

# Classification de Signaux Multidimensionnels Irrégulièrement Échantillonnés

Alexandre CONSTANTIN<sup>1</sup>, Mathieu FAUVEL<sup>2</sup>, Stéphane GIRARD<sup>1</sup>, Serge IOVLEFF<sup>3</sup>

<sup>1</sup>Université Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000 Grenoble, France

<sup>2</sup>CESBIO, Université de Toulouse, CNES/CNRS/IRD/UPS/INRA, Toulouse, France

<sup>3</sup>Laboratoire Paul Painlevé - Université Lille 1, CNRS, Inria, France

Alexandre.Constantin@inria.fr, Mathieu.Fauvel@inra.fr, Stephane.Girard@inria.fr,  
serge.iovleff@inria.fr

**Résumé** – Nous proposons une nouvelle approche statistique basée sur les processus gaussiens pour la classification de signaux échantillonnés irrégulièrement en s’affranchissant du recalage temporel. Ce modèle offre de plus un cadre théorique pour l’imputation des valeurs manquantes. Les premiers résultats sur des données simulées montrent des performances prometteuses en termes de classification et d’imputation.

**Abstract** – A new statistical approach using Gaussian processes is proposed to to classify irregularly sampled signals without temporal rescaling. Moreover, the model offers a theoretical framework to impute missing values. First experiments on simulated data show promising results both in terms of classification and imputation accuracy. Good robustness properties with respect to the modelling assumptions are also observed.

## 1 Introduction

Dans un contexte de changement climatique, la surveillance de l’état et le contrôle de l’évolution des surfaces continentales sont des enjeux majeurs pour les années à venir. L’utilisation de séries temporelles de natures diverses (températures, pluviométrie, image de télédétection optique et radar ...) fournit un ensemble de données précieuses pour ce type de diagnostic. Cependant, le volume des données impose l’utilisation de modèles et d’algorithmes automatiques. On trouve ainsi une vaste littérature scientifique sur cette thématique, incluant la modélisation statistique des séries temporelles [1, 2] ou l’utilisation de séries temporelles en apprentissage statistique [3].

Dans le contexte de cette étude<sup>1</sup>, nous travaillons avec des séries temporelles d’images satellitaires issues de la mission Copernicus Sentinel-2<sup>2</sup> pour la classification de l’occupation des sols. Ces satellites fournissent tous les 5 jours une acquisition complète de la France à une résolution de 10m par pixel pour 4 bandes spectrales. En raison de l’orbite des satellites et de la présence de nuages lors des acquisitions, les pixels observés ont un échantillonnage temporel irrégulier sur l’ensemble du territoire. Chaque pixel a alors une dimension variable (la dimension étant le nombre d’instant d’acquisition multiplié par le nombre de bandes spectrales). En général, une étape de ré-échantillonnage temporel sur une grille commune est effectuée

en pré-traitement [4] pour pouvoir appliquer des algorithmes d’apprentissage sur des données vectorielles de taille fixe. Cependant, cette étape est coûteuse en temps de calcul et peut affecter négativement la précision finale de la classification.

Nous présentons ici une méthode de classification permettant de s’affranchir du recalage temporel. La méthode proposée est basée sur une modélisation des pixels/signaux par des processus gaussiens multivariés. Elles permet d’estimer les paramètres des classes, de prédire les probabilités d’appartenance et de reconstruire les signaux sur une grille temporelle quelconque. Le modèle est décrit paragraphe 2 et l’estimation des paramètres est présentée paragraphe 3. La Partie 4 détaille les procédures de classification et d’imputation des données manquantes. Des résultats de simulations sont résumés paragraphe 5 et des perspectives sont discutées paragraphe 6.

## 2 Processus Gaussiens Multivariés

Notons  $\mathcal{S} = \{(y_i, z_i)\}_{i=1}^n$  un ensemble de signaux aléatoires multivariés labellisés et irrégulièrement échantillonnés. Le signal est modélisé par un vecteur aléatoire  $Y$  contenant  $p$  fonctions (ou processus) aléatoires de carré intégrable  $\mathcal{T} \rightarrow \mathbb{R}^p$ , avec  $\mathcal{T} = [0, T]$  l’intervalle d’acquisition du signal. Cette propriété est notée  $Y \in L_2^p(\mathcal{T})$ . La classe associée à chaque signal est modélisée par la variable aléatoire discrète  $Z$  prenant ses valeurs dans  $\{1, \dots, \mathcal{C}\}$ . Le nombre d’échantillons disponibles est noté  $n$ .

