

Dé entrelacement des signaux radars



CentraleSupélec

AVANTIX

Objectifs & Plan

Objectifs

- ▶ Déterminer le nombre d'émetteur(s) présent(s) dans un signal
- ▶ Classifier chaque observation de ce signal par rapport au(x) émetteur(s) présent(s)

Plan

- I. Présentation des données
- II. Stratégie de dé entrelacement
 - 1. Clustering classique
 - 2. Traitement des clusters
- III. Conclusions & Perspectives

Présentation des données

- ▶ Récupération de tables au format Matlab regroupant des informations sur 2 628 signaux regroupant eux-mêmes 2 000 observations :
 - Fréquence
 - Durée d'impulsion
 - Niveau
 - dTOA
 - Identifiant du radar
- ▶ Reconstruction de la TOA à partir des données initiales
- ▶ Création d'une base de données unique regroupant l'intégralité des caractéristiques précédentes

Aperçu des données

		Index	PDW	0	1	2	3	4	} Observations
Signal	}	0	TOA	0	1387.83	2648.36	3899.45	5404.4	
		0	ID	1835	1835	1835	1835	1835	
		0	dTOA	1387.83	1260.52	1251.09	1504.94	1361.75	
		0	DI	1378.46	1382.1	1379.27	1378.53	1378.58	
		0	FN	2846.57	2819.08	3003.23	2939.18	2955.33	
		0	LVL	-161.709	-156.509	-153.712	-150.916	-146.669	
Signal	}	1	LVL	56.8632	54.0588	58.846	58.3316	59.9799	
		1	TOA	0	1653.18	3303.05	4950.42	6601.89	
		1	ID	1783	1783	1783	1783	1783	
		1	DI	1607.18	1608.18	1608.39	1610.31	1608.08	
		1	dTOA	1653.18	1649.86	1647.38	1651.46	1651.73	
		1	FN	2355.65	2355.77	2355.58	2355.96	2357.03	
		2	ID	3577	3577	3577	3577	3577	

