

# Diagnostic par suivi de point de fonctionnement

Faïcel Chamroukhi



Institut National de Recherche sur les Transports et leur Sécurité (INRETS)  
Laboratoire des Technologies Nouvelles (LTN)

Université de Technologie de Compiègne (UTC)  
Laboratoire HEUDIASYC, UMR CNRS 6599

Directeurs de thèse : Gérard Govaert (UTC)  
Patrice Aknin (INRETS)

Encadrement : Allou Samé (INRETS)

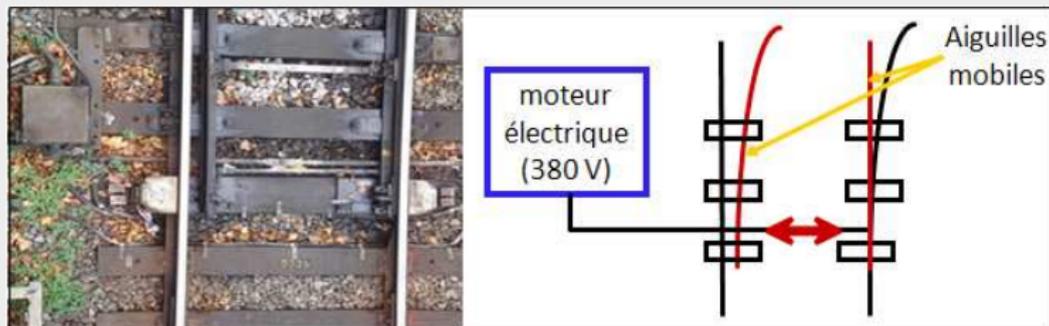
9 Juillet 2009

- 1 Introduction
- 2 Aspects bibliographiques
  - Modèle de régression par morceaux
  - Modèle de régression régi par une chaîne de Markov caché (HMM)
- 3 Modèle de régression à processus latent proposé
- 4 Expérimentations
- 5 Apprentissage à partir de plusieurs signaux
  - Expérimentations
- 6 Conclusion et perspectives

# Contexte applicatif

- ▶ Collaboration avec la SNCF
- ▶ Diagnostic par suivi de point de fonctionnement du mécanisme d'aiguillage
- ▶ Application des méthodes développées au circuit de voie

## Mécanisme d'aiguillage



# Objectifs visés

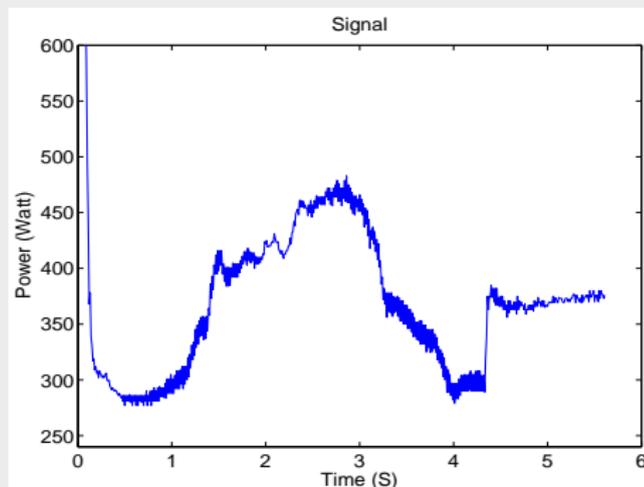
- ▶ Suivi du point de fonctionnement :  
Estimer la dynamique d'évolution de l'état de fonctionnement du système
- ▶ Détecter précisément le type défaut (défaut mécanique, défaut électrique, défaut de graissage,...) afin de déployer l'équipe concernée

# Approche proposée

- ① 1ère étape : Caractérisation de chaque signal dans un espace de représentation spécifique
  - Estimation des paramètres d'un modèle probabiliste de signal
  - Avantages :
    - Apte à prendre en compte les différentes phases des signaux étudiés
    - Flexibilité
    - Possibilité de simuler facilement
- ② 2ème étape : Développement d'une méthode de suivi temporel s'appuyant sur les paramètres estimés (2ème partie de la thèse)

# Données exploitées

- ▶ Signaux de puissance consommée par le moteur d'aiguille durant chaque manœuvre



- ▶ Particularité des signaux : plusieurs régimes se succédant de manière brusque ou lente
- ▶ Solution proposée : utiliser un modèle de régression adapté, dont les paramètres vont servir de caractéristiques des signaux.

# Modèle de régression par morceaux [McGee & Carleton 70]

- ▶ Échantillon  $((x_1, t_1), \dots, (x_n, t_n))$ 
  - $x_i$  : Variable dépendante réelle ;
  - $t_i$  : Variable indépendante représentant le temps
- ▶ Le modèle de régression par morceaux générant le signal  $\mathbf{x}$  est :

$$\forall i = 1, \dots, n, \quad x_i = \beta_k^T \mathbf{r}_i + \sigma_k \epsilon_i \quad ; \quad \epsilon_i \sim \mathcal{N}(0, 1)$$

- $k$  vérifiant  $i \in I_k = (\gamma_k, \gamma_{k+1}]$  : indices des éléments du segment  $k$
- $\mathbf{r}_i = (1, t_i, \dots, t_i^p)^T$  : vecteur-covariable dans  $\mathbf{R}^{p+1}$
- $\beta_k$  : vecteur de coefficients de régression  $\in \mathbf{R}^{(p+1)}$  pour le  $k^{ieme}$  segment

## Paramètres du modèle

$(\psi, \gamma)$  avec  $\psi = (\beta_1, \dots, \beta_K, \sigma_1^2, \dots, \sigma_K^2)$  et  $\gamma = (\gamma_1, \dots, \gamma_{K+1})$ .

# Estimation des paramètres du modèle de régression par morceaux

- ▶ Maximiser la vraisemblance de  $(\psi, \gamma)$  ou, par équivalence, minimiser par rapport à  $(\psi, \gamma)$  :

$$J(\psi, \gamma) = \sum_{k=1}^K \sum_{i \in I_k} \left[ \log \sigma_k^2 + \frac{(x_i - \beta_k^T \mathbf{r}_i)^2}{\sigma_k^2} \right]$$

- ▶ Optimization globale par l'algorithme de Fisher [Fisher 58] qui est basé sur la programmation dynamique [Bellman 61 ; Lechevallier 90] puisque  $J$  est additif sur  $k$

# Estimation des paramètres du modèle de régression par morceaux

- ▶ Maximiser la vraisemblance de  $(\psi, \gamma)$  ou, par équivalence, minimiser par rapport à  $(\psi, \gamma)$  :

$$J(\psi, \gamma) = \sum_{k=1}^K \sum_{i \in I_k} \left[ \log \sigma_k^2 + \frac{(x_i - \beta_k^T \mathbf{r}_i)^2}{\sigma_k^2} \right]$$

- ▶ Optimization globale par l'algorithme de Fisher [Fisher 58] qui est basé sur la programmation dynamique [Bellman 61 ; Lechevallier 90] puisque  $J$  est additif sur  $k$

## Approximation et segmentation d'un signal

- ▶  $\hat{x}_i = \sum_{k=1}^K \hat{z}_{ik} \hat{\beta}_k^T \mathbf{r}_i$  ;  $\forall i = 1, \dots, n$
- ▶  $\hat{z}_{ik} = 1$  si  $i \in (\hat{\gamma}_k, \hat{\gamma}_{k+1}]$  ( $x_i$  appartient au  $k$ ème segment) et  $\hat{z}_{ik} = 0$  sinon

- ▶ La programmation dynamique peut s'avérer coûteuse en temps de calcul
- ▶ Fournit une segmentation dure  $\Rightarrow$  adapté pour les changements brusques

