_d(\mu_k, \Sigma_k) \quad \text{où} \quad \alpha_k = \left( \underbrace{\mu_k}_{\text{moyenne}}, \underbrace{\Sigma_k}_{\text{matrice de variance}} \right)$$



### Des modèles multinomiaux pour les données qualitatives

$$f(\mathbf{x}_i; \theta) = \sum_k \pi_k m_k(\mathbf{x}_i; \alpha_k) = \sum_k \pi_k \prod_{j,h} (\alpha_k^{jh})^{x_i^{jh}}$$

où  $\theta = (\pi_1, \dots, \pi_g, \alpha_1^{11}, \dots, \alpha_g^{d m_d})$  avec

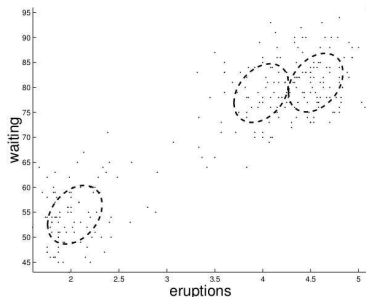
- $\alpha_k^{jh}$  : probabilité que la variable  $j$  présente la modalité  $h$  dans la classe  $k$ ,
- $\pi_k$  : proportions du mélange

# 3<sup>e</sup> intérêt des modèles de mélange

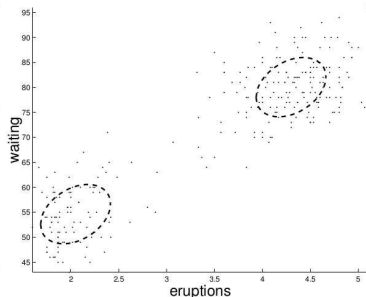
## Des critères de sélection

$$\underbrace{\text{Critère}}_{\text{à maximiser}} = \underbrace{\text{maximum log vraisemblance}}_{\text{adéquation modèle-données}} - \underbrace{\text{pénalité}}_{\text{"coût" du modèle}}$$

## Exemple : the old Faithful Geyser



BIC :  $(K = 3, [\pi \lambda D A D'])$



ICL :  $(K = 2, [\pi_k \lambda_k D A_k D'])$

# MIXMOD

## 1 Pourquoi les modèles de mélanges en classification des données

- La classification des données
- Intérêts des modèles de mélanges

## 2 Mixmod

- Principales fonctionnalités
- Illustrations
- Quelques outils

## 3 Diffusion et valorisation

- Prototype ou logiciel ?
- Un long fleuve tranquille ?
- C'est en marchant qu'on apprend à marcher
- Diffusion, valorisation : des difficultés et des actions concrètes
- La diffusion, la valorisation se mesurent-elles ?

## 4 Perspectives

# MIXMOD

## 1 Pourquoi les modèles de mélanges en classification des données

- La classification des données
- Intérêts des modèles de mélanges

## 2 Mixmod

- Principales fonctionnalités
- Illustrations
- Quelques outils

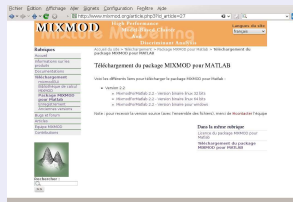
## 3 Diffusion et valorisation

- Prototype ou logiciel ?
- Un long fleuve tranquille ?
- C'est en marchant qu'on apprend à marcher
- Diffusion, valorisation : des difficultés et des actions concrètes
- La diffusion, la valorisation se mesurent-elles ?

## 4 Perspectives

## Fiche d'identité

- Projet débuté en **2001**
- Compétences complémentaires en **informatique** et **statistiques**
- Site web : **www.mixmod.org**
- Diffusion : **Licence GNU GPL**



## Un Comité de Pilotage

Des compétences complémentaires en **statistiques** et **informatique**

- **C. Biernacki** (Labo. Paul Painlevé - Univ. Lille 1-CNRS)
- **G. Celeux** (Projet Select - INRIA Saclay)
- **G. Govaert** (Labo. Heudiasyc - UTC/CNRS)
- **F. Langrognnet** (LMB - UFC-CNRS)



# Fonctionnalités (1)

## Problématiques traitées

- Classification non supervisée
- Classification supervisée (analyse discriminante)

## Cadre de travail - Type de données traitées

### Modèles de mélanges

- Gaussiens (données quantitatives)
- Multinomiaux (données qualitatives)
- Modèles spécifiques pour les données en grande dimension

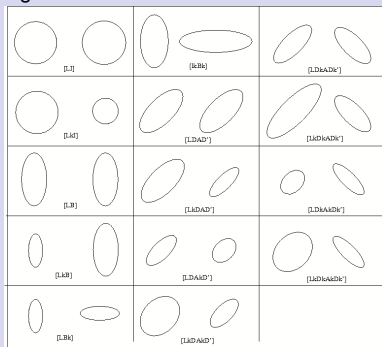
# Fonctionnalités (2)

## Modèles et métriques

### Données quantitatives

#### 14 modèles gaussiens

basés sur la décomposition en valeur sigilière de la matrice de variance



### Données quantitatives en grande dimension

#### 8 modèles spécifiques pour la grande dimension

$$\Delta_k = \left( \begin{array}{ccc|ccc} a_{k1} & & 0 & & & \\ & \ddots & & & & \\ 0 & & a_{k\delta_k} & & & \\ \hline & & & b_k & & 0 \\ & & & & \ddots & \\ & & 0 & & & b_k \end{array} \right) \left. \begin{array}{l} \} \\ \} \\ \} \end{array} \right\} \begin{array}{l} \delta_k \\ \\ (d - \delta_k) \end{array}$$

### Données qualitatives

#### 5 modèles multinomiaux

basés sur une reparamétrisation de la distribution de Bernoulli



# Fonctionnalités (3)

## Algorithmes

Maximisation de la vraisemblance (ou vraisemblance complétée)

- **EM** (Expectation Maximisation)
- **SEM** (Stochastic EM)
- **CEM** (Classification EM)

## Critères

- **BIC** (Bayesian Information Criterion)
- **ICL** (Integrated Completed Likelihood)
- **NEC** (Normalized Entropy Criterion)
- **CV** (Cross Validation)

## Initialisations et Stratégies

- **6 initialisations**  
Ex : 'random', 'short runs of EM',...
- **Algorithmes chaînés**  
Ex : 100 iterations de **SEM** puis 50 iterations de **EM**

## Et aussi...

- Connaissance partielle des labels des individus (**semi-supervisé**)
- Individus **pondérés**

# MIXMOD

## 1 Pourquoi les modèles de mélanges en classification des données

- La classification des données
- Intérêts des modèles de mélanges

## 2 Mixmod

- Principales fonctionnalités
- **Illustrations**
- Quelques outils

## 3 Diffusion et valorisation

- Prototype ou logiciel ?
- Un long fleuve tranquille ?
- C'est en marchant qu'on apprend à marcher
- Diffusion, valorisation : des difficultés et des actions concrètes
- La diffusion, la valorisation se mesurent-elles ?