Les données : caractéristiques des signaux

► Tailles différentes

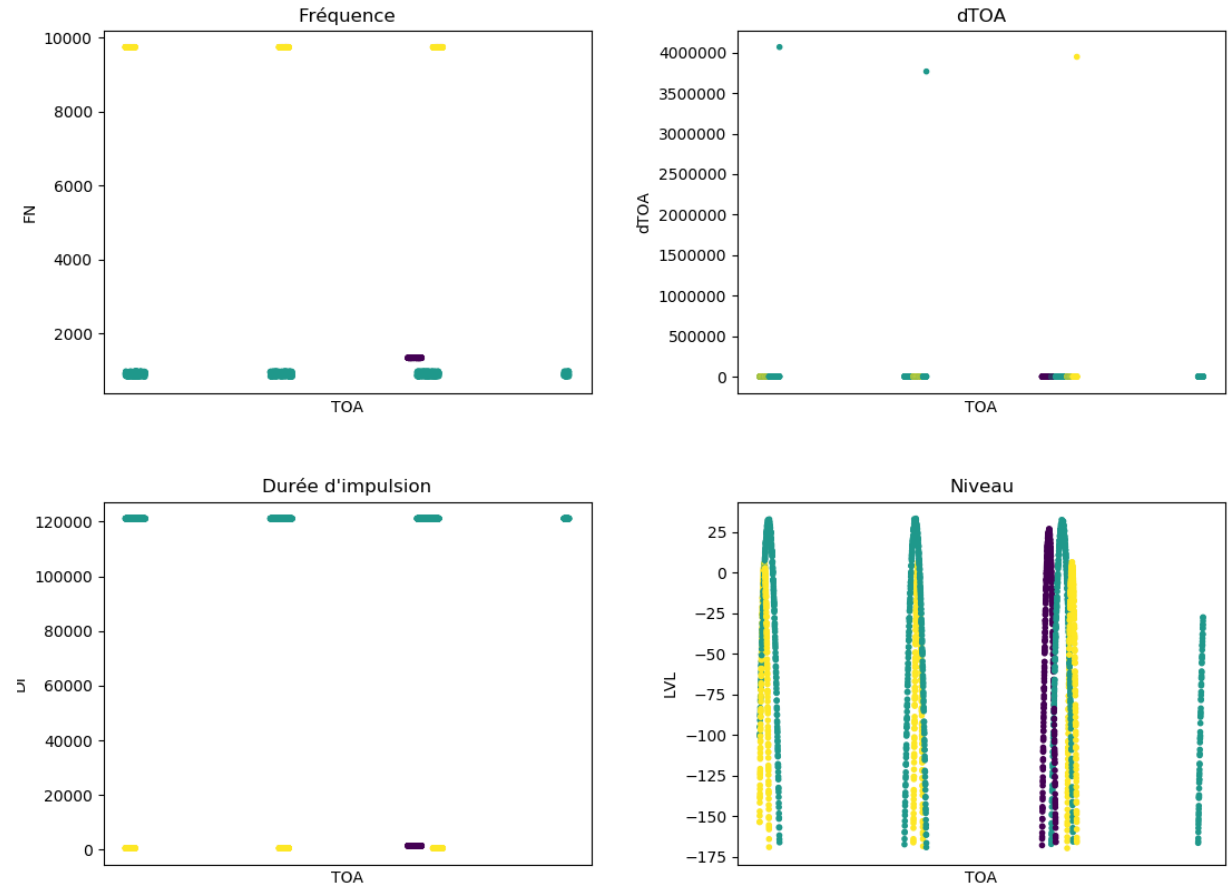
- Signal n°2 : 621 observations
- Signal n°39 : 579 observations
- Signal n°59 : 2 000 observations

► Nombre de radar différent

- 1 radar : 1 365 signaux
- 2 radars : 774 signaux
- 3 radars : 373 signaux
- 4 radars : 116 signaux