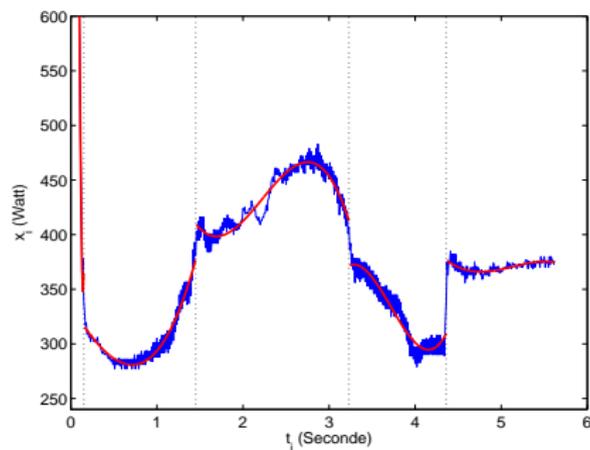


FIG.: Exemple de résultat graphique obtenu avec le modèle de régression par morceaux

# Modèle de régression régi par un HMM [Fridman 93]

Le modèle :

$$\forall i = 1, \dots, n, \quad x_i = \beta_{z_i}^T \mathbf{r}_i + \sigma_{z_i} \varepsilon_i \quad ; \quad \varepsilon_i \sim \mathcal{N}(0, 1),$$

- ▶  $z_i$  une variable latente discrète qui prend ses valeurs dans  $\{1, \dots, K\}$  représentant le label du modèle de régression de  $x_i$
- ▶ La séquence cachée  $\mathbf{z} = (z_1, \dots, z_n)$  est une chaîne de markov de matrice de transitions  $A$  et de loi initiale  $\pi$

# Contraintes sur la matrice de transition

- ▶ Les signaux d'aiguillages sont dus aux mouvements successifs liés à la mécanique du système
- ▶ On suppose une contrainte d'ordre sur les états du HMM

$$p(z_i = k | z_{i-1} = \ell) = 0 \quad \text{if } k < \ell ,$$

et

$$p(z_i = k | z_{i-1} = \ell) = 0 \quad \text{if } k > \ell + 1.$$

⇒ Pas de saut d'état et pas de retour à un état précédent

- ▶ Cas particulier du modèle gauche-droit [Rabiner 89]

# Estimation des paramètres du modèle

▶  $\Psi = (\pi, A, \beta_1, \dots, \beta_K, \sigma_1^2, \dots, \sigma_K^2)$  vecteur paramètre du modèle

▶ Méthode du maximum de vraisemblance

▶ log-vraisemblance de  $\Psi$  :

$$\begin{aligned} L(\Psi; \mathbf{x}) &= \log p(\mathbf{x}; \Psi) \\ &= \log \sum_{\mathbf{z}} p(\mathbf{z}_1; \pi) \prod_{i=2}^n p(z_i | z_{i-1}; A) \prod_{i=1}^n \mathcal{N}(x_i; \beta_{z_i}^T, \sigma_{z_i}^2). \end{aligned}$$

▶ Maximisation de la log-vraisemblance par l'algorithme EM (Baum-Welch)

# Approximation et segmentation de signaux avec le modèle markovien

## Approximation d'un signal

- Somme de polynômes pondérés par les probabilités de filtrage  $\omega_{ik}$  :

$$\hat{x}_i = \sum_{k=1}^K \omega_{ik}(\hat{\Psi}) \hat{\beta}_k^T \mathbf{r}_i ; i = 1, \dots, n,$$

où  $\hat{\Psi} = (\hat{\pi}, \hat{A}, \hat{\beta}_1, \dots, \hat{\beta}_K, \hat{\sigma}_1^2, \dots, \hat{\sigma}_K^2)$  et  $\omega_{ik}(\hat{\Psi}) = p(z_i = k | x_1, \dots, x_i; \hat{\Psi})$

## Segmentation d'un signal

Calculer le label  $\hat{z}_i$  du composant générant  $x_i$  par la règle du MAP :

$$\hat{z}_i = \arg \max_{1 \leq k \leq K} \hat{\tau}_{ik} ; \forall i = 1, \dots, n, \quad (1)$$

où  $\tau_{ik} = p(z_i = k | \mathbf{x}; \Psi)$  est la probabilité a posteriori que  $x_i$  soit issu de la  $k$ ème composante régressive.

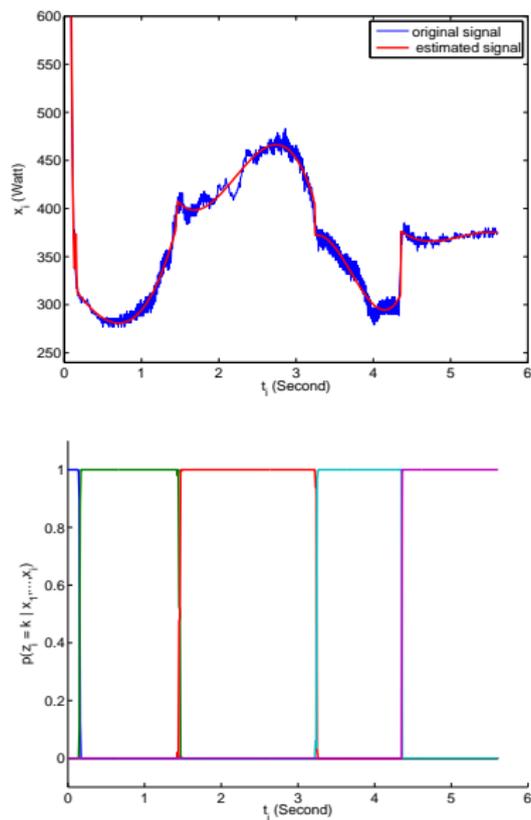


FIG.: Exemple de résultat graphique obtenu avec le modèle markovien

# Modèle de régression à processus latent proposé

## Le modèle global

$$\forall i = 1, \dots, n, \quad x_i = \beta_{z_i}^T \mathbf{r}_i + \sigma_{z_i} \epsilon_i \quad ; \quad \epsilon_i \sim \mathcal{N}(0, 1),$$

$\mathbf{z} = (z_1, \dots, z_n)$  est un processus logistique caché

# Modèle de régression à processus latent proposé

## Le modèle global

$$\forall i = 1, \dots, n, \quad x_i = \beta_{z_i}^T \mathbf{r}_i + \sigma_{z_i} \epsilon_i \quad ; \quad \epsilon_i \sim \mathcal{N}(0, 1),$$

$\mathbf{z} = (z_1, \dots, z_n)$  est un processus logistique caché

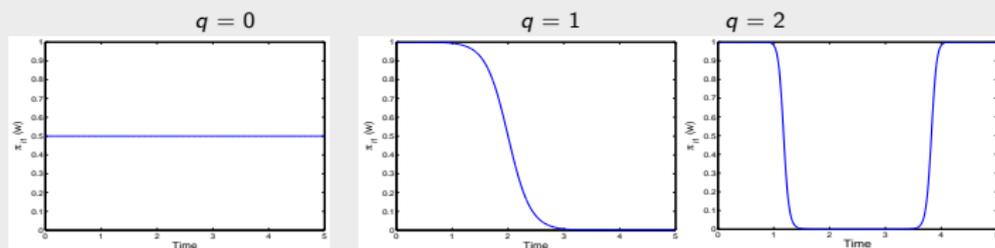
$z_i \sim \mathcal{M}(1, \pi_{i1}(\mathbf{w}), \dots, \pi_{iK}(\mathbf{w}))$ ; où

$$\pi_{ik}(\mathbf{w}) = p(z_i = k; \mathbf{w}) = \frac{\exp(\mathbf{w}_k^T \mathbf{v}_i)}{\sum_{\ell=1}^K \exp(\mathbf{w}_\ell^T \mathbf{v}_i)},$$