## 4 Perspectives

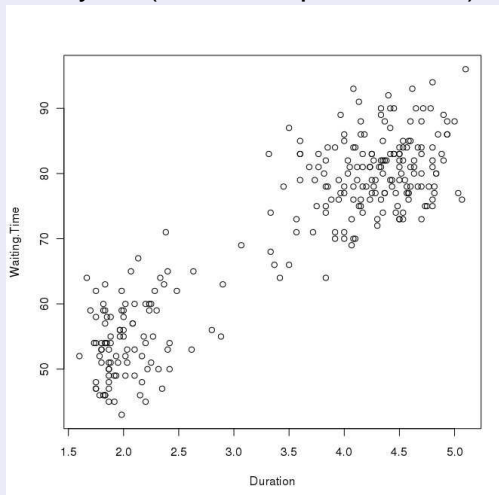
## Illustration 1

Classification non supervisée sur données quantitatives

# Classification non supervisée

## Illustration

### Geyser (données quantitatives)



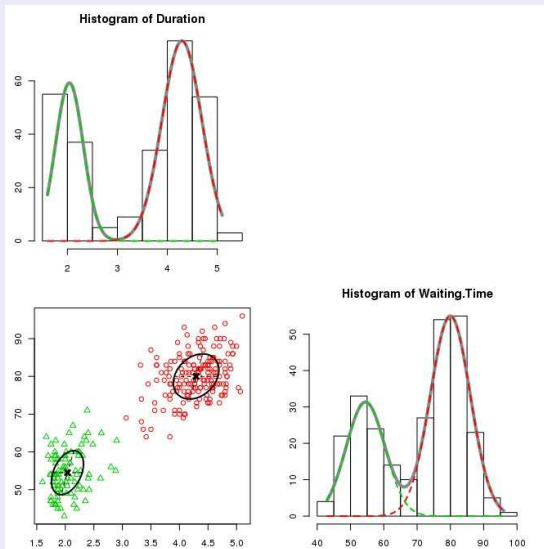
# Illustration 1

## Commandes Rmixmod

```
> data(geyser)
> out<-mixmodCluster(geyser, nbCluster=2)
> summary(out)
*****
* Number of samples      = 272
* Problem dimension      = 2
*****
*      Number of cluster = 2
*      Model Type        = Gaussian_pk_Lk_C
*      Criterion         = BIC(2322.9719)
*      Parameters        = list by cluster
*      Cluster 1 :
*          Proportion    = 0.6429
*          Means         = 4.2922 79.9964
*          Variances     = | 0.1453 0.8301 |
*                       | 0.8301 40.9022 |
*      Cluster 2 :
*          Proportion    = 0.3571
*          Means         = 2.0397 54.5171
*          Variances     = | 0.0984 0.5618 |
*                       | 0.5618 27.6831 |
*      Log-likelihood    = -1136.2599
*****
```

# Illustration 1

## Plot



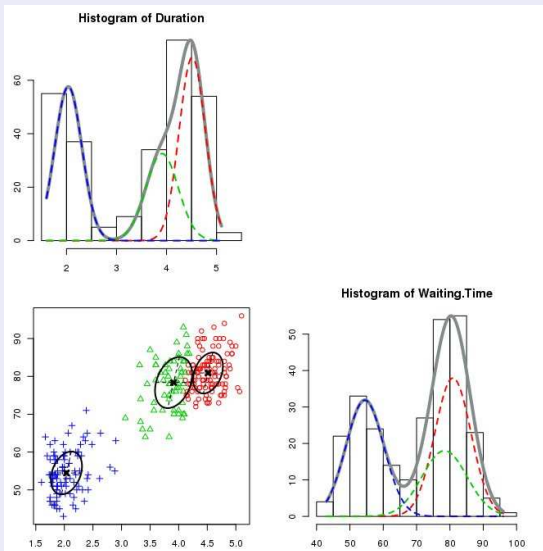
# Plusieurs modèles et plusieurs critères

## Commandes Rmixmod

```
> out<-mixmodCluster(geyser, nbCluster=2:3, criterion=c("BIC","ICL"),
+ models=mixmodGaussianModel())
> summary(out)
*****
* Number of samples      = 272
* Problem dimension      = 2
*****
*      Number of cluster = 3
*      Model Type        = Gaussian_p_L_C
*      Criterion         = BIC(2312.6006) ICL(2377.6923)
*      Parameters        = list by cluster
*      Cluster 1 :
*          Proportion    = 0.3333
*          Means          = 3.9765 78.7195
*          Variances     = | 0.0798 0.5341 |
*                       | 0.5341 34.2108 |
*
*      Cluster 2 :
*          Proportion    = 0.3333
*          Means          = 4.5545 81.0528
*          Variances     = | 0.0798 0.5341 |
*                       | 0.5341 34.2108 |
*
*      Cluster 3 :
*          Proportion    = 0.3333
*          Means          = 2.0390 54.5083
*          Variances     = | 0.0798 0.5341 |
*                       | 0.5341 34.2108 |
*
*      Log-likelihood    = -1131.0742
*****
```

# Illustration 1 bis

## plot (3 classes)





# Illustration 1 bis

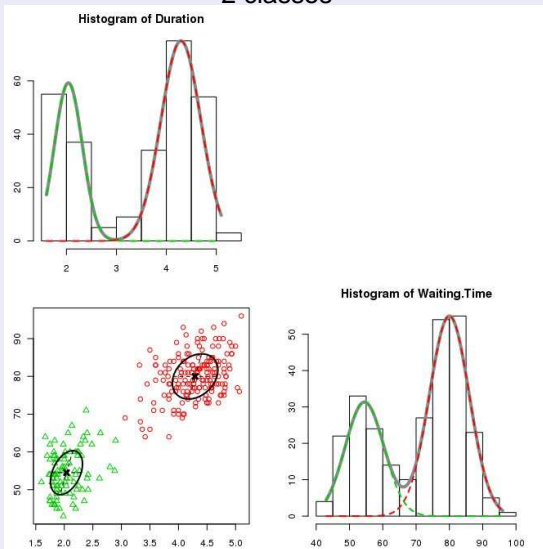
## Tri selon ICL

```
> out_icl<-sortByCriterion(out, "ICL")
> summary(out_icl)
*****
* Number of samples      = 272
* Problem dimension      = 2
*****
*
*   Number of cluster = 2
*   Model Type = Gaussian_pk_Lk_D_Ak_D
*   Criterion = BIC(2320.2833) ICL(2320.5794)
*   Parameters = list by cluster
*   Cluster 1 :
*       Proportion = 0.6432
*       Means = 4.2915 79.9893
*       Variances = | 0.1588 0.6810 |
*                   | 0.6810 35.7667 |
*
*   Cluster 2 :
*       Proportion = 0.3568
*       Means = 2.0387 54.5041
*       Variances = | 0.0783 0.6467 |
*                   | 0.6467 33.8930 |
*
*   Log-likelihood = -1132.1126
*****
```

# Illustration 1 bis

plot (2 classes)