Le modèle introduit ici repose sur deux hypothèses : 1) Les processus coordonnées  $Y_b$ ,  $b \in \{1, \dots, p\}$  de  $Y$  sont indépen-

1. Ce travail bénéficie du soutien de l’Agence Nationale de la Recherche par le programme Investissements d’Avenir (ANR-15-IDEX-02) et du Centre National d’Etudes Spatiales (CNES).

2. [https://www.esa.int/Our\\_Activities/Observing\\_the\\_Earth/Copernicus/Sentinel-2](https://www.esa.int/Our_Activities/Observing_the_Earth/Copernicus/Sentinel-2)

dants, 2) Chaque processus  $Y_b$  est, conditionnellement à  $Z = c$ , un processus gaussien. Cela conduit à

$$Y_b(t)|Z = c \sim \mathcal{GP}(m_{b,c}(t), K_{b,c}(t, s)), \quad (1)$$

avec  $t \in \mathcal{T}$ ,  $s \in \mathcal{T}$ ,  $c = \{1, \dots, C\}$ ,  $b = \{1, \dots, p\}$  et où  $\mathcal{GP}(m_{b,c}, K_{b,c})$  désigne un processus gaussien de moyenne  $m_{b,c} \in L_2(\mathcal{T})$  et d'opérateur de covariance  $K_{b,c} : L_2(\mathcal{T}) \times L_2(\mathcal{T}) \rightarrow \mathbb{R}$ . Nous supposons de plus que nous observons une version bruitée du signal :

$$Y_b = \tilde{Y}_b + \eta_b \quad (2)$$

avec  $\eta_b(t)|Z = c \sim \mathcal{GP}(0, \sigma_{b,c}^2 \delta_{t,s})$ . Dans le contexte d'échantillonnage irrégulier, un signal  $\mathbf{y}_i$  est observé aux  $T_i$  instants  $\{t_1^i, \dots, t_{T_i}^i\} \in \mathcal{T}$  et son bème processus coordonnée est représenté par un vecteur de  $\mathbb{R}^{T_i}$

$$\mathbf{y}_{i,b} = [Y_b^i(t_1^i), \dots, Y_b^i(t_{T_i}^i)]^\top. \quad (3)$$

Dans le cadre des données Sentinel-2,  $\mathbf{y}_{i,b}$  est le vecteur des observations de la bande  $b$  du pixel  $i$  aux instants  $t_1^i$  à  $t_{T_i}^i$ . Par définition des processus gaussiens, ce vecteur est, conditionnellement à  $Z = c$ , un vecteur gaussien

$$\mathbf{y}_{i,b}|Z_i = c \sim \mathcal{N}_{T_i}(\boldsymbol{\mu}_{i,b,c}, \boldsymbol{\Sigma}_{b,c}^i), \quad (4)$$

où  $\mathcal{N}_{T_i}$  est la loi normale multivariée sur  $\mathbb{R}^{T_i}$ ,  $\boldsymbol{\mu}_{i,b,c} = [m_{b,c}(t_1^i), \dots, m_{b,c}(t_{T_i}^i)]^\top$  et  $\boldsymbol{\Sigma}_{b,c}^i[t_j^i, t_{j'}^i] = K_{b,c}(t_j^i, t_{j'}^i)$ .

Ainsi, sous l'hypothèse d'indépendance entre les bandes, le  $i$ ème échantillon  $\mathbf{y}_i = [\mathbf{y}_{i,1}, \dots, \mathbf{y}_{i,p}]^\top \in \mathbb{R}^{pT_i}$  est, conditionnellement à  $Z = c$ , un vecteur gaussien  $pT_i$ -varié dont la distribution s'écrit comme un produit de lois normales :

$$\mathbf{y}_i|Z_i = c \sim \prod_{b=1}^p \mathcal{N}_{T_i}(\boldsymbol{\mu}_{i,b,c}, \boldsymbol{\Sigma}_{b,c}^i). \quad (5)$$