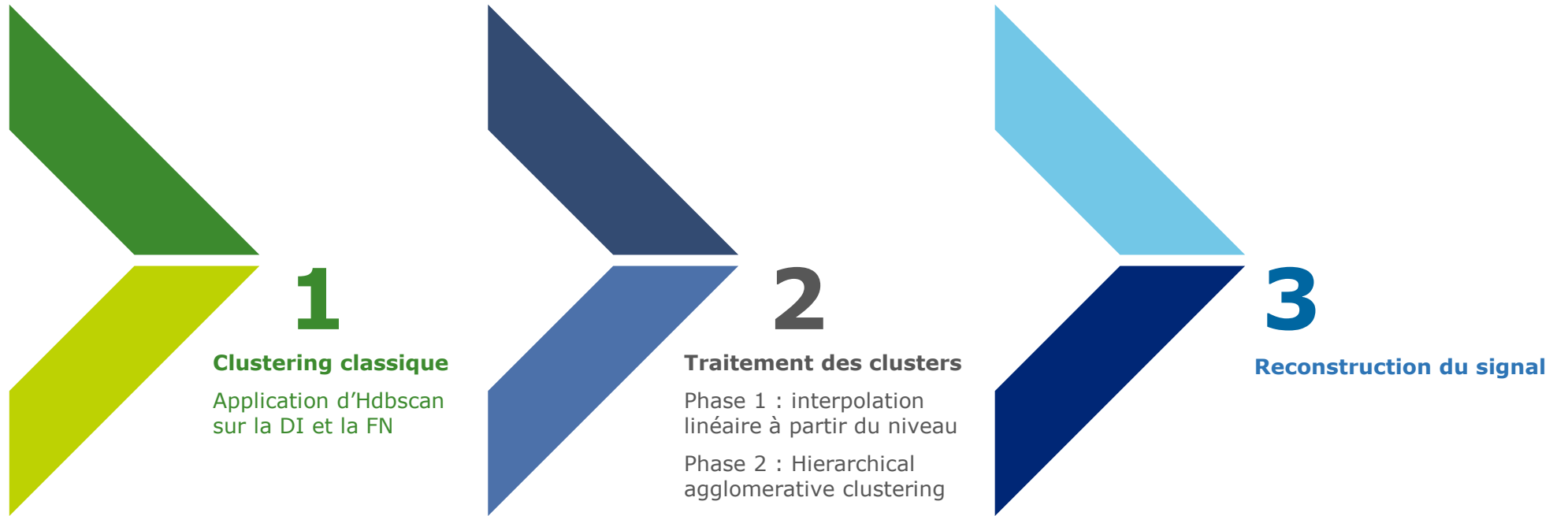
► Données bruitées

Signal n° 678 , 2000 obs et 3 émetteur(s)



Dé entrelacement des signaux

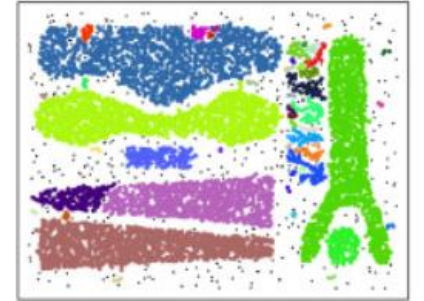
Stratégie mise en place



Clustering classique

- ▶ Pourquoi Hdbscan ?
 - Successeur de Dbscan et d'Optic, 2 algorithmes ayant prouvé leur efficacité en apprentissage non-supervisé
 - Permet de détecter n'importe quelle forme de cluster
 - Robuste aux valeurs aberrantes et permet de détecter le bruit
 - Autorise la recherche de clusters à densités variables (grosse différence par rapport à Dbscan)
- ▶ Réflexion autour du paramètre « min_cluster_size » de l'algorithme
 - Ce paramètre permet de définir la taille minimale à partir de laquelle un groupe de point peut être considéré comme un cluster et non comme du bruit
 - Une taille trop grande ne nous permettra pas de capter les radars émettant de façon irrégulière
 - Une taille trop petite nous renverra un très grand nombre de clusters
 - Fixation du paramètre à 20
- ▶ Application d'Hdbscan sur la fréquence et la durée d'impulsion
 - 2 paramètres très discriminants