2 classes

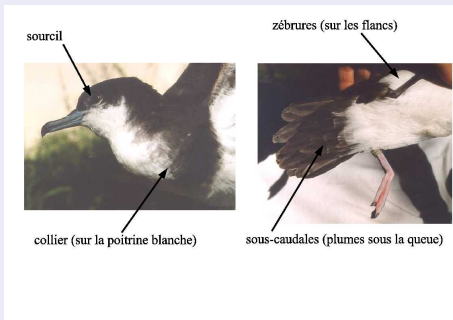


## Illustration 2

### Classification supervisée sur données qualitatives

# Classification supervisée

## Illustration (données qualitatives) : puffins



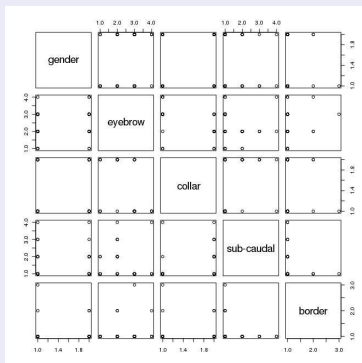
variable	nombre de niveaux de réponse	valeurs
sexe	2	mâle, femelle
sourcils	5	absent -> très prononcé
collier	6	absent -> continu
sous-caudales	5	blanc, noir, noir&blanc, noir&BLANC, NOIR&blanc
liseret	4	absent, . . . , beaucoup

# Observations

## Données

- Nombre d'**individus** :  $n = 69$
- Nombre d'**espèces (classes)** :  $K = 2$
- Nombre de **variables** :  $d = 5$
- Individu  $i$  :  $(x_i, z_i) = ((x_i^j)_{j=1, \dots, d}, z_i)$

$n^0$	$z_i$	$x_i^1$	$x_i^2$	$x_i^3$	$x_i^4$	$x_i^5$
1	1	1	2	2	2	2
2	1	2	1	3	3	1
⋮	⋮			⋮		
68	2	1	4	1	2	1
69	2	1	3	1	2	1



# Illustration 2 - Apprentissage

## Apprentissage (1)

```
> data(birds)
> learn<-mixmodLearn(birds, birdsLabel)
> summary(learn)
*****
* Number of samples = 69
* Problem dimension = 5
*****
*      Number of cluster = 2
*      Model Type = Binary_pk_EKjh
*      Criterion = CV(0.9855)
*      Parameters = list by cluster
*      Cluster 1 :
*          Proportion = 0.6667
*          Center = 1.0000 3.0000 1.0000 1.0000 1.0000
*          Scatter = |      0.4787      0.4787 |
*                   |      0.0691      0.0266      0.1862      0.0904 |
*                   |      0.1660      0.1532      0.0043      0.0043      0.0043 |
*                   |      0.0383      0.0043      0.0043      0.0255      0.0043 |
*                   |      0.0780      0.0496      0.0284 |
*      Cluster 2 :
*          Proportion = 0.3333
*          Center = 2.0000 2.0000 2.0000 2.0000 1.0000
*          Scatter = |      0.3958      0.3958 |
*                   |      0.1354      0.1562      0.0104      0.0104 |
*                   |      0.0500      0.0750      0.0083      0.0083      0.0083 |
*                   |      0.3417      0.5333      0.1333      0.0500      0.0083 |
*                   |      0.0694      0.0556      0.0139 |
*      Log-likelihood = -203.5308
*****
```

# Illustration 2 - Apprentissage

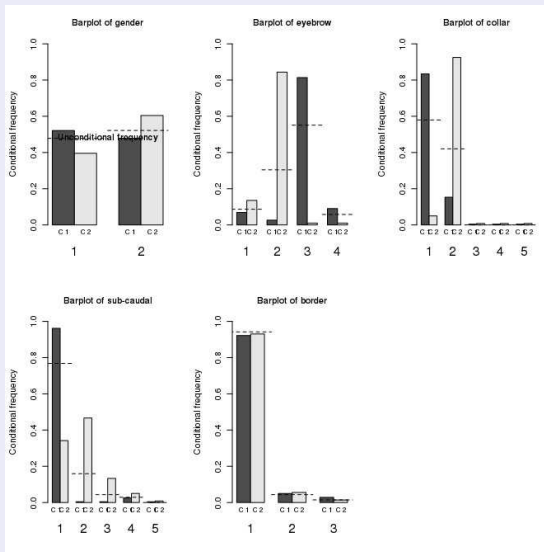
## Apprentissage (2)

```
> learn@bestResult
* nbCluster = 2
* model name = Binary_pk_Ekjh
* criterion = CV(0.9855)
* likelihood = -203.5308
*****
* number of modalities = 2 4 5 3
*** Cluster 1
* proportion = 0.6667
* center = 1.0000 3.0000 1.0000 1.0000 1.0000
* scatter = |
            | 0.4787 0.4787 |
            | 0.0691 0.0266 | 0.1862 0.0904 | |
            | 0.1660 0.1532 | 0.0043 0.0043 | 0.0043 |
            | 0.0383 0.0043 | 0.0043 0.0255 | 0.0043 |
            | 0.0780 0.0496 | 0.0284 |
*** Cluster 2
* proportion = 0.3333
* center = 2.0000 2.0000 2.0000 2.0000 1.0000
* scatter = |
            | 0.3958 0.3958 |
            | 0.1354 0.1562 | 0.0104 0.0104 | |
            | 0.0500 0.0750 | 0.0083 0.0083 | 0.0083 |
            | 0.3417 0.5333 | 0.1333 0.0500 | 0.0083 |
            | 0.0694 0.0556 | 0.0139 |
*****
* Classification with CV:
      | Cluster 1 | Cluster 2 |
-----|-----|-----|
Cluster 1 | 45 | 0 |
Cluster 2 | 1 | 23 |
-----|-----|-----|
* Error rate with CV = 1.45 %

* Classification with MAP:
      | Cluster 1 | Cluster 2 |
-----|-----|-----|
Cluster 1 | 46 | 0 |
Cluster 2 | 0 | 23 |
-----|-----|-----|
* Error rate with MAP = 0.00 %
*****
```

# Illustration 2 - Apprentissage

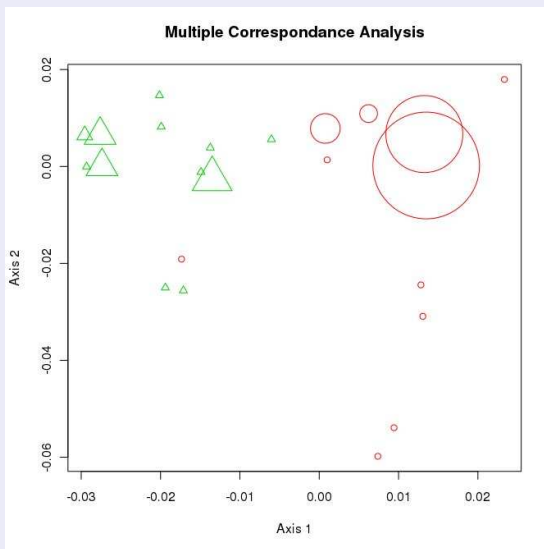
## Visualisation (1)





