

Qui suis-je?

Clémentine Barreyre @airbus.com

- Ingénieure Génie Mathématiques et Modélisation INSA Toulouse, Diplômée en 2014.
- Stage de fin d'études réalisé à Airbus Defence and Space en 2014
- Thèse CIFRE « Statistique en grande dimension pour la détection d'anomalies dans les données fonctionnelles issues des satellites », réalisée à Airbus Defence and Space de février 2015 à février 2018 – thèse soutenue en mai 2018 – en partenariat avec l'institut de Mathématiques de Toulouse (IMT)
- Depuis 2018, responsable de l'atelier de science des données au sein de la future plateforme Big Data d'Airbus Defence and Space.



SOMMAIRE

1. INTRODUCTION

- 1. Quels satellites? Pour quelles applications?
- 2. Quelles sont les grandes étapes de la vie des satellites ?
- 3. Historique de la data science à Airbus Defence and Space.

2. LA SCIENCE DES DONNEES AU SERVICE DE LA CONCEPTION, DES TESTS ET DES OPERATIONS REALISES DES SATELLITES

- 1. Un exemple d'application : validation de la structure des guides d'onde.
- 2. Un exemple d'application : validation automatique des tests réalisés sur les satellites
- 3. Un exemple d'application : surveillance automatique et détection de pannes à bord des satellites
- 4. Un exemple d'application : Analyse des images

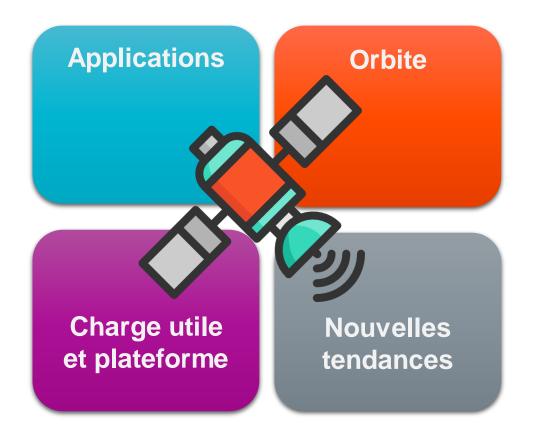
3. MISE EN ŒUVRE DE LA DATA SCIENCE CHEZ AIRBUS DEFENCE AND SPACE



Qu'est-ce qu'un satellite

Pour quelles applications?

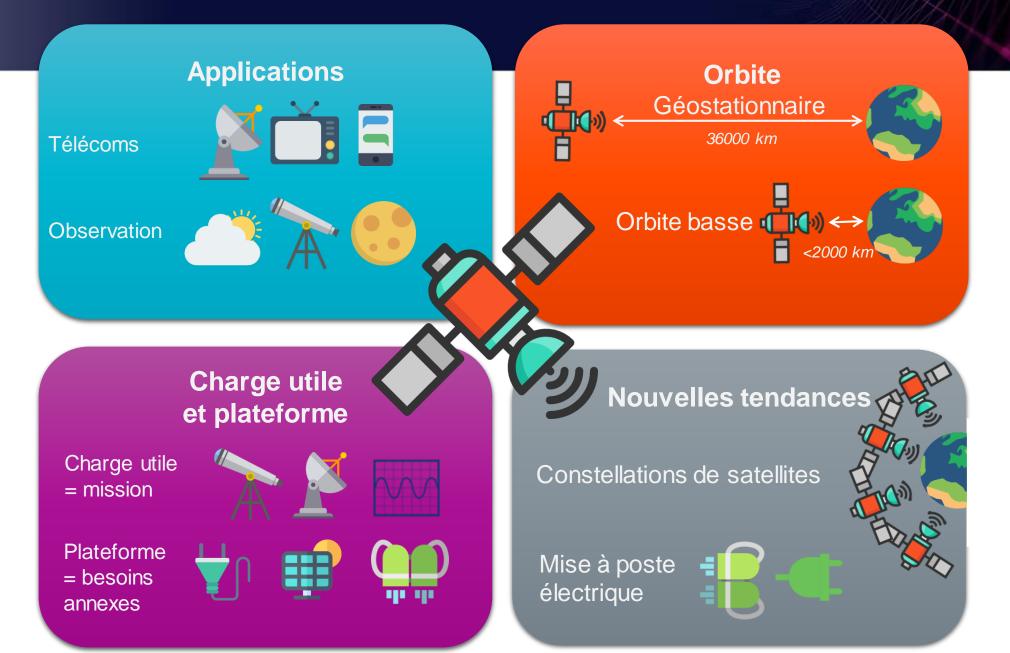
De quoi est-il fait



Sur quelle orbite évoluent-ils ?