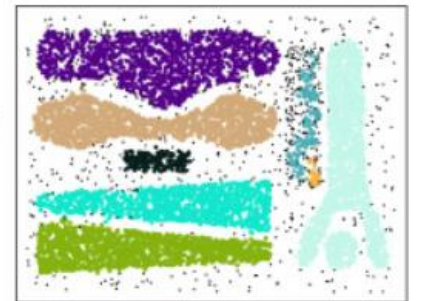
DBSCAN



OPTICS



HDBSCAN



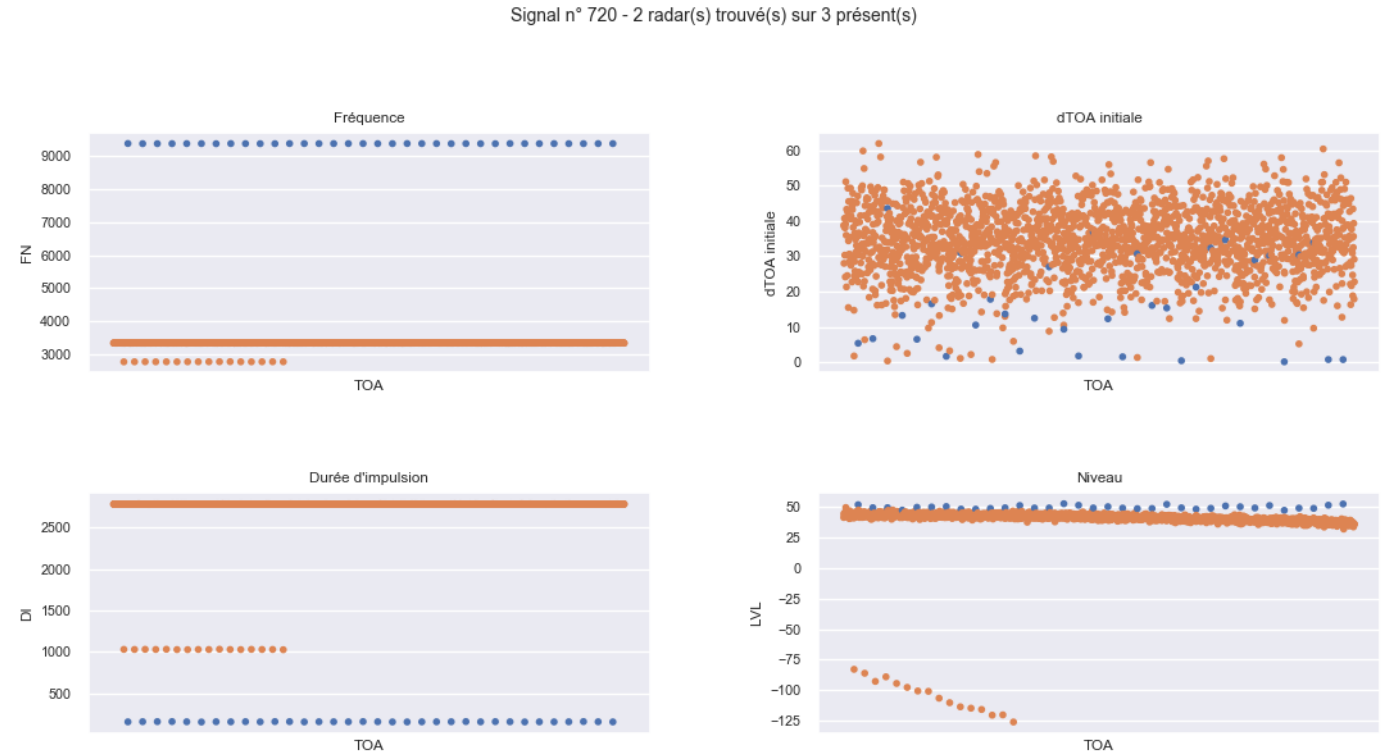
Clustering classique

- Résultats : analyse de la matrice de confusion sur la présence du nombre d'émetteur(s) par signal
 - Très bon résultats dès la première phase de clustering :
 - Classification correcte pour 98% des signaux
 - Quid sur le réalisme des données
 - Surestimation pour 0,8% des signaux du nombre d'émetteurs présents
 - Application de la phase 2 de la stratégie de dé entrelacement

		Valeurs prédites								Total
		1	2	3	4	5	6	7	9	
Valeurs réelles	1	1361	4	0	0	0	0	0	0	1365
	2	12	754	4	1	1	1	1	0	774
	3	0	14	354	2	0	1	2	0	373
	4	0	1	4	106	2	1	1	1	116