# Illustration 2 - Apprentissage

## Visualisation (2)



# Illustration 2 - Classement

## Classement

```
> other_birds
  gender      poor eyebrow collar sub-caudal border
1  male poor pronounced dotted      white  few
2  female      none dotted      black  none
3  female pronounced none      white  none
4  male pronounced dotted      white  none
5  male pronounced dotted      white  none
6  male pronounced dotted      white  none
7  female pronounced dotted      white  none
8  female poor pronounced dotted      white  none
9  female poor pronounced dotted      white  none
10 female poor pronounced dotted      white  none
> predict<-mixmodPredict(other_birds, classificationRule=learn["bestResult"])
> summary(predict)
*****
* partition      = 2 2 1 1 1 1 1 2 2 2
* probabilities = | 0.0334 0.9666 |
                  | 0.0012 0.9988 |
                  | 0.9998 0.0002 |
                  | 0.9896 0.0104 |
                  | 0.9896 0.0104 |
                  | 0.9896 0.0104 |
                  | 0.9828 0.0172 |
                  | 0.0226 0.9774 |
                  | 0.0226 0.9774 |
                  | 0.0226 0.9774 |
*****
_
```

# MIXMOD

## 1 Pourquoi les modèles de mélanges en classification des données

- La classification des données
- Intérêts des modèles de mélanges

## 2 Mixmod

- Principales fonctionnalités
- Illustrations
- **Quelques outils**

## 3 Diffusion et valorisation

- Prototype ou logiciel ?
- Un long fleuve tranquille ?
- C'est en marchant qu'on apprend à marcher
- Diffusion, valorisation : des difficultés et des actions concrètes
- La diffusion, la valorisation se mesurent-elles ?

## 4 Perspectives

# Système de gestion de version

## De CVS à SVN

Passage de CVS à SVN en 2008

- Serveur CVS géré en local **Vs** Utilisation de SVN à travers une Forge
- Avantages de SVN :
  - ▶ Renommage et déplacement de fichiers sans perte de l'historique
  - ▶ Gestion des répertoires
  - ▶ Gestion des metadonnées (ex : permissions)
  - ▶ ...
- Historique conservé grâce à **cvs2svn**

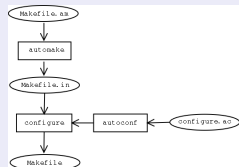


# Outils de construction

## Des autotools à cmake

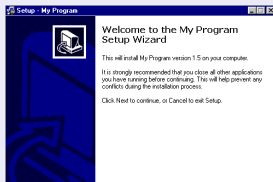
### Migration des autotools à cmake vers 2007

- Lourdeur des autotools Vs Simplicité de cmake
- cmake
  - ▶ Multi OS
  - ▶ Intégré dans les IDE
  - ▶ Outils complémentaires : cpack, ctest



## Innosetup

### Construction d'exécutables d'installation sous windows



# Environnements de Développement Intégré

## IDE

En fonction des habitudes, des goûts des développeurs. . .

KDevelop



Eclipse



## INRIAGForge

The screenshot shows the INRIAGForge web interface. At the top, there is a navigation bar with tabs for 'Accueil', 'Ma page', and 'Projets'. Below this, there are sub-tabs for 'Résumé', 'Administration', 'Activité', 'Forums', 'Outil de suivi', 'Listes', 'Tâches', 'Documents', 'Annonces', 'Sources', and 'Fichiers'. The main content area is titled 'Dépôt de sources' and includes a link to 'Voir le code source | Rapport | Administration'. It contains two sections: 'Accès Subversion pour les développeurs via SSH' and 'Accès Subversion pour les développeurs via WebDAV'. On the right side, there is a 'Historique du dépôt' section with a table of statistics.

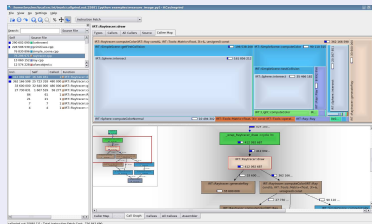
Historique du dépôt		
Données sur les états présent et passé du dépôt		
Statistiques sur le dépôt		
Nom	Ajouts	Envois
Florent Langrognet	21957	6031
...	960	7314
...	682	3246
...	558	1123
...	252	776
...	189	683
...	0	396
...	115	244
...	267	26
...	0	21
...	3	2
<b>Total:</b>	<b>24983</b>	<b>19862</b>

- SVN
- Forums
- Gestion des tâches
- Suivi de bugs
- Liste de diffusion

# Outils de profilage

## et d'études de problèmes liés à la mémoire

### Valgrind, KCacheGrind

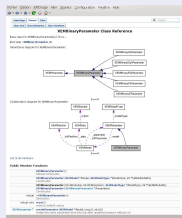




# Documentation

## Quels documents ?

- Guide utilisateur
- Documentation statistique
- Documentation logicielle



## Outils



Package ROxygen

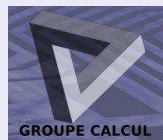


# Quels sont les 'bons' outils ?

## Comment trouver l'information ?

ou l'intérêt de la veille technologique

- Internet
  - ▶ Recherche avec mots clés
  - ▶ Sites spécialisés (PLUME)
- Réseaux métiers
  - ▶ Réseaux régionaux (DevelopR6) - Retours d'expériences
  - ▶ Réseaux nationaux (Calcul, DevLog)
- Livres, revues
- Ecole (ENVOL), Congrès, ...



# MIXMOD

## 1 Pourquoi les modèles de mélanges en classification des données

- La classification des données
- Intérêts des modèles de mélanges

## 2 Mixmod

- Principales fonctionnalités
- Illustrations
- Quelques outils

## 3 Diffusion et valorisation

- Prototype ou logiciel ?
- Un long fleuve tranquille ?
- C'est en marchant qu'on apprend à marcher
- Diffusion, valorisation : des difficultés et des actions concrètes
- La diffusion, la valorisation se mesurent-elles ?

## 4 Perspectives

# MIXMOD

## 1 Pourquoi les modèles de mélanges en classification des données

- La classification des données
- Intérêts des modèles de mélanges

## 2 Mixmod

- Principales fonctionnalités
- Illustrations
- Quelques outils

## 3 Diffusion et valorisation

- **Prototype ou logiciel ?**
- Un long fleuve tranquille ?
- C'est en marchant qu'on apprend à marcher
- Diffusion, valorisation : des difficultés et des actions concrètes
- La diffusion, la valorisation se mesurent-elles ?

## 4 Perspectives

# Développement : prototype ou logiciel ?

Concilier des approches, des intérêts parfois (très) différents ?



Prototype

- Rapidement (pour hier !)
- Résultats numériques à intégrer dans un article
- Quid de la diffusion, la licence ?
- Quid de la fiabilité, reproductibilité ?



Logiciel

- Besoin de temps
- Architecture solide permettant des évolutions
- Logiciel validé
  - ▶ Absence de bugs (on presque)
  - ▶ Tests unitaires, non régression
  - ▶ Reproductibilité
- Diffusion, licence
- Support utilisateurs, documentations, ...

# Le développement et la diffusion de logiciels

## Une production scientifique

- **Valorisation** du travail de recherche et de développement
- **Diffusion** du savoir faire
- Rôle central pour la **reproductibilité**
- **Publications**, reconnaissance
  - ▶ Journaux dédiés
  - ▶ Comment juger de la qualité d'un logiciel ?  
Rewieving



**Les compétences complémentaires du chercheur et de l'informaticien sont nécessaires**

# MIXMOD

## 1 Pourquoi les modèles de mélanges en classification des données

- La classification des données
- Intérêts des modèles de mélanges

## 2 Mixmod

- Principales fonctionnalités
- Illustrations
- Quelques outils

## 3 Diffusion et valorisation

- Prototype ou logiciel ?
- **Un long fleuve tranquille ?**
- C'est en marchant qu'on apprend à marcher
- Diffusion, valorisation : des difficultés et des actions concrètes
- La diffusion, la valorisation se mesurent-elles ?