- ▶  $\mathbf{v}_i = (1, t_i, \dots, t_i^q)^T \in \mathbf{R}^{(q+1)}$
- ▶  $\mathbf{w}_k = (\mathbf{w}_{k0}, \dots, \mathbf{w}_{kq})^T$  vecteur paramètre de la  $k$ ème fonction logistique  $\in \mathbf{R}^{q+1}$
- ▶  $\mathbf{w} = (\mathbf{w}_1, \dots, \mathbf{w}_K)$  vecteur paramètre pour les  $K$  fonctions logistiques  $\in \mathbf{R}^{K \times (q+1)}$

# Flexibilité de la transformation logistique : Exemple pour $K = 2$ :

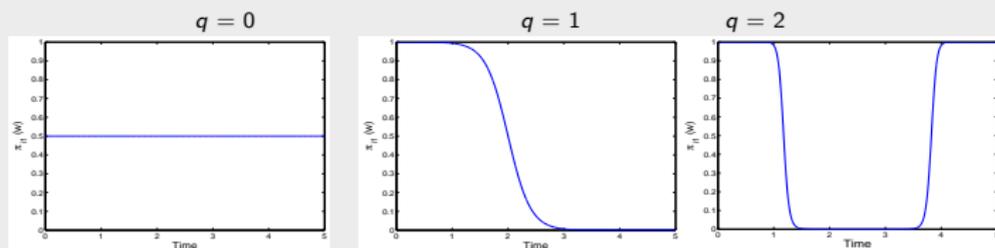
- ①  $\pi_{ik}(\mathbf{w})$  en fonction de la dimension  $q$  de  $\mathbf{w}_k$  :



$\Rightarrow q = 1$  garantit une segmentation en régions contiguës

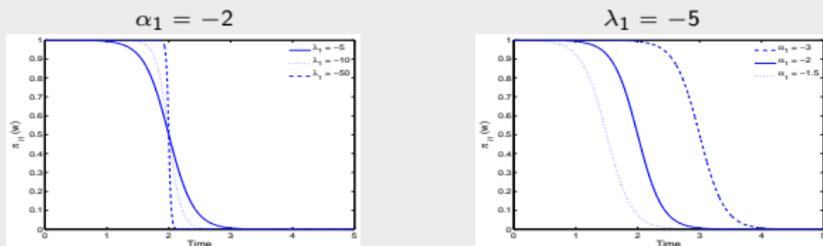
# Flexibilité de la transformation logistique : Exemple pour $K = 2$ :

- ❶  $\pi_{ik}(\mathbf{w})$  en fonction de la dimension  $q$  de  $\mathbf{w}_k$  :



$\Rightarrow q = 1$  garantit une segmentation en régions contiguës

- ❷  $\pi_{ik}(\mathbf{w})$  en fonction de  $\mathbf{w}_k$  pour  $q = 1$ . Paramétrisation de  $\mathbf{w}_k$  par  $\mathbf{w}_k = \lambda_k(\alpha_k, 1)^T$



- $\Rightarrow$  Le paramètre  $\lambda_k$  contrôle la qualité des transitions entre les régimes  
 $\Rightarrow$  Le paramètre  $\alpha_k$  contrôle l'instant de transition

# Estimation des paramètres du modèle par maximum de vraisemblance

- Densité mélange

$$p(x_i; \theta) = \sum_{k=1}^K \pi_{ik}(\mathbf{w}) \mathcal{N}(x_i; \beta_k^T \mathbf{r}_i, \sigma_k^2)$$

- Vecteur paramètre du modèle :

$$\theta = (\mathbf{w}, \beta_1, \dots, \beta_K, \sigma_1^2, \dots, \sigma_K^2)$$

- Log-vraisemblance de  $\theta$  :

$$L(\theta; \mathbf{x}) = \sum_{i=1}^n \log \sum_{k=1}^K \pi_{ik}(\mathbf{w}) \mathcal{N}(x_i; \beta_k^T \mathbf{r}_i, \sigma_k^2).$$

- Maximisation de  $L(\theta; \mathbf{x})$  par un algorithme Expectation-Maximization (EM) dédié [Dempster et al. 77].

# Algorithme EM dédié

**Initialisation :**  $\theta^{(0)}$

Itérer jusqu'à la convergence les deux étapes suivantes :

# Algorithme EM dédié

**Initialisation** :  $\theta^{(0)}$

Itérer jusqu'à la convergence les deux étapes suivantes :

**1 Étape E : Espérance** ( à l'itération  $q$  )

Calcul de l'espérance conditionnelle de la log-vraisemblance de  $\theta$  pour les données complétées  $L(\theta; \mathbf{x}, \mathbf{z})$

$$\begin{aligned}
 Q(\theta, \theta^{(q)}) &= E \left[ L(\theta; \mathbf{x}, \mathbf{z}) | \mathbf{x}, \theta^{(q)} \right] \\
 &= \underbrace{\sum_{i=1}^n \sum_{k=1}^K \tau_{ik}^{(q)} \log \pi_{ik}(\mathbf{w})}_{Q_1(\mathbf{w})} + \underbrace{\sum_{i=1}^n \sum_{k=1}^K \tau_{ik}^{(q)} \log \mathcal{N}(x_i; \beta_k^T \mathbf{r}_i, \sigma_k^2)}_{Q_2(\beta_k, \sigma_k^2; k=1, \dots, K)},
 \end{aligned}$$

$$\text{avec } \tau_{ik}^{(q)} = p(z_i = k | x_i; \theta^{(q)}) = \frac{\pi_{ik}(\mathbf{w}^{(q)}) \mathcal{N}(x_i; \beta_k^T \mathbf{r}_i, \sigma_k^2)}{\sum_{\ell=1}^K \pi_{i\ell}(\mathbf{w}^{(q)}) \mathcal{N}(x_i; \beta_\ell^T \mathbf{r}_i, \sigma_\ell^2)}$$

# Algorithme EM dédié

**Initialisation** :  $\theta^{(0)}$

Itérer jusqu'à la convergence les deux étapes suivantes :

**1 Étape E : Espérance** ( à l'itération  $q$  )

Calcul de l'espérance conditionnelle de la log-vraisemblance de  $\theta$  pour les données complétées  $L(\theta; \mathbf{x}, \mathbf{z})$

$$\begin{aligned}
 Q(\theta, \theta^{(q)}) &= E \left[ L(\theta; \mathbf{x}, \mathbf{z}) \mid \mathbf{x}, \theta^{(q)} \right] \\
 &= \underbrace{\sum_{i=1}^n \sum_{k=1}^K \tau_{ik}^{(q)} \log \pi_{ik}(\mathbf{w})}_{Q_1(\mathbf{w})} + \underbrace{\sum_{i=1}^n \sum_{k=1}^K \tau_{ik}^{(q)} \log \mathcal{N}(x_i; \beta_k^T \mathbf{r}_i, \sigma_k^2)}_{Q_2(\beta_k, \sigma_k^2; k=1, \dots, K)},
 \end{aligned}$$

avec  $\tau_{ik}^{(q)} = p(z_i = k \mid x_i; \theta^{(q)}) = \frac{\pi_{ik}(\mathbf{w}^{(q)}) \mathcal{N}(x_i; \beta_k^T \mathbf{r}_i, \sigma_k^2)}{\sum_{\ell=1}^K \pi_{i\ell}(\mathbf{w}^{(q)}) \mathcal{N}(x_i; \beta_\ell^T \mathbf{r}_i, \sigma_\ell^2)}$