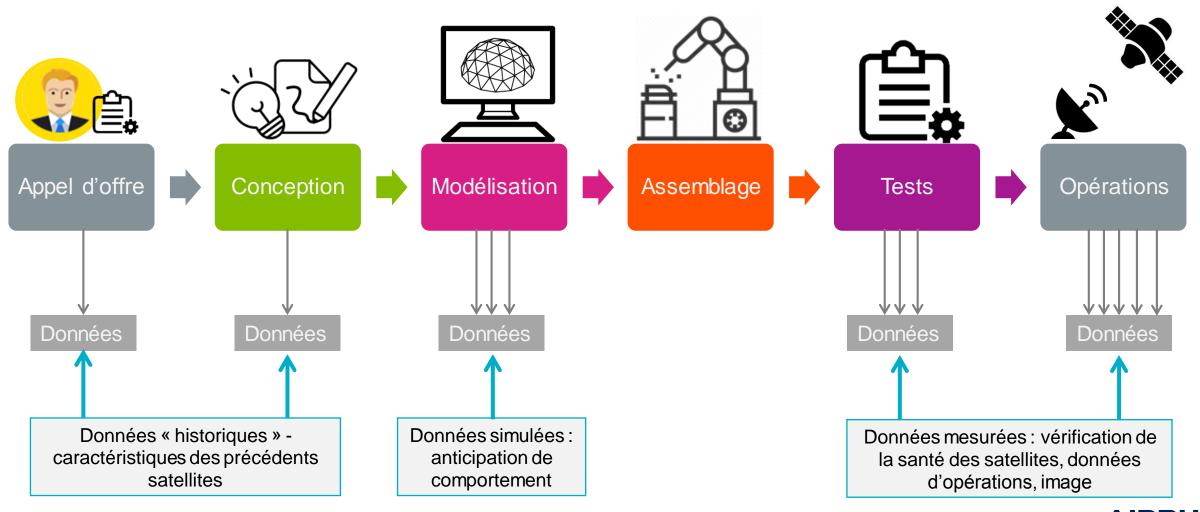
Quel est le futur des satellites ?







Cycle de vie d'un satellite





La Data Science à Airbus Defence and Space

***** 2013 :

- Un logiciel par phase de vie du satellite, stockage des données en XML, CSV, archives. Un serveur par type de données.
- Suivi « automatique » des tests et des opérations réalisé à partir de règles simples : dépassements de seuils.

3 2014 :

- Première initiative autour de l'analyse des données sur les données de tests satellites : validation des satellites après les tests.
- Etude parallèle sur la **dégradation des batteries en orbite** (étude réalisée avec le CNES).

***** 2015 :

- Thèse sur la détection d'anomalies à bord des satellites
- Début de l'analyse des images : détection automatique des nuages sur les images. Utilisation de la Google Cloud Plateform.

2016 :

- Création du **Datalab** à Airbus Defence and Space, recensant les initiatives autour de la donnée.
- Premiers cas d'applications aux segments sols : suivi de performance de traitement...

***** 2017 :

- Explosion du nombre de data scientists à Airbus DS.
- **SKYWISE** chez Airbus Commercial plateforme big data pour les avions, gérée par Palantir.

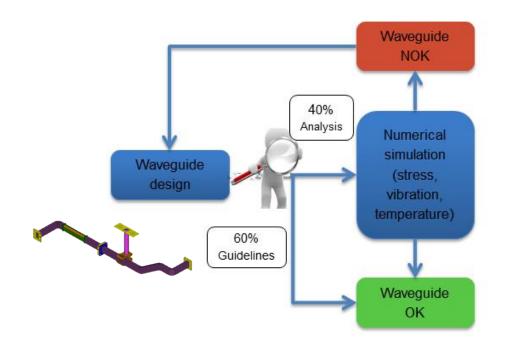
2018:

- Création d'une entité « Analytics » transverse, afin de supporter les initiatives autour de l'IA.
- Début du projet OASIS, destiné à créer la plateforme big data pour les données satellites.