## 4 Perspectives

# Diffusion d'un logiciel

## La diffusion (et la valorisation) n'est pas un long fleuve tranquille

Diffusion et valorisation peuvent évoluer dans le temps en fonction :

- des **objectifs**
- des **envies**
- des **rencontres** (congrès, réseaux, ...)
- des **connaissances** en terme de
  - ▶ Actions de valorisation
  - ▶ Acteurs de la valorisation
  - ▶ Licences
- Des demandes des utilisateurs
- Du succès (ou non) de la diffusion





# Diffuser : oui mais pour qui ?

## Quelques questions à se poser

- A t-on identifié les utilisateurs potentiels (type et nombre) ?
- Connait-on leurs attentes ?
- A t-on pris suffisamment en compte l'activité de support ?

## La réalité

- On découvre des utilisateurs qui n'étaient pas la cible de départ
- Le nombre d'utilisateurs peut être (très) différent des estimations

## Adapter la diffusion aux utilisateurs

Moyens de diffusions, produits logiciels, type de support,...

# MIXMOD

## 1 Pourquoi les modèles de mélanges en classification des données

- La classification des données
- Intérêts des modèles de mélanges

## 2 Mixmod

- Principales fonctionnalités
- Illustrations
- Quelques outils

## 3 Diffusion et valorisation

- Prototype ou logiciel ?
- Un long fleuve tranquille ?
- **C'est en marchant qu'on apprend à marcher**
- Diffusion, valorisation : des difficultés et des actions concrètes
- La diffusion, la valorisation se mesurent-elles ?

## 4 Perspectives

# 1<sup>re</sup> période : à partir de 2000

## Objectif

Développer un logiciel de classification des données efficace, robuste et  
le diffuser au sein de la communauté scientifique

## Composants logiciels

- Une bibliothèque de calcul : **mixmodLib**
- Des packages pour des environnements tiers : **mixmodForMatlab** et **mixmodForScilab**

## Moyens de diffusion

- Site web [www-math.univ-fcomte.fr](http://www-math.univ-fcomte.fr) devenu **[www.mixmod.org](http://www.mixmod.org)**
- Licence GNU GPL

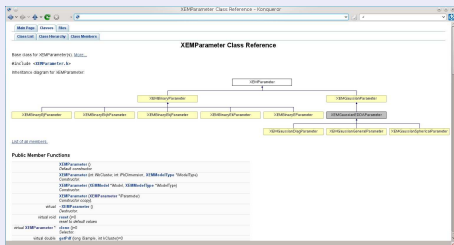
# 1<sup>er</sup> composant logiciel

mixmodLib

Bibliothèque de calcul

## Bibliothèque de calcul

- Ensemble de classes C++
  - ▶ Architecture évolutive
  - ▶ env. 100 classes et 35 000 lignes de code C++
- Qualités
  - ▶ Rapidité
  - ▶ Robustesse
  - ▶ Evolutivité
- Env. 500 téléchargements par an depuis 2001



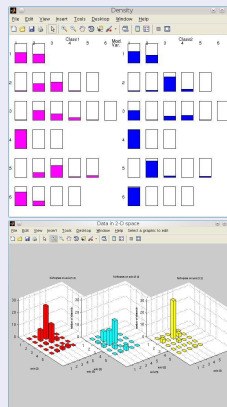
## 2<sup>e</sup> composant logiciel

mixmodForMatlab

Package pour Matlab

## Package pour Matlab

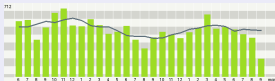
- Interface de mixmodLib pour Matlab
- Ensemble de fonctions Matlab :
  - ▶ Classification supervisée et non supervisée
  - ▶ Outils de visualisation
- Les calculs sont effectués dans mixmodLib
- Disponible depuis 2003  
env. 300 téléchargements par an



## 2<sup>e</sup> étape : à partir de 2003

### La situation évolue

- La “renommée” de Mixmod croît  
Environ 500 visites et 150 téléchargements par mois



- Mixmod touche de plus en plus d’**utilisateurs “non spécialistes”**

### Conséquences

**Demande croissante** en terme de

- Support
- Demandes d’évolution
- Utilisation de Mixmod  
Les outils initialement mis à disposition ne conviennent pas toujours à cette nouvelle population d’utilisateurs



# Conséquences

## Composants logiciels

- **Enrichissement** des packages pour Matlab et Scilab
  - ▶ Utilisation plus aisée (pour les utilisateurs non spécialistes)
  - ▶ Fonctions graphiques enrichies

## Diffusion

- **Dépôt à l'APP** (Agence de Protection des Programmes)
  - ▶ A conseiller vivement
  - ▶ Procédure **simple** (en théorie) mais **longue** (en pratique) avec 5 organismes concernés
- Licence **GNU GPL** favorise la diffusion, l'utilisation au sein de la communauté scientifique



## 3<sup>e</sup> étape (à partir de 2007-2008)

### Constats

- Hétérogénéité encore plus grande des utilisateurs  
Comment proposer des outils pour le **spécialiste** et le **novice** ?
- Nécessité de répondre à certaines demandes (**package pour R**)

Mais aussi :

- Difficulté à développer, maintenir plusieurs composants logiciels très différents

### Evolutions des composants logiciels

- Abandon de **mixmodForScilab**
- Développement d'un logiciel avec interface graphique : **mximodGUI**
- Développement d'un package pour R : **RMixmod**