Les paramètres inconnus du modèle sont les fonctions moyennes  $m_{c,b}$ , les opérateurs de covariances  $K_{b,c}$  et les niveaux de bruit  $\sigma_{b,c}^2$ . Les opérateurs de covariances sont exprimés classiquement par des fonctions définies positives paramétriques sous la forme suivante :

$$K_{b,c}(t, s) = \gamma_{b,c}^2 k(t, s|h_{b,c}) + \sigma_{b,c}^2 \delta_{t,s} \quad (6)$$

avec par exemple  $k(t, s|h_{b,c}) = \exp\left\{-\frac{(t-s)^2}{2h_{b,c}}\right\}$ . On regroupe l'ensemble des hyperparamètres sous le symbole  $\boldsymbol{\theta}_{b,c} = \{\gamma_{b,c}^2, h_{b,c}, \sigma_{b,c}^2\}$ , on note alors  $\boldsymbol{\Sigma}_{b,c}^i = \boldsymbol{\Sigma}^i(\boldsymbol{\theta}_{b,c})$ .

Afin de gérer l'échantillonnage irrégulier, nous introduisons un ensemble de  $J_b$  bases fixes de  $L_2(\mathcal{T})$ ,  $\{\varphi_{b,j}\}_{j=1}^{J_b}$ , sur lesquelles les fonctions moyennes sont exprimées :

$$m_{b,c}(t) = \sum_{j=1}^{J_b} \alpha_{b,c,j} \varphi_{b,j}(t), \quad t \in \mathcal{T}, \quad (7)$$

où  $\alpha_{b,c,j}$  est le coefficient de la moyenne  $m_{b,c}$  sur la fonction la base  $j$ . En notant  $\boldsymbol{\alpha}_{b,c} = [\alpha_{b,c,1}, \dots, \alpha_{b,c,J_b}]^\top \in \mathbb{R}^{J_b}$  et  $\mathbf{B}_b^i$  la matrice de *design* de taille  $T_i \times J_b$  définie par  $B_{\ell,j}^i = \varphi_j(t_\ell^i)$

pour  $\ell = \{1, \dots, T_i\}$ , nous avons  $\boldsymbol{\mu}_{i,b,c} = \mathbf{B}_b^i \boldsymbol{\alpha}_{b,c}$  et l'eq. (5) devient

$$\mathbf{y}_i|Z_i = c \sim \prod_{b=1}^p \mathcal{N}_{T_i}(\mathbf{B}_b^i \boldsymbol{\alpha}_{b,c}, \boldsymbol{\Sigma}^i(\boldsymbol{\theta}_{b,c})). \quad (8)$$

Déjà étudié en regression [5, Section 2.7], cette utilisation des processus gaussiens en classification est, à notre connaissance, originale. Les paramètres  $\boldsymbol{\alpha}_{b,c}$ ,  $\boldsymbol{\theta}_{b,c}$ ,  $\gamma_{b,c}^2$  et  $\sigma_{b,c}^2$  pour  $c = \{1, \dots, C\}$ ,  $b = \{1, \dots, p\}$  sont estimés en maximisant la log-vraisemblance du modèle, comme détaillé dans la partie suivante.

### 3 Inférence par maximum de vraisemblance

En écrivant la log-vraisemblance selon l'équation (8) et en simplifiant selon les hypothèses du modèle, on obtient  $C \times p$  problèmes indépendants de minimisation :

$$(\hat{\boldsymbol{\alpha}}_{b,c}, \hat{\boldsymbol{\theta}}_{b,c}) = \arg \min_{\boldsymbol{\alpha}_{b,c}, \boldsymbol{\theta}_{b,c}} \ell_{b,c}(\boldsymbol{\alpha}_{b,c}, \boldsymbol{\theta}_{b,c}) \quad (9)$$

où  $\ell_{b,c}(\boldsymbol{\alpha}_{b,c}, \boldsymbol{\theta}_{b,c})$  est donnée par

$$\sum_{i|Z_i=c} \log |\boldsymbol{\Sigma}^i(\boldsymbol{\theta}_{b,c})| + (\mathbf{y}_{i,b} - \mathbf{B}_b^i \boldsymbol{\alpha}_{b,c})^\top \boldsymbol{\Sigma}^i(\boldsymbol{\theta}_{b,c})^{-1} (\mathbf{y}_{i,b} - \mathbf{B}_b^i \boldsymbol{\alpha}_{b,c}). \quad (10)$$