Validation de la structure des guides d'ondes

- Entrainé sur 8 satellites de telecommunications, sur une base de données comprenant 5390 géométries de guides d'ondes.
- Aide à la decision se basant sur l'historique de la géométrie de ces guides d'ondes.
- Estimer la conformité du guide d'onde par rapport aux spécifications de stress et thermo-élastiques
- Eviter des simulations inutiles et coûteuses
- Appliqué et testé sur une archive de données
- Algorithme déployé et opérationnel
- Minimise le taux de fausses détections



Analyse de sensibilité

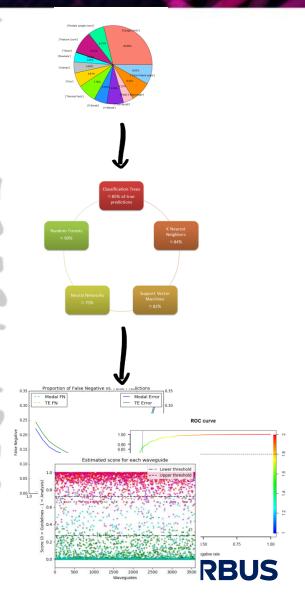
Identifier les facteurs les plus influents. Enrichir les données avec d'autres paramètres

Benchmark prediction methods

Arbres de classifications, K plus proches voisins, Support Vector Machines, Réseaux de neurones, forêts aléatoires.

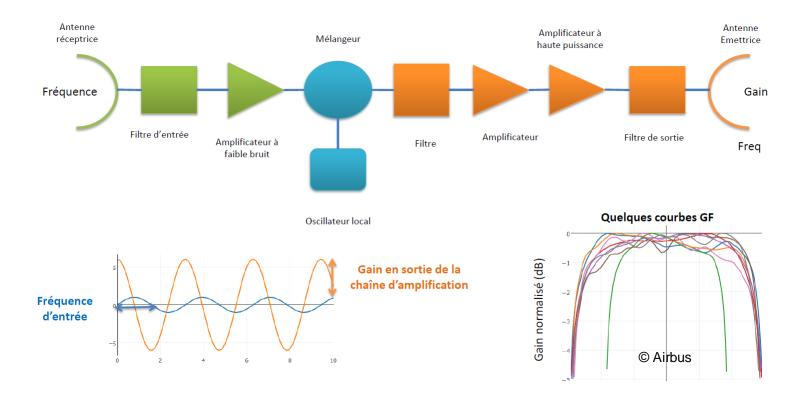
Amélioration du modèle

Optimisation des hyperparamètres, validation, contrôle du taux de faux négatifs, indégration de deux seuils (5% à valider manuellement)



Validation automatique des tests : Application au test Gain-Fréquence

- Les tests Gain-Fréquence représentent une des plus grandes source de données des tests réalisés sur les satellites de télécommunications.
- On mesure la performance de la chaîne d'amplifications en mesurant le gain à la sortie d'un signal envoyé en entrée, pour une certaine fréquence.
- Objectif: valider automatiquement les tests normaux identifier les tests pouvant avoir subi des dégradations





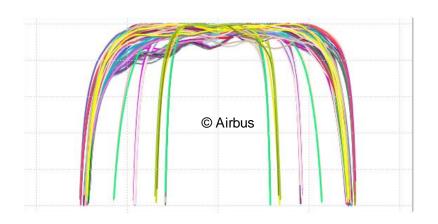
Validation automatique des tests : Application au test Gain-Fréquence

Lecture des données : résultats des tests

Préparation / normalisation des données

Réduction de dimension:
Représentation des 1000 signaux en deux dimensions, résumé de l'information

Application d'une méthode de détection d'événements atypiques : One-Class SVM – 90% des tests validés





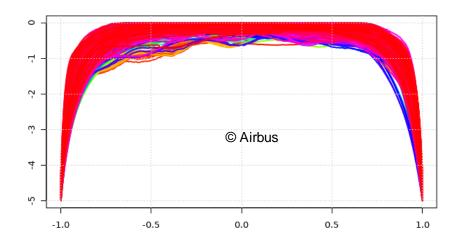
Validation automatique des tests : Application au test Gain-Fréquence

Lecture des données : résultats des tests

Préparation / normalisation des données

Réduction de dimension:
Représentation des 1000 signaux en deux dimensions, résumé de l'information