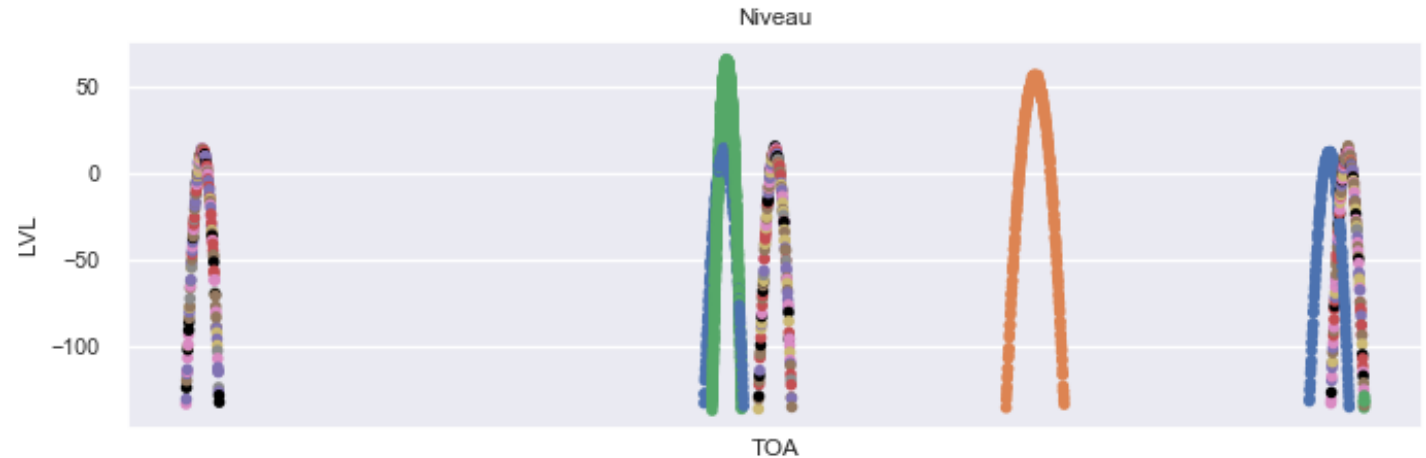
Clustering classique

- Résultats à manipuler avec précaution
 - Identification d'un grand nombre d'outliers lorsque le signal représente 1 seul radar
 - Sous-estimation pour 1,2% des signaux du nombre d'émetteurs présents
 - Peu d'observations pour le radar (min_cluster_size est trop grand)
 - Cas « atypiques »



Traitement des clusters

- ▶ **Objectif** : effectuer des regroupements entre les clusters en comparant leurs informations relatives
- ▶ **Méthode** : stratégie en 2 étapes
 - Interpolation
 - Hierarchical agglomerative clustering
- ▶ **Inconvénient** :
 - Les clusters doivent être uniformément répartis



Traitement des clusters

Phase 1 : interpolation

► Phase 1 : Interpolation à partir du niveau

- Interpoler les valeurs d'un cluster de référence dans le plan Niveau x TOA afin d'approximer une fonction du type