Chaque sous-problème de (9) peut être résolu par une technique de descente de gradient classiquement employée pour les processus gaussiens de régression [5, Chapitre 5]. Dans un premier temps, en annulant le gradient en  $\boldsymbol{\alpha}$  dans (10), on obtient la formule explicite suivante :

$$\hat{\boldsymbol{\alpha}}_{b,c}(\boldsymbol{\theta}_{b,c}) = \left[ \sum_{i|Z_i=c} \mathbf{B}_b^{i\top} \boldsymbol{\Sigma}^i(\boldsymbol{\theta}_{b,c})^{-1} \mathbf{B}_b^i \right]^{-1} \left[ \sum_{i|Z_i=c} \mathbf{B}_b^{i\top} \boldsymbol{\Sigma}^i(\boldsymbol{\theta}_{b,c})^{-1} \mathbf{y}_{i,b} \right]. \quad (11)$$

Dans un second temps, la minimisation en  $\boldsymbol{\theta}_{b,c}$  à  $\hat{\boldsymbol{\alpha}}_{b,c}$  fixé est réalisée numériquement par un algorithme de gradient sous contraintes de positivité [6] pour certains hyperparamètres en utilisant le gradient de (10) :

$$\frac{\partial}{\partial \theta_k} \ell_{b,c} = \sum_{i|Z_i=c} \text{tr} \left( \left( \boldsymbol{\Sigma}^i(\boldsymbol{\theta}_{b,c})^{-1} - \boldsymbol{\beta}_i \boldsymbol{\beta}_i^\top \right) \frac{\partial \boldsymbol{\Sigma}^i(\boldsymbol{\theta}_{b,c})}{\partial \theta_k} \right) \quad (12)$$

avec  $\boldsymbol{\beta}_i = \boldsymbol{\Sigma}^i(\boldsymbol{\theta}_{b,c})^{-1} (\mathbf{y}_i - \mathbf{B}_b^i \boldsymbol{\alpha}_{b,c})$ .

### 4 Classification et imputation

Le modèle proposé permet à la fois de prédire la classe d'appartenance du signal multivarié et d'imputer ses valeurs manquantes (par exemple pour des signaux satellites, de retrouver

une information qui serait masquée par un nuage). Dans la suite de ce paragraphe,  $\mathbf{y}_j$  représente un nouveau signal à classer dont le nombre d'instants d'observation est noté  $T_j$ .

## 4.1 Classification par MAP

On note  $f_d(\cdot, \boldsymbol{\mu}, \boldsymbol{\Sigma})$  la densité gaussienne  $d$ -variée de moyenne  $\boldsymbol{\mu}$  et de matrice de covariance  $\boldsymbol{\Sigma}$ , et  $\pi_c = P(Z = c)$  les probabilités *a priori* des classes. Pour calculer la probabilité *a posteriori*  $P(Z = c | \mathbf{y}_j)$ , nous utilisons le modèle de mélange paramétrique gaussien qui découle de l'équation (8) obtenue pour les  $T_j$  instants d'observation de  $\mathbf{y}_j$  :

$$P(Z = c | \mathbf{y}_j) = \frac{\hat{\pi}_c \prod_{b=1}^p f_{T_j}(\mathbf{y}_j, \mathbf{B}_b^j \hat{\boldsymbol{\alpha}}_{b,c}, \boldsymbol{\Sigma}^j(\hat{\boldsymbol{\theta}}_{b,c}))}{\sum_{\ell=1}^K \hat{\pi}_\ell \prod_{b=1}^p f_{T_j}(\mathbf{y}_j, \mathbf{B}_b^j \hat{\boldsymbol{\alpha}}_{b,\ell}, \boldsymbol{\Sigma}^j(\hat{\boldsymbol{\theta}}_{b,\ell}))}. \quad (13)$$

Les  $\pi_c$  sont estimés par leur valeur empirique  $\hat{\pi}_c = n_c/n$  où  $n_c$  est le nombre d'échantillons de la classe  $c$  dans  $\mathcal{S}$ . Le signal  $\mathbf{y}_j$  est alors affecté à la classe la plus probable au sens du *maximum a posteriori* (MAP).