Application d'une méthode de détection d'événements atypiques :
One-Class SVM – 90% des tests validés





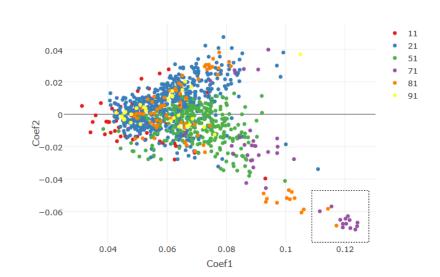
Validation automatique des tests : Application au test Gain-Fréquence

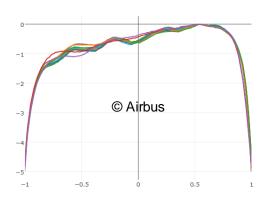
Lecture des données : résultats des tests

Préparation / normalisation des données

Réduction de dimension:
Représentation des 1000 signaux en deux dimensions, résumé de l'information

Application d'une méthode de détection d'événements atypiques :
One-Class SVM – 90% des tests validés







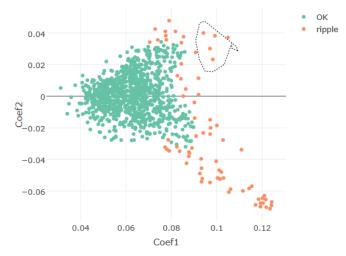
Validation automatique des tests : Application au test Gain-Fréquence

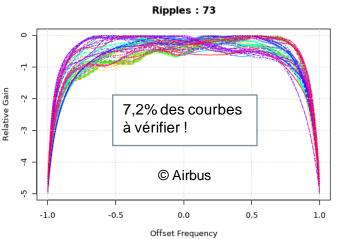
Lecture des données : résultats des tests

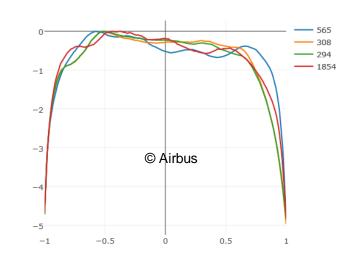
Préparation / normalisation des données

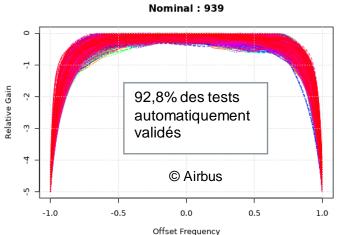
Réduction de dimension:
Représentation des 1000 signaux en deux dimensions, résumé de l'information

Application d'une méthode de détection d'événements atypiques : One-Class SVM – 90% des tests validés











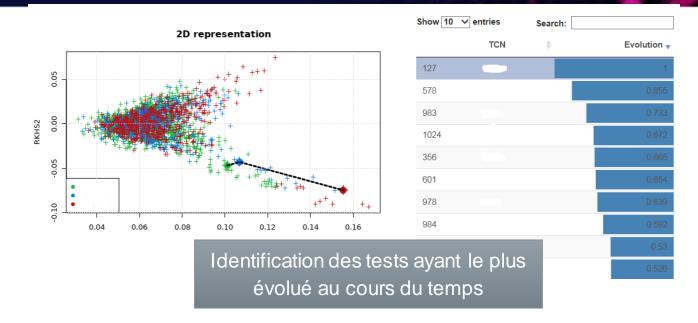
Validation automatique des tests : Application au test Gain-Fréquence

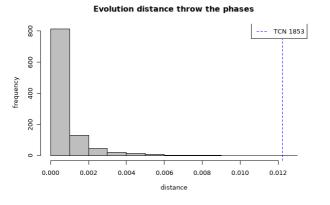
Lecture des données : résultats des tests

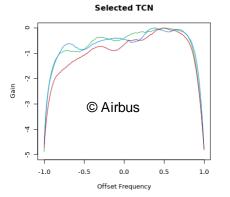
Préparation / normalisation des données

Réduction de dimension:
Représentation des 1000 signaux en deux dimensions, résumé de l'information

Application d'une méthode de détection d'événements atypiques : One-Class SVM – 90% des tests validés