$$f : y = f(x)$$

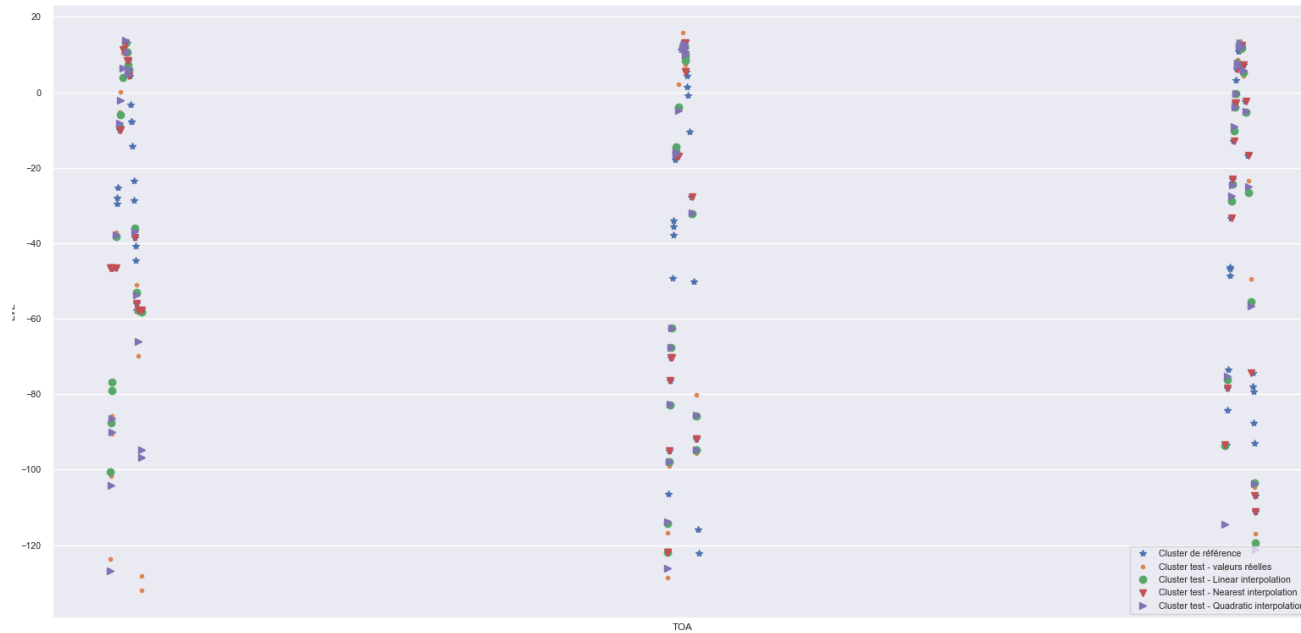
avec y : Niveau et x : TOA

- Estimer à l'aide de cette fonction les valeurs du niveau des autres clusters
- Calculer une matrice de distance afin de comparer les valeurs estimées du niveau avec les valeurs réelles à l'aide de la distance euclidienne

Traitement des clusters

Phase 1 : interpolation

Signal n° 277



Signal n° 277



Traitement des clusters

Phase 2 : HAC

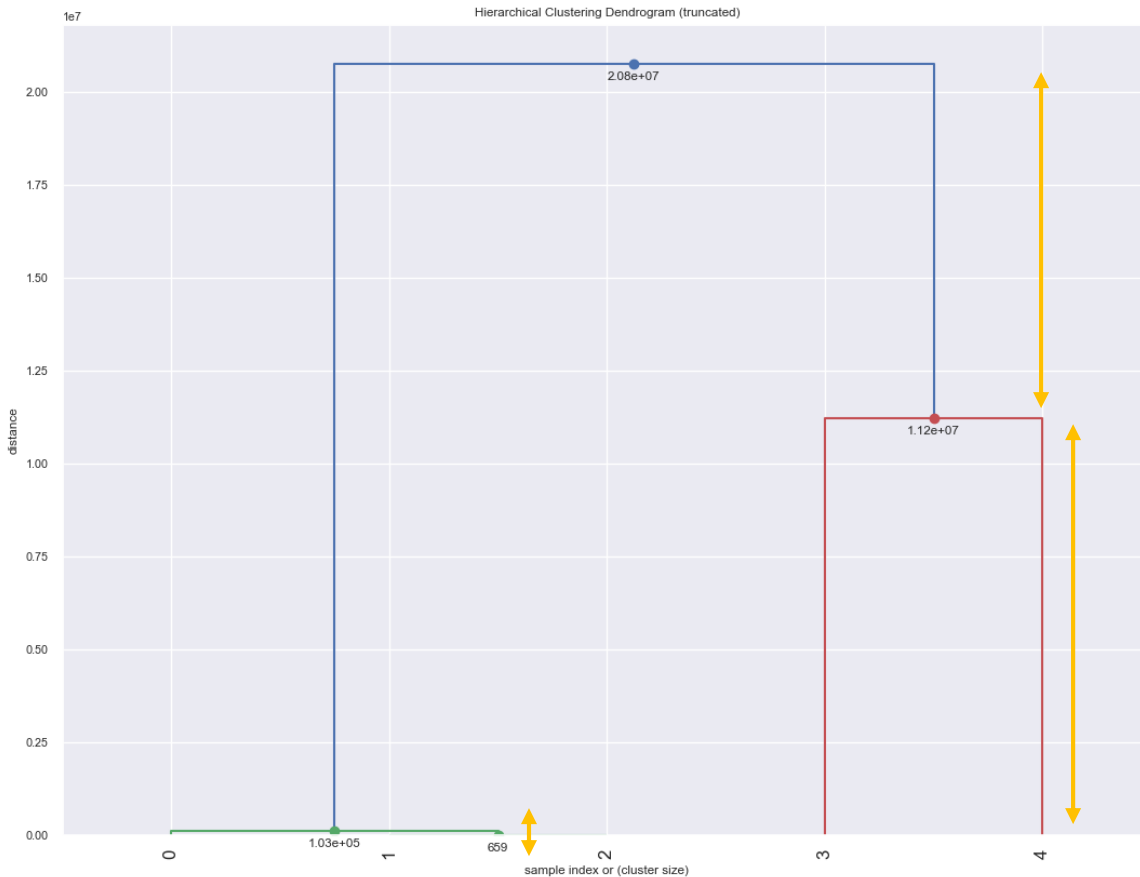
► Phase 2 : Hierarchical agglomerative clustering