**2 Étape M : Maximisation** (à l'itération  $q$ )

Calcul de  $\theta^{(q+1)} = \arg \max_{\theta} Q(\theta, \theta^{(q)})$

## Détails de l'étape M

Maximisation séparée de  $Q_1(\mathbf{w})$  et  $Q_2(\beta_k, \sigma_k^2; k = 1, \dots, K)$  :

- 1 Maximisation de  $Q_2$  par rapport aux  $\beta_k$  : résolution analytique de  $K$  problèmes de régression polynomiale pondérés

- $\beta_k^{(q+1)} = (\mathbf{M}^T \Gamma_k^{(q)} \mathbf{M})^{-1} \mathbf{M}^T \Gamma_k^{(q)} \mathbf{x}$

avec  $\mathbf{M}$  la matrice de régression et  $\Gamma_k^{(q)} = \text{diag}(\tau_{1k}^{(q)}, \dots, \tau_{nk}^{(q)})$

Maximisation de  $Q_2$  par rapport aux  $\sigma_k^2$  :

- $\sigma_k^{2(q+1)} = \frac{1}{\sum_{i=1}^n \tau_{ik}^{(q)}} \sum_{i=1}^n \tau_{ik}^{(q)} (x_i - \beta_k^{T(q+1)} \mathbf{r}_i)^2$

## Détails de l'étape M

Maximisation séparée de  $Q_1(\mathbf{w})$  et  $Q_2(\beta_k, \sigma_k^2; k = 1, \dots, K)$  :

- 1 Maximisation de  $Q_2$  par rapport aux  $\beta_k$  : résolution analytique de  $K$  problèmes de régression polynomiale pondérés

- $\beta_k^{(q+1)} = (\mathbf{M}^T \Gamma_k^{(q)} \mathbf{M})^{-1} \mathbf{M}^T \Gamma_k^{(q)} \mathbf{x}$

avec  $\mathbf{M}$  la matrice de régression et  $\Gamma_k^{(q)} = \text{diag}(\tau_{1k}^{(q)}, \dots, \tau_{nk}^{(q)})$

Maximisation de  $Q_2$  par rapport aux  $\sigma_k^2$  :

- $\sigma_k^{2(q+1)} = \frac{1}{\sum_{i=1}^n \tau_{ik}^{(q)}} \sum_{i=1}^n \tau_{ik}^{(q)} (x_i - \beta_k^{T(q+1)} \mathbf{r}_i)^2$

- 2 Maximisation de  $Q_1$  par rapport  $\mathbf{w}$  : problème convexe de régression logistique multinomiale pondéré

⇒ Algorithme IRLS (Iterative Reweighted Least Squares) [Green 84]

$$\mathbf{w}^{(m+1)} = \mathbf{w}^{(m)} - \left[ \frac{\partial^2 Q_1(\mathbf{w})}{\partial \mathbf{w} \partial \mathbf{w}^T} \right]_{\mathbf{w}=\mathbf{w}^{(m)}}^{-1} \frac{\partial Q_1(\mathbf{w})}{\partial \mathbf{w}} \Big|_{\mathbf{w}=\mathbf{w}^{(m)}} ,$$

⇒ L'algorithme IRLS fournit le paramètre  $\mathbf{w}^{(q+1)}$  de l'itération courante.

# Approximation et segmentation de signaux avec le modèle proposé

## Approximation d'un signal

- ▶ Comme dans la régression classique,  $x_i$ ,  $i = 1, \dots, n$  est approximé par son espérance :

$$\hat{x}_i = E(x_i; \hat{\theta}) = \int_R x_i p(x_i; \hat{\theta}) dx_i = \sum_{k=1}^K \pi_{ik}(\hat{\mathbf{w}}) \hat{\beta}_k^T \mathbf{r}_i .$$

Une somme de polynômes pondérés par des fonctions logistiques  $\pi_{ik}(\hat{\mathbf{w}})$ .

⇒ Adapté pour les changements brusques ou souples entre les différents composants

⇒ Assure la continuité du signal estimé

## Segmentation d'un signal

- ▶ Estimer le label  $\hat{z}_i$  du composant générant  $x_i$  par la règle suivante :

$$\hat{z}_i = \arg \max_{1 \leq k \leq K} \pi_{ik}(\hat{\mathbf{w}}) .$$

# Expérimentations sur des données simulées

## Critères d'évaluation

- ▶ Écart quadratique moyen entre la courbe de régression estimée et la courbe de régression simulée (erreur de débruitage) :

$$\frac{1}{n} \sum_{i=1}^n [E(x_i; \theta) - \hat{x}_i]^2$$

$\theta$  étant le vrai paramètre.

- ▶ Erreur de segmentation
- ▶ Comparaison de l'algorithme proposé avec
  - la méthode de régression par morceaux
  - le modèle de régression avec HMM

# Protocole de simulation

Trois expériences effectuées à partir de signaux simulés selon le modèle de régression à processus logistique caché

- 1 Expérience 1 : effet du niveau de souplesse des transitions sur la qualité d'estimation.

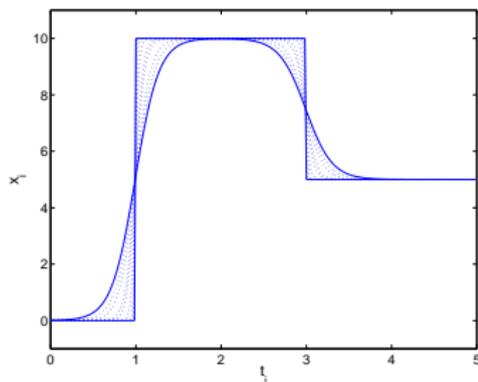
La qualité des transitions entre les différents régimes est contrôlée par les paramètres  $\lambda_k$  de la fonction logistique

- 2 Expérience 2 : effet de la taille d'échantillon  $n$  sur la qualité des estimations
- 3 Expérience 3 : effet du niveau de bruit  $\sigma$  sur la qualité des estimations.

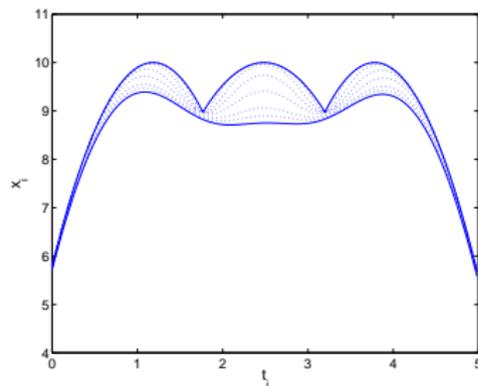
20 jeux de données différents ont été simulés pour chaque situation et les résultats ont été moyennés

Situation 1	$\beta_1 = 0$ $\beta_2 = 10$ $\beta_3 = 5$	$w_1 = [3341.33, -1706.96]$ $w_2 = [2436.97, -810.07]$ $w_3 = [0, 0]$
Situation 2	$\beta_1 = [-0.64, 14.4, -6]$ $\beta_2 = [-21.25, 25, -5]$ $\beta_3 = [-78.64, 45.6, -6]$	$w_1 = [3767.58, -1510.19]$ $w_2 = [2468.99, -742.55]$ $w_3 = [0, 0]$

Niveau de souplesse des transitions	1	2	3	4	5	6	7	8	9	10
(a) $ \lambda_k $ divisé par :	1	2	5	10	20	40	50	80	100	125
(b) $ \lambda_k $ divisé par :	1	10	50	100	150	200	250	275	300	400