## 4.2 Imputation de données

Dans le cadre de l'imputation de données manquantes, deux cas sont envisagés, selon que la classe d'appartenance est connue ou estimée.

**Imputation conditionnelle** Pour un instant  $t^*$ ,  $Y_b^j(t^*)$  est imputée par l'espérance conditionnelle :

$$\hat{Y}_{b,c}^j(t^*) := \mathbb{E}[Y_b^j(t^*) | \mathbf{y}_j, Z_j = c]. \quad (14)$$

En utilisant l'hypothèse d'indépendance entre les bandes

$$\mathbb{E}[Y_b^j(t^*) | \mathbf{y}_j, Z_j = c] = \mathbb{E}[Y_b^j(t^*) | \mathbf{y}_{j,b}, Z_j = c]. \quad (15)$$

Par ailleurs, pour une classe donnée  $Z_j = c$ , la loi jointe entre valeurs observées et non observées est donnée par

$$\begin{pmatrix} Y_b^j(t^*) \\ \mathbf{y}_{j,b} \end{pmatrix} \sim \mathcal{N} \left( \begin{bmatrix} m_{b,c}(t^*) \\ \mathbf{B}_b^j \boldsymbol{\alpha}_{b,c} \end{bmatrix}, \begin{bmatrix} K(t^*, t^* | \boldsymbol{\theta}_{b,c}) & \mathbf{k}(t^*, t_{1:T_j}^j | \boldsymbol{\theta}_{b,c})^\top \\ \mathbf{k}(t^*, t_{1:T_j}^j | \boldsymbol{\theta}_{b,c}) & \boldsymbol{\Sigma}^j(\boldsymbol{\theta}_{b,c}) \end{bmatrix} \right), \quad (16)$$

où  $\mathbf{k}(t^*, t_{1:T_j}^j | \boldsymbol{\theta}_{b,c}) = [K(t^*, t_1^j | \boldsymbol{\theta}_{b,c}), \dots, K(t^*, t_{T_j}^j | \boldsymbol{\theta}_{b,c})]$ . En appliquant les propriétés des lois conditionnelles des vecteurs gaussiens [7, page 63], la valeur imputée et sa variance associée s'écrivent :

$$\hat{Y}_{b,c}^j(t^*) = \mathbf{B}_b^j \boldsymbol{\alpha}_{b,c} + \mathbf{k}(t^*, t_{1:T_j}^j | \boldsymbol{\theta}_{b,c}) \boldsymbol{\Sigma}^j(\hat{\boldsymbol{\theta}}_{b,c})^{-1} (\mathbf{y}_{i,b} - \mathbf{B}_b^j \boldsymbol{\alpha}_{b,c}) \quad (17)$$

$$\text{var}\{\hat{Y}_{b,c}^j(t^*)\} = K(t^*, t^* | \hat{\boldsymbol{\theta}}_{b,c}) - \mathbf{k}(\hat{\boldsymbol{\theta}}_{b,c} | t^*, t_{1:T_j}^j) \boldsymbol{\Sigma}^j(\hat{\boldsymbol{\theta}}_{b,c})^{-1} \mathbf{k}(t^*, t_{1:T_j}^j | \hat{\boldsymbol{\theta}}_{b,c})^\top. \quad (18)$$

**Imputation globale** La valeur manquante est imputée en utilisant l'espérance conditionnelle suivante

$$\hat{Y}_b^j(t^*) = \mathbb{E}[Y_b^j(t^*) | \mathbf{y}_j] = \sum_{c=1}^c P(Z_j = c | \mathbf{y}_j) \hat{Y}_{b,c}^j(t^*), \quad (19)$$

et la variance associée est donnée par

$$\sum_{c=1}^c P(Z_j = c | \mathbf{y}_j) \mathbb{E}[Y_b^j(t^*)^2 | \mathbf{y}_j, Z_j = c] - \mathbb{E}[Y_b^j(t^*) | \mathbf{y}_j]^2. \quad (20)$$

Après simplification, nous obtenons

$$\text{var}\{\hat{Y}_b^j(t^*)\} = \sum_{c=1}^c P(Z_j = c | \mathbf{y}_j) \text{var}\{\hat{Y}_{b,c}^j(t^*)\} - \hat{Y}_b^j(t^*)^2 + \sum_{c=1}^c P(Z_j = c | \mathbf{y}_j) \hat{Y}_{b,c}^j(t^*)^2, \quad (21)$$

où les quantités  $\hat{Y}_{b,c}^j(t^*)$  et  $\text{var}\{\hat{Y}_{b,c}^j(t^*)\}$  sont données par (17) et (18). La variance de la valeur prédite se décompose classiquement en un premier terme de variance intra-classes et un second terme de variance inter-classes. Enfin, ces calculs s'étendent à l'imputation simultanée de plusieurs valeurs.

