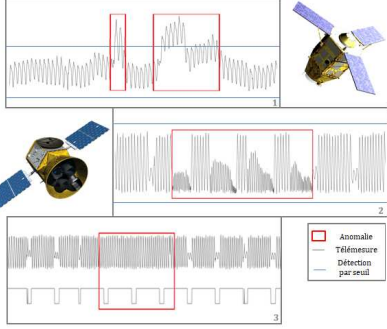


## 1. Détection d'anomalies dans la télémétrie satellite



- Des milliers de paramètres de télémétrie.
- Détection d'anomalies univariées (#1, #2) et multivariées (#3).
- Application de méthodes de machine learning ou apprentissage statistique semi-supervisé.
- Evaluation des méthodes étudiées sur une base de test composée de 10 signaux (3 discrets et 7 continus) et munie d'une vérité terrain.

## 2. Représentation parcimonieuse pour la détection d'anomalies

### 2.1. ADDICT : Anomaly Detection using a Dictionary [1]

$$\begin{bmatrix} y_D \\ y_C \end{bmatrix} = \begin{bmatrix} \Phi_D \\ \Phi_C \end{bmatrix} \begin{bmatrix} x_D \\ x_C \end{bmatrix} + \begin{bmatrix} e_D \\ e_C \end{bmatrix} + \begin{bmatrix} b_D \\ b_C \end{bmatrix}$$

■ Élément nul  
■ Élément non nul  
■ Élément non sélectionné

#### Représentation parcimonieuse discrète :

$$\min_{x_D, e_D} \min_{e_C} \frac{1}{2} \|y_D - \Phi_D x_D - e_D\|_2^2 + b_D \sum_k \|e_{D,k}\|_2 \quad (1)$$

où  $\mathcal{B}$  est la base canonique de  $\mathbb{R}^L$ . Le modèle proposé décrit un problème combinatoire où le signal discret  $y_D$  est approximé par un seul atome du dictionnaire discret  $\Phi_D$ .

#### Représentation parcimonieuse continue :

$$\min_{x_C, e_C} \frac{1}{2} \|y_C - \Phi_C x_C - e_C\|_2^2 + a_C \|x_C\|_1 + b_C \sum_k \|e_{C,k}\|_2 \quad (2)$$

où

- $\mathbf{y} = [y_D^T, y_C^T]^T \in \mathbb{R}^N$  est un vecteur de données mixtes divisé en un sous-signal discret noté  $y_D \in \mathbb{R}^{N_D}$  et un sous-signal continu noté  $y_C \in \mathbb{R}^{N_C}$ .
- $\Phi \in \mathbb{R}^{N \times 2L}$  est un dictionnaire appris à partir de données nominales (sans anomalie).
- $\mathbf{x} = [x_D^T, x_C^T]^T \in \mathbb{R}^{2L}$  est un vecteur parcimonieux de coefficients.
- $\mathbf{e} = [e_D^T, e_C^T]^T \in \mathbb{R}^N$  est un vecteur d'anomalie ( $\mathbf{e} = 0$  en l'absence d'anomalie).
- $\mathbf{e}_{C,k} \in \mathbb{R}^m, k = 1, \dots, K_C$  (resp.  $\mathbf{e}_{D,k} \in \mathbb{R}^m, k = 1, \dots, K_D$ ) est le  $k^{\text{ème}}$  sous-signal de  $\mathbf{e}_C$  (resp.  $\mathbf{e}_{D,k}$ ) associé au  $k^{\text{ème}}$  paramètre continu (resp. discret).
- $\mathbf{b} \in \mathbb{R}^N$  est un bruit additif.
- $\mathcal{M} = \{l \in \{1, \dots, L\} \mid \|\mathbf{e}_{D,l}\|_2 = 0\}, l = 1, \dots, L$ .
- $\Phi_{\mathcal{M}} = \{\phi_{l+L}, l \in \mathcal{M}\}$
- $a_C, b_C$  et  $b_D$  sont les hyperparamètres contrôlant la parcimonie de  $x_C, e_C$  et  $e_D$ .