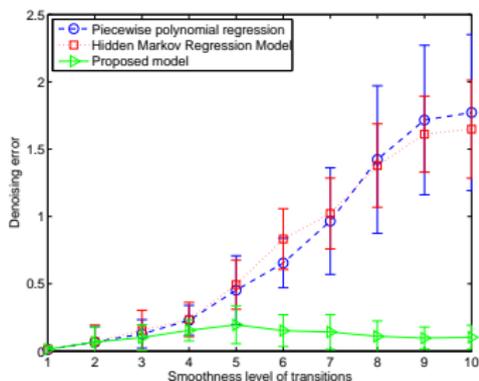


(a)

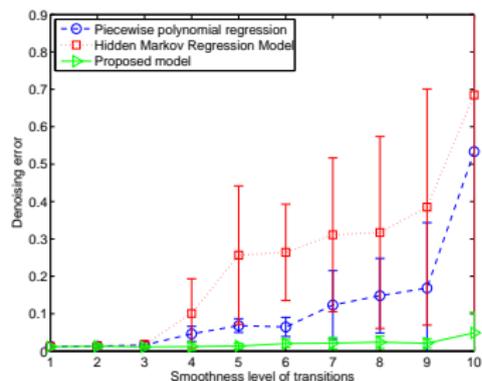


(b)

Résultats :  $EQM$  entre la courbe de régression estimée et la courbe de régression simulée en fonction de la qualité des transitions, obtenue pour la situation (a) et la situation (b) :

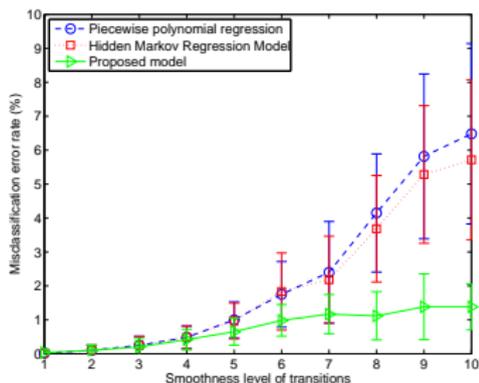


(a)

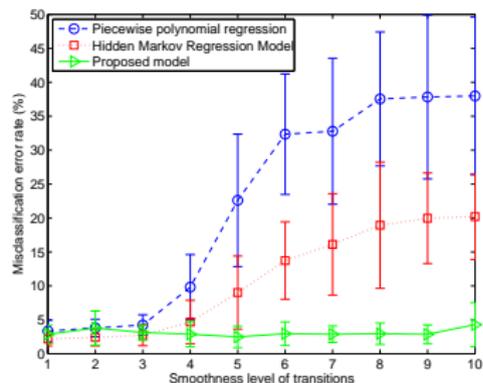


(b)

Résultats : erreur de segmentation en fonction de la qualité des transitions, obtenue pour la situation (a) et la situation (b) :



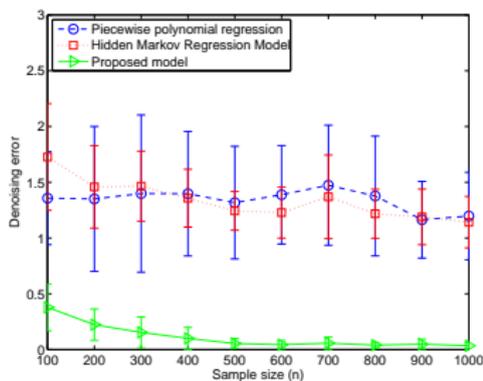
(a)



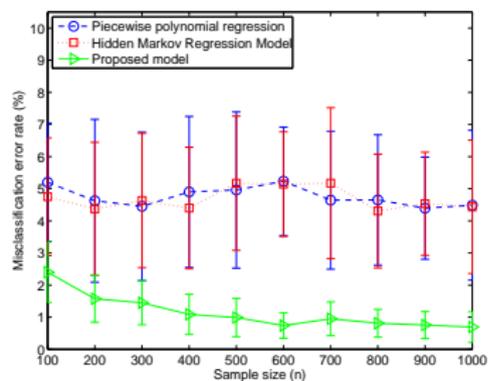
(b)