- Chaque observation est représentée par un cluster
- A chaque itération les 2 groupes les plus proches sont fusionnés en un seul cluster (ceux ayant la plus petite mesure de dissimilarité)
- Mise à jour de la matrice de distance
- Répétition des étapes précédentes jusqu'à obtention d'un cluster unique

Traitement des clusters

Phase 2 : HAC

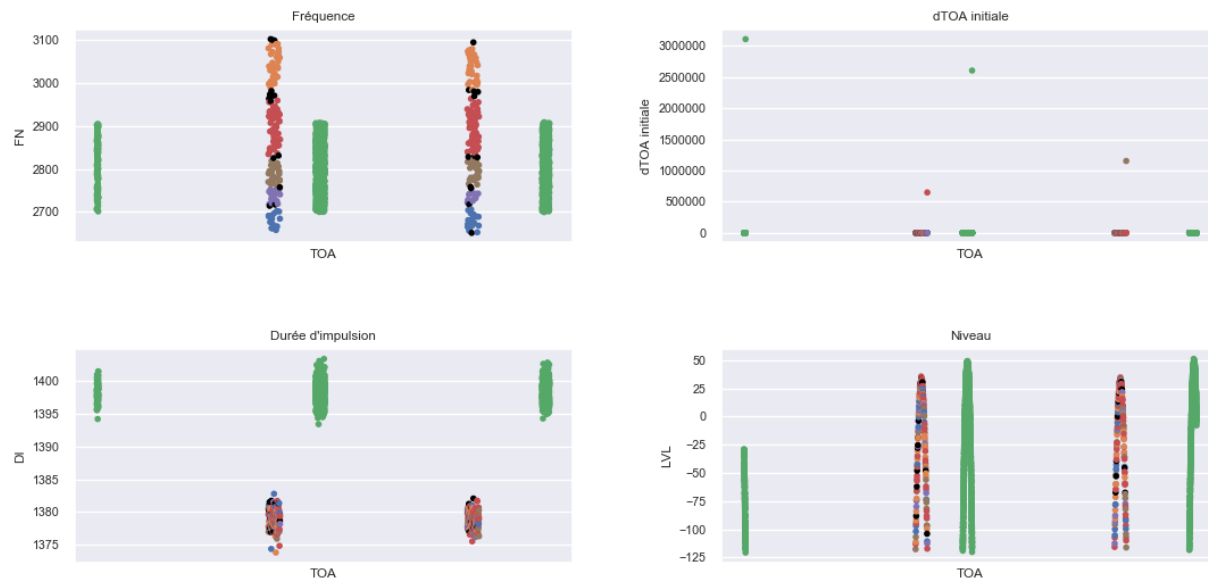
Signal n° 1095 - 5 radar(s) trouvé(s) sur 4 présent(s)