### 2.2. Règle de détection

$$\text{anomalie détectée si } \begin{cases} \|\hat{\mathbf{e}}_D\|_2 > 0 & \text{(Discrète)} \\ \|\hat{\mathbf{e}}_C\|_2 > S_{\text{PFA}} & \text{(Continue/Multivariée)} \end{cases} \quad (3)$$

### 2.3. Algorithme ADDM [2]

Introduction d'une variable auxiliaire  $\mathbf{z}$  et d'une contrainte d'égalité.

$$\hat{\mathbf{x}}, \hat{\mathbf{e}} = \arg \min_{\mathbf{x}, \mathbf{e}} \left[ \frac{1}{2} \|\mathbf{y} - \Phi \mathbf{x} - \mathbf{e}\|_2^2 + a \|\mathbf{z}\|_1 + b \sum_k \|e_k\|_2 \right]$$

$$\text{s.t. } \mathbf{z} = \mathbf{x} \quad (4)$$

La mise à jour de  $\mathbf{x}, \mathbf{e}$  et  $\mathbf{z}$  est obtenue par minimisation du Lagrangien augmenté

$$\mathcal{L}_a(\mathbf{x}, \mathbf{z}, \mathbf{e}, \mathbf{m}, \mu) = \frac{1}{2} \|\mathbf{y} - \Phi \mathbf{x} - \mathbf{e}\|_2^2 + a \|\mathbf{z}\|_1 + b \sum_k \|e_k\|_2 + \langle \mathbf{m}, \mathbf{z} - \mathbf{x} \rangle + \frac{\mu}{2} \|\mathbf{z} - \mathbf{x}\|_2^2 \quad (5)$$

L'algorithme ADDM est itératif et estime  $\mathbf{x}, \mathbf{z}, \mathbf{e}$  et  $\mathbf{m}$  de manière alternée jusqu'à convergence. Les mises à jour des différentes variables à l'itération  $k+1$  sont données ci-dessous.

- $\mathbf{x}^{k+1} = (\Phi^T \Phi + \mu^k I)^{-1} [\Phi^T (\mathbf{y} - \mathbf{e}^k) + \mathbf{m}^k + \mu^k \mathbf{z}^k]$
- $\mathbf{z}^{k+1} = S_\gamma(\mathbf{x}^{k+1} - \frac{1}{\mu^k} \mathbf{m}^k)$
- $\mathbf{e}^{k+1} = T_b(\mathbf{y} - \Phi \mathbf{x}^{k+1})$
- $\mathbf{m}^{k+1} = \mathbf{m}^k + \mu^k (\mathbf{z}^{k+1} - \mathbf{x}^{k+1})$
- $\mu^{k+1} = \rho \mu^k$

où  $S_\gamma$  est l'opérateur de seuillage par élément avec  $\gamma = \frac{a}{\mu}$  et  $T_b$  est l'opérateur de seuillage par groupe.

## 3. Premiers résultats & comparaison à l'état de l'art

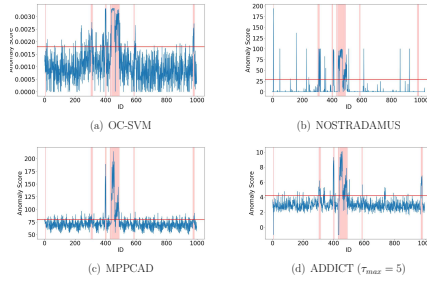


Figure 1: Scores d'anomalie retournés par OC-SVM [3] (a), MPPCAD [4] (b), NOSTRADAMUS [5] (c) et ADDICT [1] (d) pour le jeu de données de test. Les périodes d'anomalies de la vérité terrain sont indiquées en rouge et représentées dans les cadres rouges. Les seuils de détection donnés dans la Table 1 sont représentés par une droite rouge.

Table 1: Probabilité de détection  $P_D$  et probabilité de fausses alarmes  $P_{FA}$  retournées par OC-SVM, MPPCAD, NOSTRADAMUS et ADDICT.

| Méthode                            | Seuil  | $P_D$  | $P_{FA}$ |
|------------------------------------|--------|--------|----------|
| OC-SVM                             | 0.0018 | 80.85% | 7%       |
| MPPCAD                             | 10.8   | 70%    | 12%      |
| NOSTRADAMUS                        | 29     | 77.26% | 6%       |
| ADDICT ( $\tau_{\text{max}} = 0$ ) | 5.2    | 80%    | 6.16%    |
| ADDICT ( $\tau_{\text{max}} = 5$ ) | 5.4    | 81%    | 3%       |

## 4. Extensions ADDICT

### 4.1. Weighted-ADDICT (W-ADDICT)[6]

Modèle continu proposé

$$\min_{\mathbf{x}, \mathbf{e}} \frac{1}{2} \|\mathbf{y} - \Phi \mathbf{x} - \mathbf{e}\|_2^2 + a \|\mathbf{x}\|_1 + b \sum_{k=1}^{K_C} \|\mathbf{w}_k \mathbf{e}_k\|_2$$

- $\mathbf{w}_k \in \mathbb{R}^{+*}$  est un poids associé au  $k^{\text{ème}}$  paramètre. Le poids est mis à jour à chaque itération de l'algorithme à partir du coefficient de corrélation entre le signal test et sa reconstruction.

#### Intérêt de W-ADDICT

- Intégration d'informations externes par l'intermédiaire de poids permettant d'améliorer la détection.