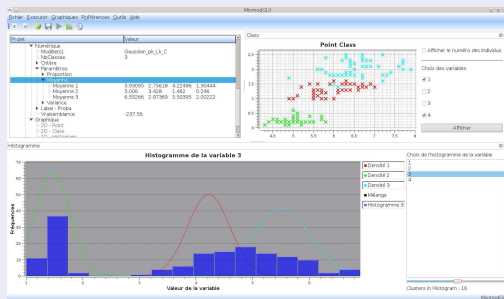
# 3<sup>e</sup> composant logiciel

mixmodGUI

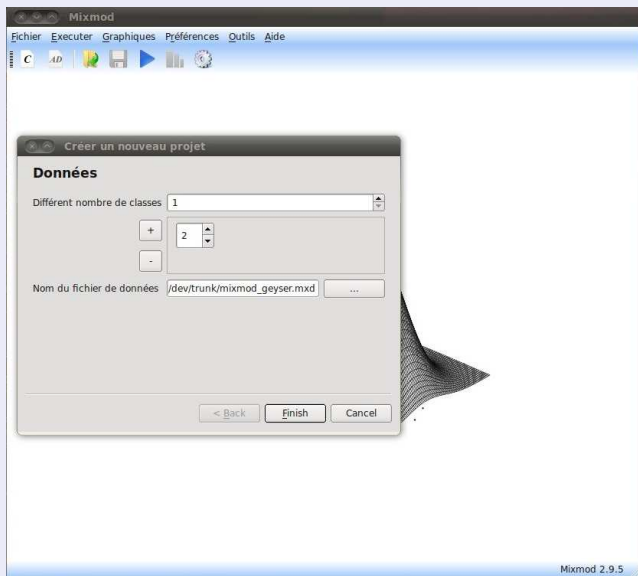
Interface graphique

## Interface graphique

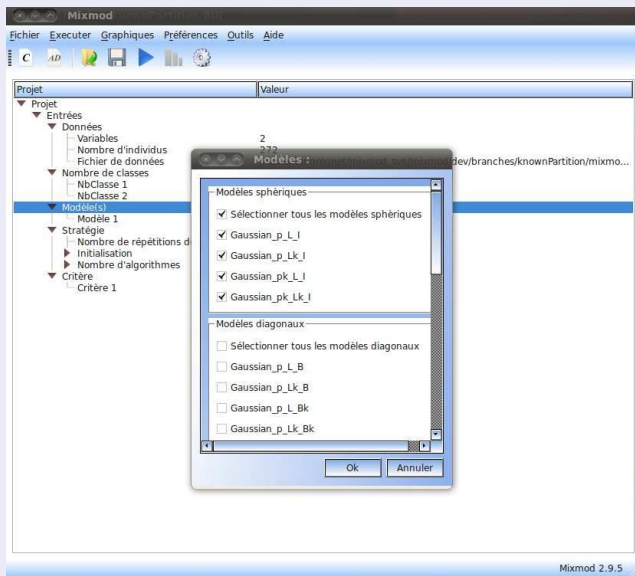
- Conviviale
- Entrées/Sorties XML
- Utilisation des bibliothèques QT et Qwt
- Disponible depuis 2011  
env. 500 téléchargements  
sur un an



## Création d'un nouveau Projet (classification)



## Choix de modèles



## Vue des inputs

The screenshot shows the Mixmod GUI interface. At the top, there is a menu bar with 'Fichier', 'Executer', 'Graphiques', 'Préférences', 'Outils', and 'Aide'. Below the menu bar is a toolbar with icons for file operations and execution. The main window is titled 'Vue des inputs' and contains a table with two columns: 'Projet' and 'Valeur'.

Projet	Valeur
▼ Projet	
▼ Entrées	
▼ Données	
Variables	2
Nombre d'individus	272
Fichier de données	/home/flangrognnet/mixmod_svn/mixmod/dev/branches/knownPartition/mixmo...
▼ Nombre de classes	
NbClasse 1	2
NbClasse 2	3
▼ Modèle(s)	
Modèle 1	Gaussian_p_L_I
Modèle 2	Gaussian_p_Lk_I
Modèle 3	Gaussian_pk_L_I
Modèle 4	Gaussian_pk_Lk_I
Modèle 5	Gaussian_pk_Lk_C
▼ Stratégie	
Nombre de répétitions de la stratégie	1
▶ Initialisation	
▶ Nombre d'algorithmes	NbAlgo
▼ Critère	
Critère 1	BIC
Critère 2	ICL

The 'Critère' row is highlighted in blue. The version number 'Mixmod 2.9.5' is visible in the bottom right corner of the window.

## Résultats numériques

Fichier Exécuter Graphiques Préférences Outils Aide

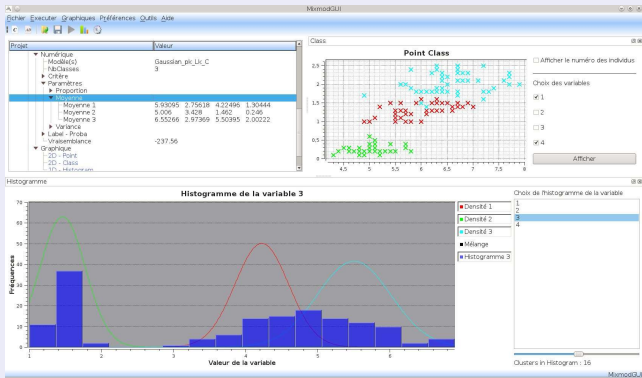
Projet Valeur

- Projet
  - Entrées
    - Données
    - Nombre de classes
    - Modèle(s)
    - Stratégie
    - Critère
  - Résultats
    - Résultat 1
      - Numérique
        - Modèle(s) Gaussian\_pk\_Lk\_C
        - NbClasses 3
        - Critère
          - Nom BIC
          - Valeur 2321.95
        - Paramètres
          - Proportion
            - Proportion 1 0.404267
            - Proportion 2 0.239898
            - Proportion 3 0.355834
          - Moyenne
            - Moyenne 1 4.51399 80.9045
            - Moyenne 2 3.9114 78.3848
            - Moyenne 3 2.0363 54.4793
          - Variance
          - Label - Proba
          - Vraisemblance -1124.54
        - Graphique
      - Résultat 2
      - Résultat 3
      - Résultat 4
      - Résultat 5
      - Résultat 6

Mixmod 2.9.5

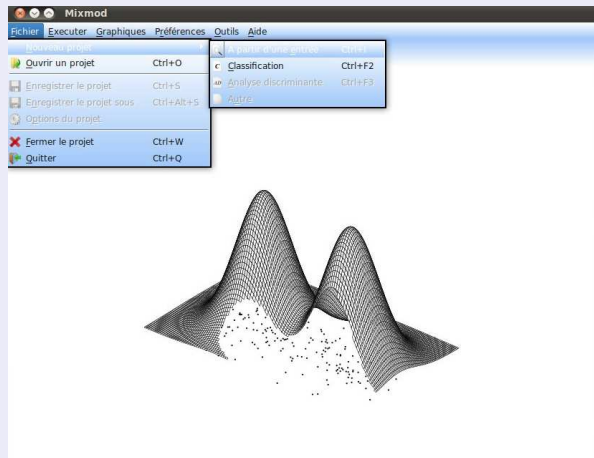


## Graphiques



## projets dans mixmodGUI

- Sauvegarde d'un projet (input/output)
- Création d'un nouveau projet à partir d'inputs enregistrées



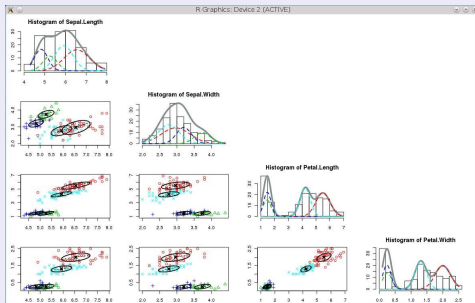
## 4<sup>e</sup> composant logiciel

RMixmod

Package pour R

## Package pour R

- Interface de mixmodLib pour R
- Ensemble de fonctions R :
  - ▶ Classification supervisée et non supervisée
  - ▶ Outils de visualisation
- Les calculs sont effectués dans mixmodLib
- Disponible depuis 2012



# MIXMOD

## 1 Pourquoi les modèles de mélanges en classification des données

- La classification des données
- Intérêts des modèles de mélanges

## 2 Mixmod

- Principales fonctionnalités
- Illustrations
- Quelques outils

## 3 Diffusion et valorisation

- Prototype ou logiciel ?
- Un long fleuve tranquille ?
- C'est en marchant qu'on apprend à marcher
- **Diffusion, valorisation : des difficultés et des actions concrètes**
- La diffusion, la valorisation se mesurent-elles ?