# Résultats

## Erreur de débruitage

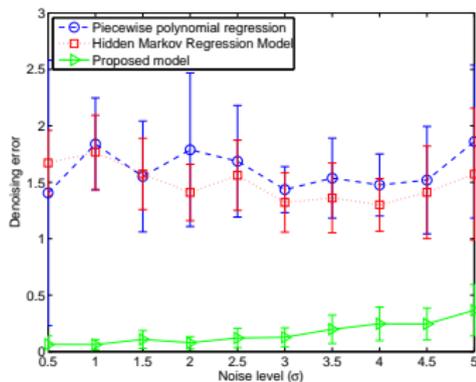


## Taux d'erreur de classification

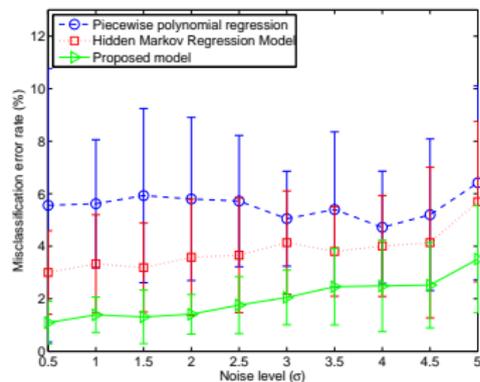


# Résultats

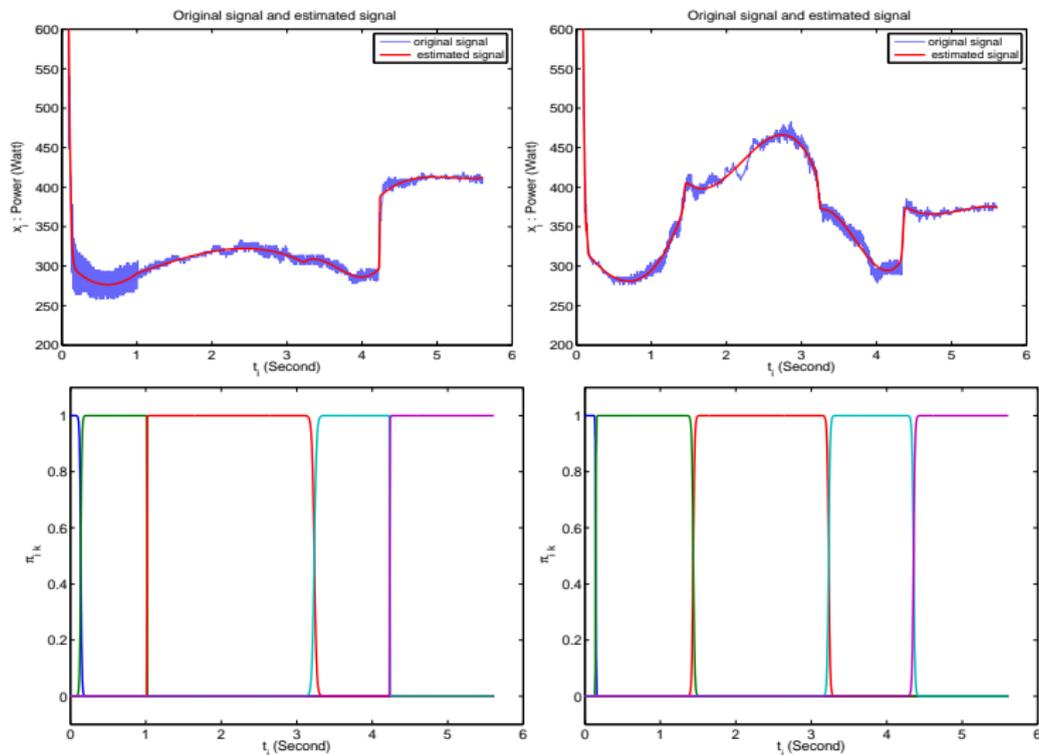
## Erreur de débruitage



## Taux d'erreur de classification



# Exemple de signaux de manœuvres d'aiguillage : Modèles estimés et fonctions logistiques associées



# Classement des signaux de manœuvres d'aiguillage

- ▶  $N = 119$  signaux de manœuvres étiquetés
- ▶ Paramétriser tous les signaux par les 3 approches
- ▶ 3 classes : sans défaut, avec défaut mineur, avec défaut critique
- ▶ Partage aléatoire en base d'apprentissage (84 signaux) et base de test (35 signaux)
- ▶ Apprentissage des paramètres des classes par MDA (Mixture Discriminant Analysis) [Hastie & Tibshirani 96] sur les caractéristiques des signaux
- ▶ classement des signaux par Maximum A Posteriori (MAP)

# Classement des signaux de manœuvres d'aiguillage

- ▶  $N = 119$  signaux de manœuvres étiquetés
- ▶ Paramétriser tous les signaux par les 3 approches
- ▶ 3 classes : sans défaut, avec défaut mineur, avec défaut critique
- ▶ Partage aléatoire en base d'apprentissage (84 signaux) et base de test (35 signaux)
- ▶ Apprentissage des paramètres des classes par MDA (Mixture Discriminant Analysis) [Hastie & Tibshirani 96] sur les caractéristiques des signaux
- ▶ classement des signaux par Maximum A Posteriori (MAP)

Taux de bonne classification (%)

Modèle régression par morceaux	83
Modèle de régression par HMM	89
Modèle de régression proposé	91

Nombre de composant pour chaque  
classe sélectionné par BIC

Classe	1	2	3
Nombre de composants	1	1	3

# Extension : Apprentissage des paramètres de classes à partir de plusieurs signaux

## Modèle dans le cas de plusieurs signaux

- ▶ Échantillon :  $n$  signaux  $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$  issus de la même classe
- ▶  $\mathbf{x}_i = ((x_{i1}, t_1), \dots, (x_{ij}, t_j), \dots, (x_{im}, t_m))$ ,  $i = 1, \dots, n$
- ▶ Modèle générant  $\mathbf{x}_i$  :

$$\forall j = 1, \dots, m, \quad x_{ij} = \beta_{z_j}^T \mathbf{r}_j + \sigma_{z_j} \epsilon_{ij} \quad ; \quad \epsilon_{ij} \sim \mathcal{N}(0, 1),$$

avec  $z_j \in \{1, \dots, K\}$  est une variable latente discrète et  $\mathbf{z} = (z_1, \dots, z_m)$  est un processus logistique.

# Évaluation en terme de modélisation de signaux

## Protocole de simulation

Trois expérimentations :

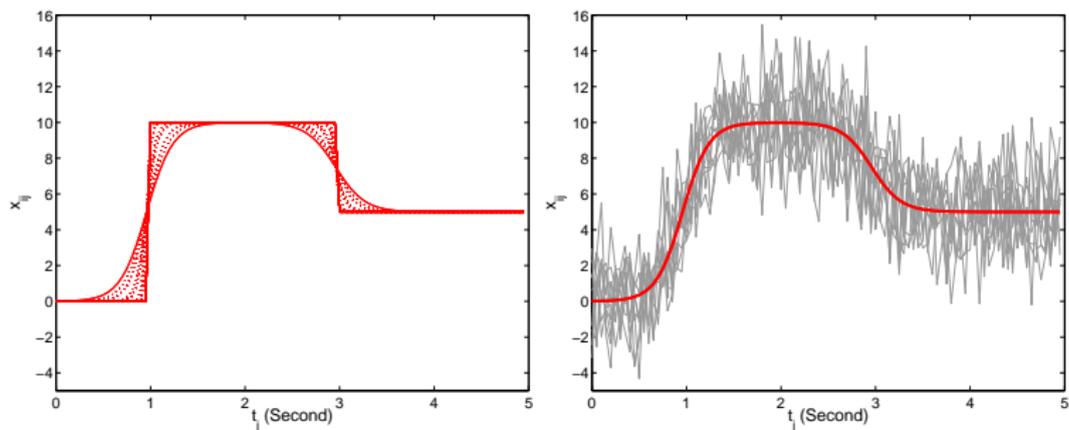
- 1 Expérience 1 : effet de la qualité des transitions sur les estimations.
- 2 Expérience 2 : effet du nombre de courbes  $n$  par échantillon sur la qualité des estimations
- 3 Expérience 3 : effet du nombre d'observations  $m$  par courbe

20 échantillons différents pour chaque situation ont été simulés et les résultats ont été moyennés

$$\beta_1 = 0 \quad \mathbf{w}_1 = [3341.33, -1706.96]$$

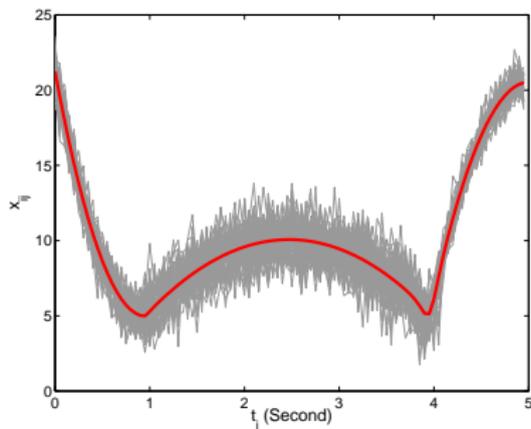
$$\beta_2 = 10 \quad \mathbf{w}_2 = [2436.97, -810.07]$$

$$\beta_3 = 5 \quad \mathbf{w}_3 = [0, 0]$$



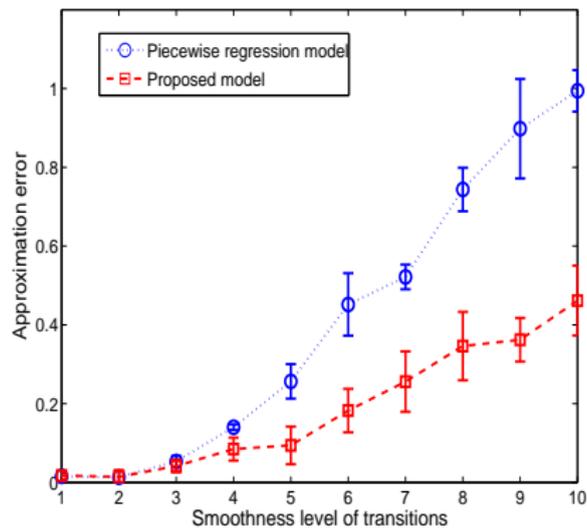
(a)

$\beta_1 = [23, -36, 18]$	$\mathbf{w}_1 = [92.72, -46.72]$	$\sigma_1 = 1$
$\beta_2 = [-3.9, 11.08, -2.2]$	$\mathbf{w}_2 = [61.16, -15.28]$	$\sigma_2 = 1.25$
$\beta_3 = [-337, 141.5, -14]$	$\mathbf{w}_3 = [0, 0]$	$\sigma_3 = 0.75$