### 4.2. Convolutional-ADDICT (C-ADDICT) [7]

$$\text{Modèle proposé } \arg \min_{\{\mathbf{x}_{l,k}\}, \{\mathbf{e}_k\}} \frac{1}{2} \sum_k \left\| \sum_l \phi_{l,k} * \mathbf{x}_{l,k} + \mathbf{e}_k - \mathbf{y}_k \right\|_2^2 + \lambda \sum_{k,l} \|\mathbf{x}_{l,k}\|_1 + b \sum_k \|e_k\|_2$$

$$\text{s.t. } \mathbf{x}_{l,k} = \mathbf{x}_{l,k} = \dots = \mathbf{x}_{l,k} \quad (6)$$

#### Intérêt de C-ADDICT

- Traitement conjoint des paramètres discrets et des paramètres continus par une unique stratégie de décomposition.
- Robustesse au choix du seuil de détection.

### 4.3. Generalized-ADDICT (G-ADDICT)

Modèle continu proposé

$$\min_{\mathbf{x}, \mathbf{e}} \frac{1}{2} \|\mathbf{y} - \Phi \mathbf{x} - \mathbf{e}\|_2^2 + a \|\mathbf{x}\|_1 + \sum_{k=1}^K b_k \|e_k\|_2$$

- $b_k \in \mathbb{R}^{+*}$  est l'hyperparamètre contrôlant la parcimonie du  $k^{\text{ème}}$  signal d'anomalie.

#### Intérêt de G-ADDICT

- Estimation automatique des hyperparamètres.
- Robustesse au choix du seuil de détection.

## 4.4. Evaluation des performances

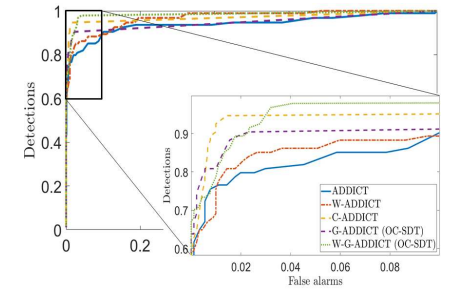


Figure 2: Courbes ROC pour ADDICT et ses extensions.

Table 2: Valeurs de  $P_D$  et  $P_{FA}$  obtenues par ADDICT, W-ADDICT, C-ADDICT, G-ADDICT et W-G-ADDICT.

| Méthode             | Seuil | $P_D$  | $P_{FA}$ |
|---------------------|-------|--------|----------|
| ADDICT              | 4.2   | 80.85% | 3%       |
| W-ADDICT            | 4.5   | 83%    | 2.6%     |
| C-ADDICT            | 0     | 94.7%  | 1.7%     |
| G-ADDICT (OC-SDT)   | 0     | 90.4%  | 2.32%    |
| W-G-ADDICT (OC-SDT) | 0     | 98%    | 4%       |

## 5. Conclusion

- Une méthode de détection d'anomalies capable de traiter de manière conjointe des données continues et des données discrètes et de détecter les anomalies univariées et multivariées.
- De nombreuses extensions proposées permettant d'améliorer les performances de détection et de répondre aux contraintes opérationnelles telles que la réduction des fausses alarmes ou le choix du seuil de détection.

## 6. References

- [1] B. Pilastre et al., "Anomaly Detection in Mixed Telemetry Data Using a Sparse Representation and Dictionary Learning", *Signal Process.*, to appear, Mai 2019.
- [2] S. Boyd et al., "Distributed optimization and statistical learning via alternating direction method of multipliers", *Foundations and Trends in Machine Learning*, vol. 3, no. 1, pp. 1-122, Mai 2010.
- [3] B. Schölkopf et al., "Estimating the support of a high-dimensional distribution", *Neural Computation*, vol. 3, no. 7, pp. 1443-1471, 2001.
- [4] T. Yairi et al., "A Data-Driven Health Monitoring Method for Satellite Housekeeping Data Based on Probabilistic Clustering and Dimensionality Reduction", *IEEE Trans. Aerosp. Electron. Syst.*, vol. 53, no. 3, pp. 1384-1401, Juin 2017.
- [5] S. Furet et al., "Improving Spacecraft Health Monitoring with Automatic Anomaly Detection Techniques", *Proc. Int. Conf. Space Operations (SpaceOps 2016)*, Coré du sud, Mai 2016.
- [6] B. Pilastre et al., "Représentation parcimonieuse pour la détection d'anomalies dans des signaux multivariés", *Groupe d'Etude de Traitement de Signal et des Images (GRETSI)*, Lille, France, Août 2019.
- [7] "Anomaly Detection in Mixed Time-Series using a Convolutional Sparse Representation with Application to Spacecraft Health Monitoring", *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, soumis.