## 4 Perspectives

## Définition

- Ensemble des actions visant à **mettre en valeur** un produit
- La valorisation de la recherche peut être définie comme l'ensemble des activités ayant pour but d'augmenter la valeur des résultats de la recherche et, plus généralement, de mettre en valeur les connaissances. La valorisation ne se résume pas uniquement à l'exploitation commerciale des résultats de la recherche : elle s'appuie également sur le déploiement et l'échange des connaissances dans tous les domaines du savoir  
Source : [fr.wikipedia.org/wiki/Valorisation\\_de\\_la\\_recherche](http://fr.wikipedia.org/wiki/Valorisation_de_la_recherche)

**La valorisation ne se limite donc à l'aspect financier**

## Valorisation via la diffusion d'un logiciel

**La diffusion d'un logiciel sous licence GNU GPL s'inscrit donc pleinement dans les missions d'un organisme de recherche**

- Diffuser du savoir-faire
- Créer de contacts/contrats
- Permettre les contributions, les développements sur la base de ce logiciel

# Comment financer les activités de développement et diffusion ?

## Activité chronophage

- Hétérogénéité des utilisateurs
- Demandes pointues de certains utilisateurs
- Aide à l'intégration

## Financement

- Peu de besoins (en dehors des moyens humains)
- Difficulté à recruter avec les financements "sur dossier" (ANR, PEPS, ...)

## Moyens humains

- 1 IR CNRS (F. Langrognnet) (env. 50%)
- Des ingénieurs CDD INRIA (7 années ingénieurs depuis 2002)
- Des stagiaires

# Quelques actions

- Nom de domaine : [www.mixmod.org](http://www.mixmod.org)
- Dépôt de marque : **MIXMOD**®
- Rencontres Mixmod : lieu d'échange entre auteurs et utilisateurs
  - ▶ 2006 : Paris
  - ▶ 2008 : Lille
  - ▶ 2010 : Lyon
  - ▶ 2013 : Paris ?
- Etudes, actions de promotion :
  - ▶ Dossier de valorisation CNRS, d'aide au transfert
  - ▶ Etude de marché (UFC)
  - ▶ Rencontre INRIA/Industrie
  - ▶ ...





# MIXMOD

## 1 Pourquoi les modèles de mélanges en classification des données

- La classification des données
- Intérêts des modèles de mélanges

## 2 Mixmod

- Principales fonctionnalités
- Illustrations
- Quelques outils

## 3 Diffusion et valorisation

- Prototype ou logiciel ?
- Un long fleuve tranquille ?
- C'est en marchant qu'on apprend à marcher
- Diffusion, valorisation : des difficultés et des actions concrètes
- **La diffusion, la valorisation se mesurent-elles ?**

## 4 Perspectives

# Pourquoi vouloir "mesurer" ?

Des arguments pour poursuivre la diffusion

## Activité parfois chronophage

- Justifier, expliquer sa participation
- Quels sont les avantages :
  - ▶ Pour mon labo
  - ▶ Pour mon employeur
  - ▶ Pour la communauté
  - ▶ Pour moi

## Rôle d'un organisme de recherche

- Diffuser le savoir faire (y compris par le biais de logiciels)
- N'est pas un éditeur logiciel

# Outils de mesure

Comment mesurer la diffusion, la valorisation ?



## Outils

- Utilisations
- Contributions
- Enrichissements et redistributions
- Collaborations
- Intégrations



# Utilisations

## Mesures

- Nombre de **téléchargements** (mesure biaisée) - env. 150 par mois
- Activité du **forum**
- **Citations** (articles)
- **Remerciements** !!

## Typologie/Qualité

**Composants logiciels** destinés à l'une ou l'autre des catégories

- Utilisation dans des projets réels, industriels
- Utilisation (/reconnaissance) par les experts

## Comment en savoir plus ?

Sans être trop intrusif (formulaire à remplir, ...)

# Contributions

- **Très peu** de contributions (en dehors du cercle de connaissance)
- **Complexité** intrinsèque
- Certaines contributions sont cependant moins *techniques*
  - ▶ *Traduction FR/US des documentations*
  - ▶ *Participation au forum, Support*
  - ▶ *Construction des packages*
- Cession des **droits**



# Enrichissement et redistribution

- **Classification** avec prise en compte d'une **composante spatiale** (IPANEMA - Synchrotron Soleil)
- **SelVarClust** (sélection de variables en classification non supervisée et supervisée)

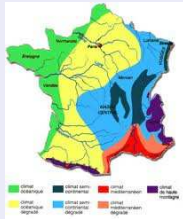


On n'est pas toujours informé !

## Utilisation avancée de MIXMOD

- **Situations complexes** ou **originales**
  - ▶ Taille des données (nombre d'individus, grande dimension)
  - ▶ Classification difficile (convergence lente et/ou non optimales)
- Nécessite parfois des (petites) **évolutions** de MIXMOD
- Contact **priviliégié** et **enrichissant** avec des utilisateurs (dans des domaines variés)

Exemple :  
Classification des climats en France



# Intégrations de mixmodLib dans un logiciel

## Motivations

- **Intérêt scientifique** de MIXMOD
- Bibliothèque **éprouvée, robuste, rapide**

## Quelques questions, contraintes

- **Licence**
  - ▶ Mixmod sous licence GNU GPL => Logiciel sous licence GNU GPL
  - ▶ Autre type de licence sur demande (**double licence**)
- **Aide** à l'intégration

## Exemple

### Plugin MIXMOD/ OpenTurns

Traitement des incertitudes (EdF, EADS, Phimeca)





# MIXMOD

## 1 Pourquoi les modèles de mélanges en classification des données

- La classification des données
- Intérêts des modèles de mélanges

## 2 Mixmod

- Principales fonctionnalités
- Illustrations
- Quelques outils

## 3 Diffusion et valorisation

- Prototype ou logiciel ?
- Un long fleuve tranquille ?
- C'est en marchant qu'on apprend à marcher
- Diffusion, valorisation : des difficultés et des actions concrètes
- La diffusion, la valorisation se mesurent-elles ?

## 4 Perspectives

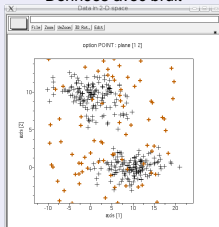
# Nouvelles fonctionnalités (statistiques)

- Traitement des **données mixtes** (quantitatives et qualitatives)

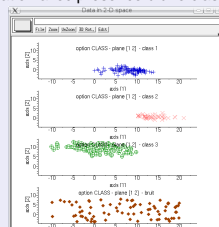
individu	hauteur	poids	sexe
1	172.5	66.3	1
2	167.1	54.1	2
⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮

- Modèles HD pour la classification non supervisée
- Traitement des données **bruitées**

Données avec bruit



Classification avec présence d'une classe de bruit



# Enrichissement de l'ensemble logiciel

## RMixmod

- Intégrer **toutes les fonctionnalités** actuelles de mixmodLib  
Toutes les méthodes d'initialisation ne sont pas disponibles
- Proposer davantage d'outils de visualisation

## mixmodGUI

- Intégrer **toutes les fonctionnalités** actuelles de mixmodLib  
Seule la classification non supervisée est disponible
- Plus de formats disponibles pour l'import et l'export des données
- Passerelles mixmodGUI / RMixmod

## mixmodForMatlab (/Scilab)

- Evolution dans l'utilisation des fonctions Mixmod dans l'environnement Matlab (pour se rapprocher de celle de RMixmod)
- mixmodForScilab n'est plus maintenu

## mixmodLib

- Activités de fonds
  - ▶ Optimisation des **performances**
  - ▶ **Robustesse** accrue (gestion améliorée des erreurs numériques, des messages d'erreur, ...)
  - ▶ Couverture de tests encore plus étendue
  - ▶ Précision numérique, fiabilité
  - ▶ **Interfaçage** facilité
  - ▶ ...
- Et aussi :
  - ▶ Calcul parallèle
  - ▶ Mesure des erreurs numériques, reproductibilité des résultats

# Diffusion, Valorisation

- Support, maintenance
- Nouvelle **journée Mixmod** (2013)
- Favoriser, développer
  - ▶ les contributions
  - ▶ les intégrations
- Rester à l'écoute des utilisateurs !