## 5 Illustration sur Données Simulées

Le modèle est évalué sur données simulées : 80 échantillons sont générés pour deux classes  $c \in \{0, 1\}$ . Une grille temporelle régulière est utilisée pour laquelle des instants sont rejetés aléatoirement (suivant une loi de Bernoulli) pour simuler un échantillonnage irrégulier. Pour chaque individu, et chaque classe, les signaux échantillonnés et bruités (2) sont générés selon la loi normale multivariée (4) avec  $\boldsymbol{\mu}_c = \mathbf{B}_{\text{sin}} \boldsymbol{\alpha}_c$  tels que  $B_{\text{sin}}(t) = \{\sin(j\pi t/T)\}_{0 \leq j \leq 9}$  pour tout  $t \in [0, T]$ ,  $T = 50$  et

$$\begin{aligned} \boldsymbol{\alpha}_0 &= [0, 0.99, 0.72, 0.83, 0.61, 0.57, 1, 1, 1, 1], \\ \boldsymbol{\alpha}_1 &= [0, 0.99, 0.72, 0.83, 0.61, 0.57, 0.6, 0.6, 0.6, 0.6] \end{aligned}$$

décrivant deux classes dont les moyennes sont séparables en hautes fréquences. De plus, les hyperparamètres de l'opérateur de covariance ont été fixés à  $\boldsymbol{\theta}_0 = \{\gamma_0^2, h_0, \sigma_0^2\} = \{0.05, 1, 0.3\}$  et  $\boldsymbol{\theta}_1 = \{0.08, 0.5, 0.3\}$ , le bruit est élevé pour observer un certain recouvrement entre les classes.

### 5.1 Classification Supervisée

Nous effectuons 50 simulations indépendantes. Pour chaque simulation, le modèle est entraîné sur 75% des données simulées et testé sur les 25% restantes pour un nombre moyen d'instants temporels donné  $n_t$ . Le modèle est entraîné avec deux types de bases : la première est celle utilisée pour générer les données, la seconde est une famille d'exponentielles  $B_{\text{exp}}(t) = \{\exp(-jt)\}_{0 \leq j \leq 9}$  pour tout  $t \in [0, T]$ . Elle est utilisée pour évaluer la robustesse à l'erreur de modèle. Nous notons le taux de bonne prédiction par base  $Acc_{\text{exp}}$  et  $Acc_{\text{sin}}$ .

TABLE 1 – Taux de bonne classification en fonction du nombre moyen d’instant temporels

$n_t$	5	10	25	50	75
Acc <sub>exp</sub> (%)	52.8	52.9	74.3	93.9	94.2
Acc <sub>sin</sub> (%)	64.3	85.3	100	100	100

Le Tableau 1 montre les résultats de classification en fonction de  $n_t$ . Avec une base adaptée, la classification est très bonne même avec peu d’instant temporels. La classification semble robuste au changement de base, sous condition d’un plus grand nombre d’échantillons temporels (comparé au nombre requis pour la base initiale).

## 5.2 Imputation de Données Manquantes

Pour évaluer les performances de la phase d’imputation conditionnelle, nous proposons de masquer environ 300 instant temporels répartis dans 20 échantillons par classe. Nous évaluons ensuite les erreurs de reconstruction normalisées par l’écart-type prédit qui, pour un processus gaussien, ont une distribution normale centrée-réduite  $\mathcal{N}(0, 1)$ .