- **Dendrogramme**
 - Première analyse
 - Décrit étape par étape les groupes fusionnés
- Détermination du seuil pour « couper » le dendrogramme
 - Nombre de cluster à retenir ?
 - Analyse de l'écart entre 2 niveaux d'agrégation successifs qui changerait la structure des données lors du regroupement
 - Regroupement des clusters 0, 1 et de la classe des outliers

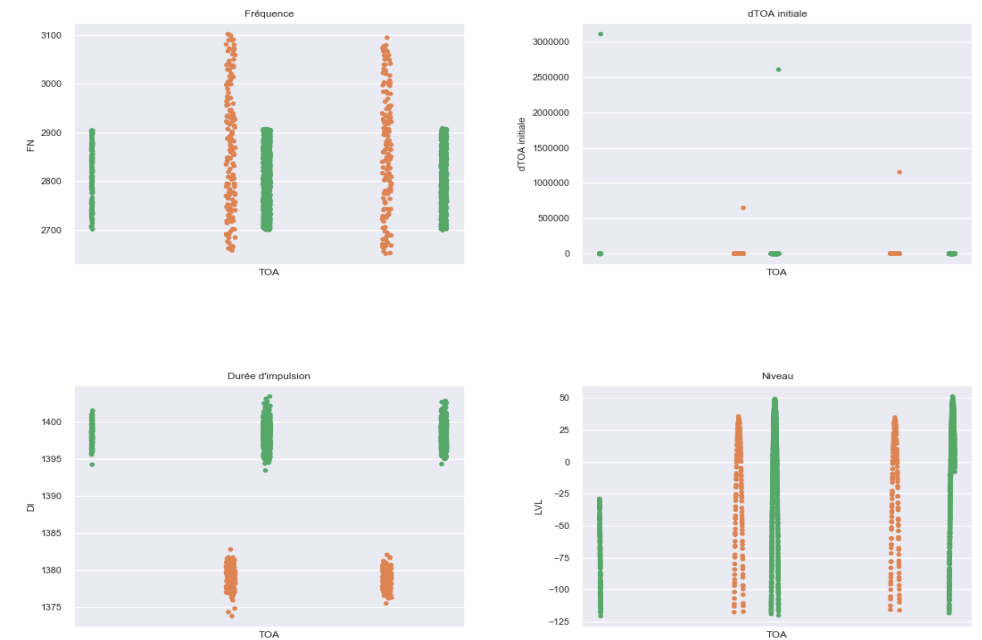
Traitement des clusters

Signal n° 40 - 6 radar(s) trouvé(s) sur 2 présent(s)



Phase 1 : clustering classique

Signal n° 40 - 2 radar(s) trouvé(s) sur 2 présent(s)



Phase 2 : Traitement des clusters

Conclusions & Perspectives

- ▶ Résultats de la stratégie:
 - Fonctionne très bien lorsque le nombre de clusters est surestimé
 - Tendance à regrouper des clusters
 - Méthodologie valide sur des cas classiques
 - Quid de la nature des données

- ▶ Perspectives :
 - Utilisation de la continuité du niveau pour valider les clusters :
 - Traitement des « mal classés » ie des siganux dont le nombre d'émetteurs est sous-estimé
 - Optimisation du seuil pour regrouper les clusters
 - Définition de règles pour déterminer s'il est nécessaire d'appliquer la phase 2 de notre stratégie

**Thank you for your
attention**

An abstract graphic of a network or mesh structure, composed of numerous small, interconnected nodes and lines, rendered in a light blue color against a darker blue background. The network appears to be a complex, interconnected web of points and lines, possibly representing a data structure or a communication network. The lines are thin and light blue, while the nodes are small, slightly glowing blue dots. The overall effect is a sense of depth and connectivity, with the network extending across the frame from the bottom left towards the top right.