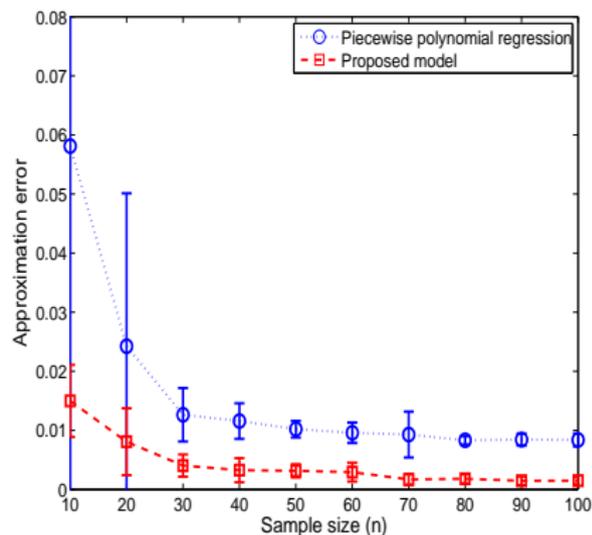
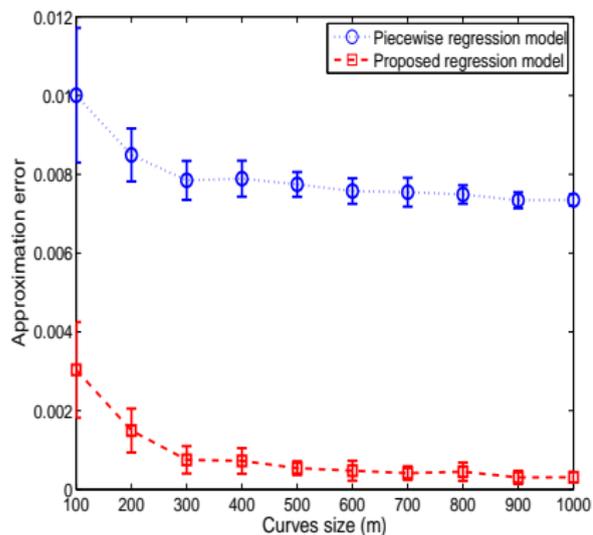


(b)

# Résultats : erreur d'approximation des courbes en fonction de la qualité des transitions (situation (a))



Résultats : erreur d'approximation des courbes en fonction la taille d'échantillon  $n$  et du nombre de points par courbe  $m$  (situation (b)) :

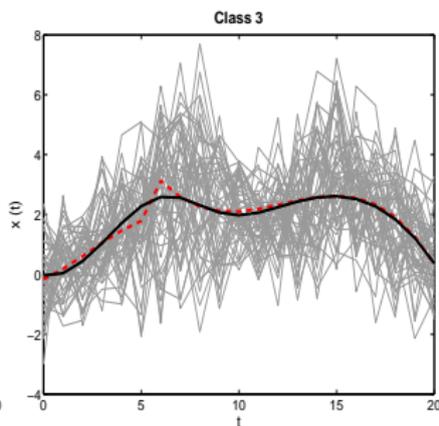
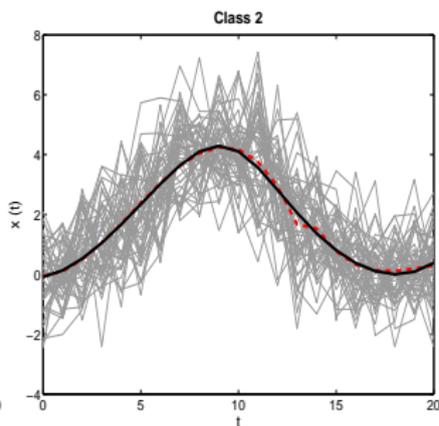
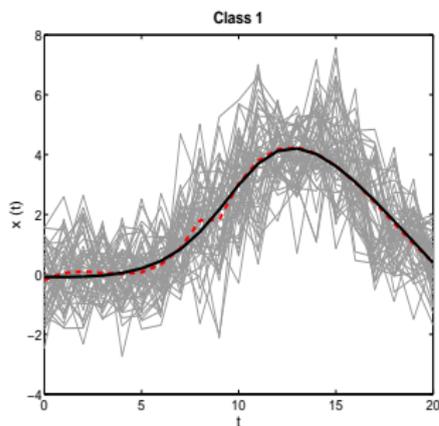


## Vagues de Breiman : [Breiman et al. 84]

- ▶ Problème de 3 classes
- ▶  $m=21$  observations dans chaque courbe

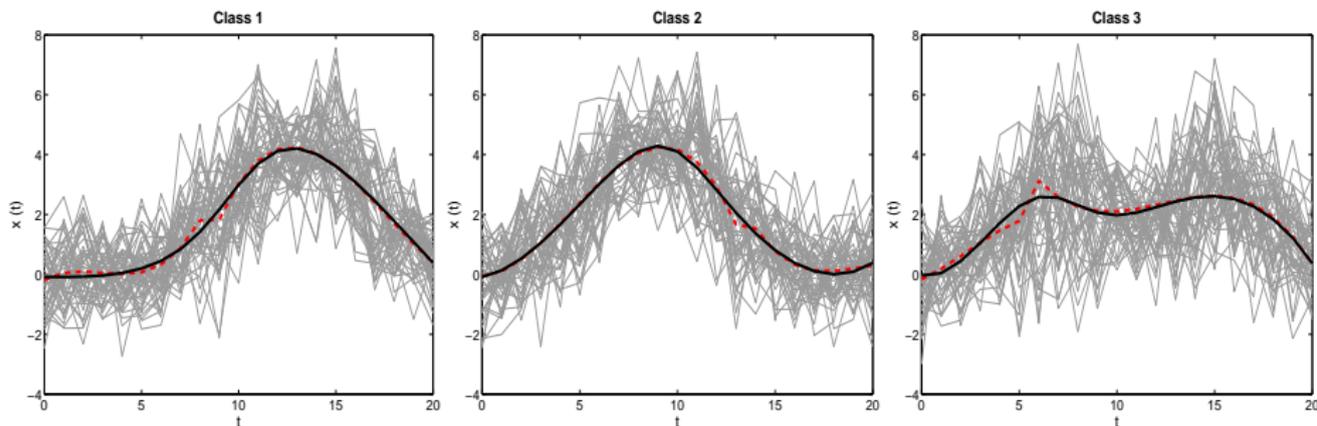
# Vagues de Breiman : [Breiman et al. 84]

- ▶ Problème de 3 classes
- ▶  $m=21$  observations dans chaque courbe
- ▶ Exemple avec 50 courbes par classe



# Vagues de Breiman : [Breiman et al. 84]

- ▶ Problème de 3 classes
- ▶  $m=21$  observations dans chaque courbe
- ▶ Exemple avec 50 courbes par classe

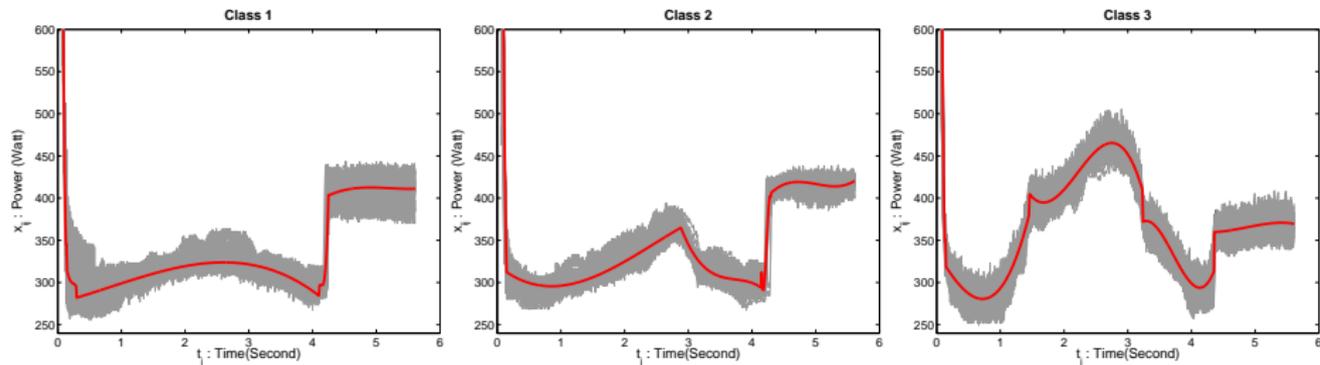


Approche de modélisation	taux d'erreur de test (%)
Régression par morceaux	2.4 (0.64)
Modèle proposé	<b>1.67 (0.84)</b>