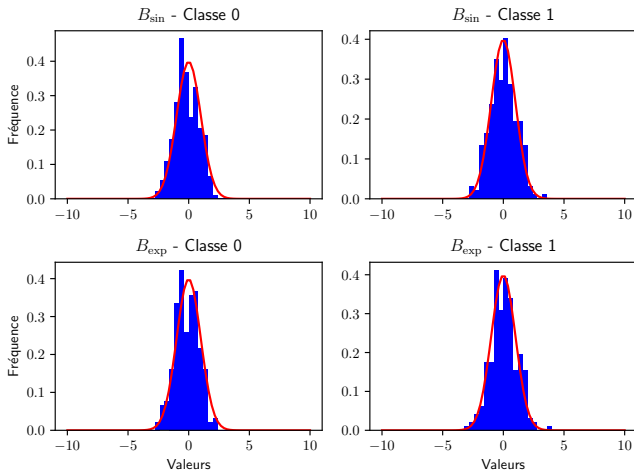


FIGURE 1 – Distribution des erreurs de reconstruction

Sur la Figure 1 nous comparons l’histogramme des erreurs (en bleu), à la densité théorique (en trait plein rouge) de la loi  $\mathcal{N}(0, 1)$ . Pour ces simulations, l’influence de la base temporelle semble moins importante que pour la classification.

La Figure 2 illustre une imputation sur l’ensemble de définition du domaine temporel pour deux signaux simulés dans une même classe. On obtient à la fois la valeur signal ainsi que l’incertitude de reconstruction. Des résultats équivalents sont obtenus par imputation globale.

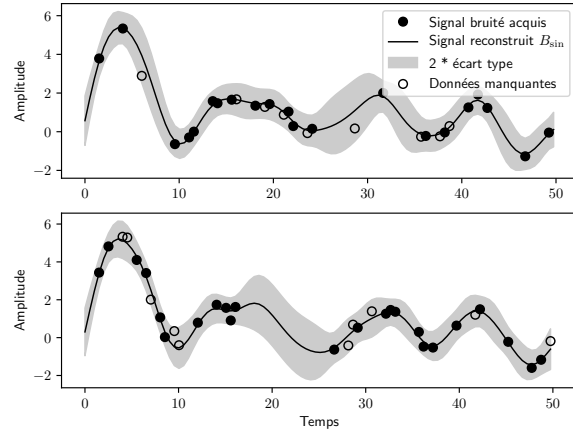


FIGURE 2 – Exemples d’imputation

## 6 Conclusion et Perspectives

Nous avons présenté un modèle permettant de gérer des échantillons multivariés irrégulièrement échantillonnés. Il repose sur une modélisation par processus gaussiens et une hypothèse d’indépendance entre les différents canaux. L’optimisation des paramètres est réalisée par maximisation de la vraisemblance et, une fois les valeurs optimales des paramètres trouvées, le modèle permet de classer les signaux et de reconstruire les valeurs manquantes. Le comportement numérique du modèle a été validé sur des simulations. Nous travaillons également à la mise en œuvre sur une base de données réelles de grande taille. Enfin, nos travaux futurs seront consacrés au choix adaptatif de bases de décomposition des moyennes ainsi qu’à la levée de l’hypothèse d’indépendance des bandes voire des pixels.

## Références

- [1] M. B. Priestley. *Spectral analysis and time series* / M.B. Priestley. Academic Press London; New York, 1981.
- [2] P. J. Brockwell and R. A. Davis. *Time Series : Theory and Methods*. Springer-Verlag, Berlin, Heidelberg, 1986.
- [3] J. Ramsay and B.W. Silverman. *Functional Data Analysis*. Springer Series in Statistics. Springer, 2005.
- [4] J. Inglada, A. Vincent, M. Arias, B. Tardy, D. Morin, and I. Rodes. Operational high resolution land cover map production at the country scale using satellite image time series. *Remote Sensing*, 9(1), 2017.
- [5] C. K. Williams and C. E. Rasmussen. Gaussian processes for machine learning. *the MIT Press*, 2(3) :4, 2006.
- [6] C. Zhu, R. H. Byrd, P. Lu, and J. Nocedal. Algorithm 778 : L-BFGS-B : Fortran subroutines for large-scale bound-constrained optimization. *ACM Transactions on Mathematical Software (TOMS)*, 23(4) :550–560, 1997.
- [7] M. Bilodeau and D. Brenner. *Theory of multivariate statistics*. Springer Science & Business Media, 2008.