Toute contribution est la bienvenue !!

# Références (1)

Disponibles sur [www.mixmod.org](http://www.mixmod.org)

- MIXMOD Statistical Documentation (rubrique documentation)
- Exposés des journées Mixmod
- *Model-Based Cluster and Discriminant Analysis with the MIXMOD Software (CSDA - 2006)*
- *MIXMOD : un logiciel de classification supervisée et non supervisée pour données quantitatives et qualitatives (La Revue de Modulad - 2009)*
- *Rmixmod : The R Package of the Model-Based Unsupervised, Supervised and Semi-Supervised Classification Mixmod Library (Preprint)*

## Fiches PLUME

### Rubrique Logiciels validés

**PROMOUVOIR LES LOGICIELS UTILES MAÎTRISÉS ET ÉCONOMISÉS DANS L'ENSEIGNEMENT SUPÉRIEUR ET LA RECHERCHE**

Recherche

Accueil | Logiciels validés | Outils Ressources | Outils Développement ESR | PLUME 1 | Outils Ressources | Aide | Contact

Continuer à PLUME!

- Actualités PLUME
- Stationner aux fils RSS
- Logiciels validés
- Logiciels à valider
- Logiciels en test
- Ressources
- Développements ESR
- Thèmes
- Projets concrets
- Fiches archives
- Recherche
- PLUME 2
- Contribuer
- Aide
- Contact

**MIXMOD**

**MIXMOD : CLASSIFICATION DE DONNÉES**

- Site web
- Système : UNIX-like, Windows
- Téléchargement
- Version actuelle : Mixmod 2.2.1 et mixmod3D 0.9.6
- Langue(s) de l'interface : Français, anglais
- Licence : GPL

Tous les produits Mixmod sont disponibles sous licence GNU GPL mais un autre type de licence peut être proposé sur demande (contact@mixmod.org).

Disposé de développement : Heulzberg, SWEA, Sackay, Labo Maths Besançon, Paul Pantano

Une fiche Dev. Ens Sup est en relation avec cette fiche, consultez-la pour plus d'informations : [Mixmod](#)

**DESCRIPTION**

**Fonctionnalités générales**

- Ensemble logiciel de classification des données :
- classification non supervisée
- classification supervisée

Cet ensemble logiciel est constitué :

- Notes-dés

**\* Fiche logiciel validé**

Logiciel validé par le comité national ESR

Statut de la fiche : Logiciel / GNU GPL

Création du Mixmod : 2002/02

Correction initiale : 17/12/02

Réaucteur de la fiche : Florent Langrognet (contact@mixmod.org), Laboratoire de Mathématiques de Besançon (UMR), Université de Franche-Comté

L'auteur accepte de faire un support léger à l'installation.

Réaucteur(s) : Violaine Louvet (Droit@univ-besancon.fr)

Annuel Scherer (MATHA.Pantano@univ-besancon.fr)

Responsable technique : Teresa Gomez-Oas (L2DM)

**Notes-dés**

Système : UNIX-like, Windows

Motiv-actuel : math

Domaine informatique : développement, reseau

OS : linux, informations perso, service informatique

Fonctionnalités principales : table, informatique, statistiques, traitement de données

Motiv-dés divers : Matlab

Motiv-dés proposée par l'auteur : classification

### Rubrique Dev Ens Sup - Recherche

**PROMOUVOIR LES LOGICIELS UTILES MAÎTRISÉS ET ÉCONOMISÉS DANS L'ENSEIGNEMENT SUPÉRIEUR ET LA RECHERCHE**

Recherche

Accueil | Logiciels validés | Outils Ressources | Outils Développement ESR | PLUME 1 | Outils Ressources | Aide | Contact

Continuer à PLUME!

- Actualités PLUME
- Stationner aux fils RSS
- Logiciels validés
- Logiciels à valider
- Logiciels en test
- Ressources
- Développements ESR
- Thèmes
- Projets concrets
- Fiches archives
- Recherche
- PLUME 2
- Contribuer
- Aide
- Contact

**MIXMOD**

**MIXMOD : ENSEMBLE LOGICIEL DE CLASSIFICATION SUPERVISÉE ET NON SUPERVISÉE DES DONNÉES**

Ce logiciel a été développé (ou est en cours de développement) dans le cadre de la recherche en enseignement supérieur et de la recherche. Son état peut être variable (cf. champs ci-dessous) dans sans garantie de bon fonctionnement.

- Site web
- Système : UNIX-like, Windows
- Version actuelle : mixmod3 2.2.1 - mixmod3D 0.9.6 - mixmod3D Matlab 2.2.1 - 2011
- Licence(s) : GPL, Licence propriétaire - Licence GPL par défaut, Licence Propriétaire sur demande
- État : validé (si sans PLUME), BETA, stable
- Support : mail,net, téléphone, développement en cours
- Contact(s) : Florent Langrognet
- Contact conceptuel(s) : mixmod@univ-besancon.fr
- Labo(s) de dev(s) : mixmod@univ-besancon.fr
- Labo(s) de dev(s) : Heulzberg, SWEA, Sackay, Labo Maths Besançon, Paul Pantano

**CONTRIBUTEURS**

- RSE du contributeur
- Charte du contributeur
- Motivations
- Motivations
- Desaccords

Une fiche de logiciel validé décrit plus en détail ce logiciel, consultez-la pour plus d'informations : [Mixmod](#)

**\* Fiche de Dev. Sup - Recherche**

Développement de la communauté de l'Enseignement Supérieur et de la Recherche

Création ou MAJ importante : 20/12/09

Correction initiale : 18/02/12

Auteur de la fiche : Florent Langrognet

Laboratoire de Mathématiques de Besançon

Personnalité(s) technique(s) : Teresa Gomez-Oas (L2DM)

**Motiv-dés**

Laboratoire(s) : Heulzberg, SWEA, Sackay, Labo Maths Besançon, Paul Pantano

Motiv-dés : CPM, SWEA, UPC, Univ Lille 1, UTC

OS : linux, informations perso

Motiv-actuel : math

Domaine informatique : reseau

OS : linux - informatique perso

Fonctionnalités principales : statistiques

Motiv-dés divers : Matlab

Motiv-dés proposée par l'auteur : méthodes de classification

MIXMOD<sup>®</sup>

Un ensemble logiciel de classification des données  
par modèles de mélange

F. Langrognet

**Merci de votre attention**