TAB.: Taux d'erreur de classement de signaux

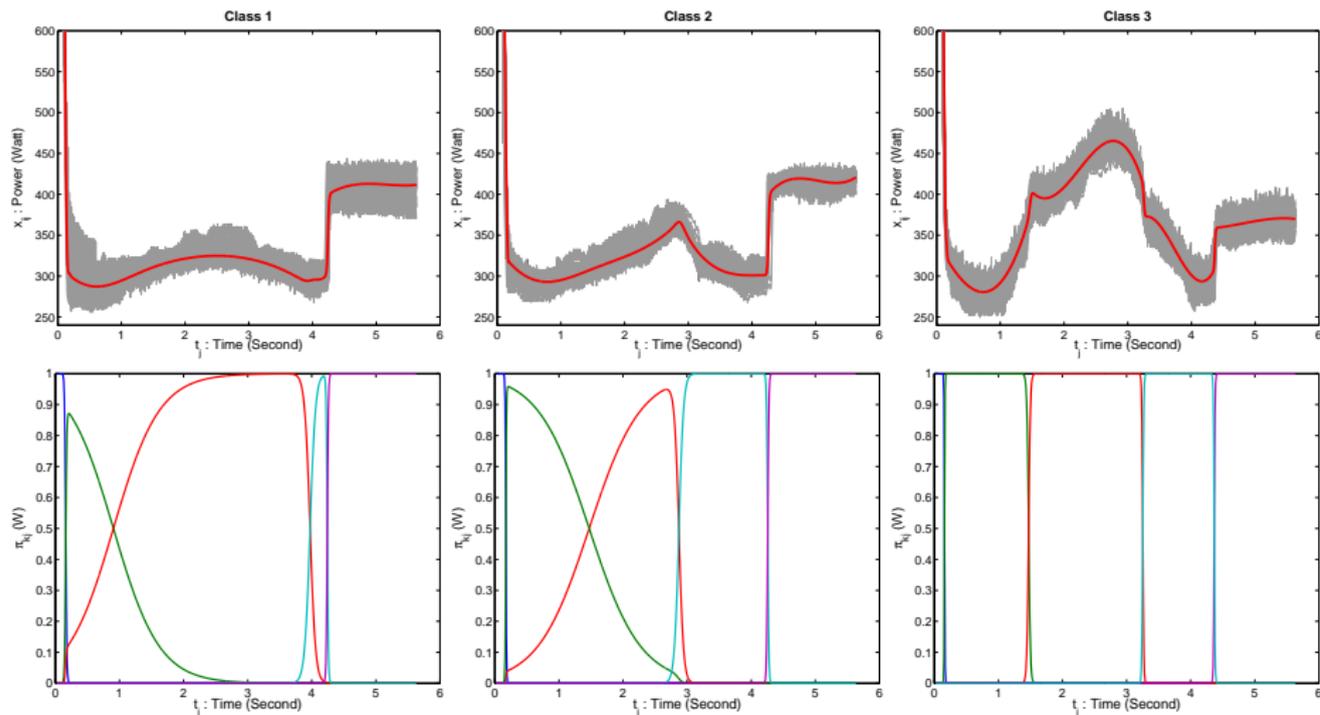
# Signaux de manœuvres d'aiguillage

Les 3 classes de signaux de manœuvres d'aiguillage et leurs modèles estimés par le modèle de régression par morceaux :



# Signaux de manoeuvres d'aiguillage

Les 3 classes de signaux de manoeuvres d'aiguillage, les modèles estimés et les fonctions logistiques associées obtenus par l'approche de régression proposée :



Approche de modélisation	Test error rates (%)
Régression par morceaux	1.82 (5.74)
Modèle proposé	<b>1.67 (2.28)</b>

TAB.: Taux d'erreur de classement des signaux de manœuvres d'aiguillage.

# Conclusion

- ▶ Modèle de régression à processus latent proposé pour modéliser les signaux de manœuvres d'aiguillage.
- ▶ Extension pour apprendre à partir de plusieurs signaux
- ▶ Avantages :
  - Le modèle proposé intègre un processus logistique qui permet de basculer de manière souple ou brusque entre différents sous-modèles polynomiaux
  - L'algorithme EM incluant une procédure de régression logistique fournit un cadre très adapté pour estimer les paramètres du modèle proposé.
- ▶ On obtient des résultats précis en terme de discrimination de signaux

# Perspectives

- ▶ Modélisation de la dynamique du système par un modèle AR à base de processus latents (HMM, logistique) à partir de données évolutives
- ▶ Mise au point d'une approche en ligne pour le suivi

## Publications acceptées

- ▶ F. Chamroukhi, A. Samé, G. Govaert and P. Aknin. "Time series modeling by a regression approach based on a latent process", *Neural Networks Elsevier, 2009.*
- ▶ F. Chamroukhi, A. Samé, G. Govaert and P. Aknin. "A regression model with a hidden logistic process for feature extraction from time series", *International Joint Conference on Neural Networks IJCNN 2009*, Atlanta, USA.
- ▶ F. Chamroukhi, A. Samé, G. Govaert and P. Aknin. "A regression model with a hidden logistic process for signal parametrization", *European Symposium on Artificial Neural Networks ESANN 2009*, Bruges, Belgium.
- ▶ F. Chamroukhi, A. Samé and P. Aknin. "A probabilistic approach for the classification of railway switch operating states", *The Sixth International Conference on Condition Monitoring and Machinery Failure Prevention Technologies 2009*, Dublin, UK.
- ▶ A. Samé, F. Chamroukhi et G. Govaert "Algorithme EM et modèle à processus latent pour la régression non linéaire", *41e journées de la SFDS 2009*, Bordeaux.
- ▶ F. Chamroukhi, A. Samé, P. Aknin and M. Antoni. "Switch mechanism diagnosis using a pattern recognition approach". *The 4th IET International Conference on Railway Condition Monitoring RCM 2008*, Derby, UK.
- ▶ F. Chamroukhi, A. Samé et P. Aknin "Régression à variable latente pour la modélisation de signaux de manoeuvres d'aiguillage". *Actes INRETS 2008.*
- ▶ A. Samé, F. Chamroukhi and P. Aknin "Détection séquentielle de défauts sur des signaux de manoeuvres d'aiguillage". *Workshop Surveillance, Sûreté et Sécurité des Grands Systèmes 3SGS'08*, 4-5 juin 2008, UTT.

## Articles soumis :

- ▶ F. Chamroukhi, A. Samé, G. Govaert and P. Aknin. "A hidden process regression model for supervised learning from curves", soumis à *Neurocomputing Elsevier.*
- ▶ F. Chamroukhi, A. Samé et G. Govaert. "Algorithme EM et modèle à processus latent pour la régression non linéaire", soumis à la *Revue des Nouvelles Technologies de l'Information RNTI.*