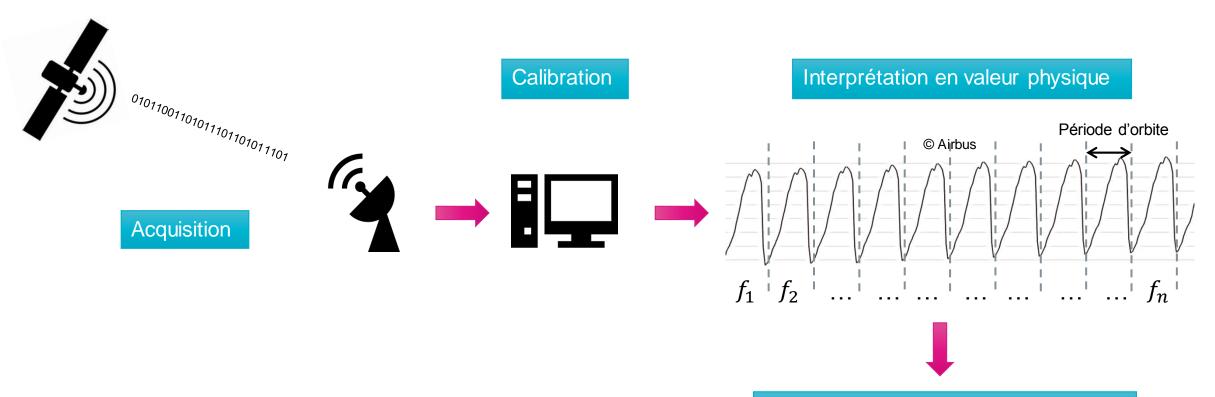








Surveillance des satellites – détection de pannes

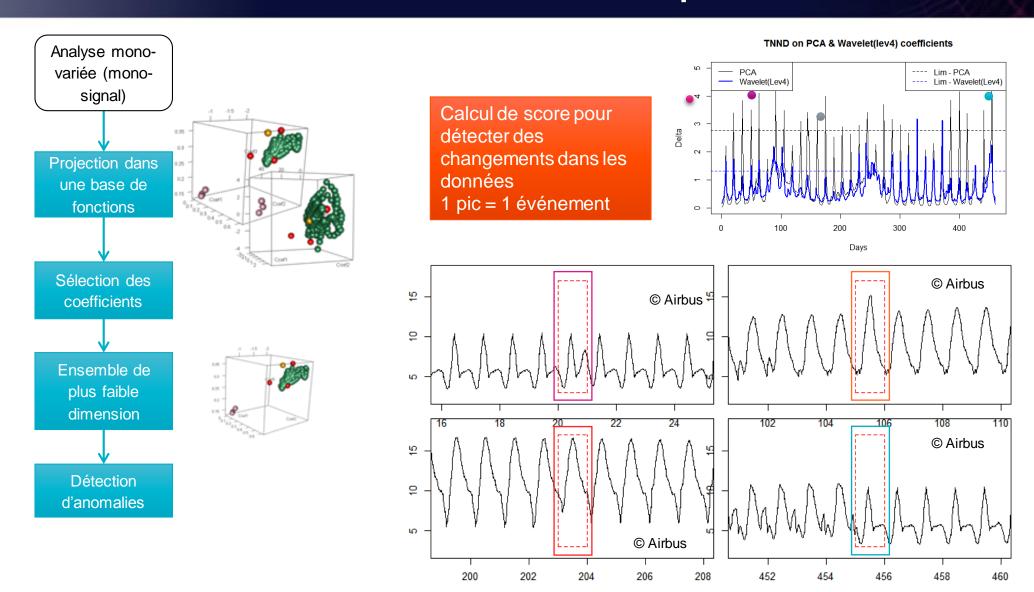


 But : détecter automatiquement des changements de comportements, pouvant expliquer des dégradations d'équipements, des pannes. Détection d'anomalies sur les courbes:

$$f_i$$
, $i = 1, \dots, n$



Surveillance des satellites – détection de pannes





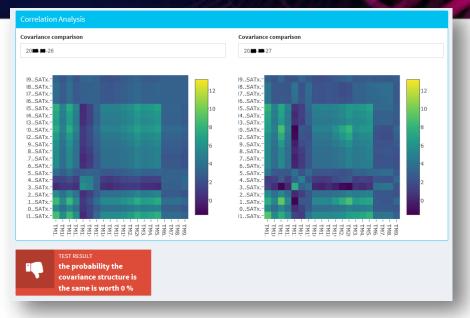
Surveillance des satellites – détection de pannes

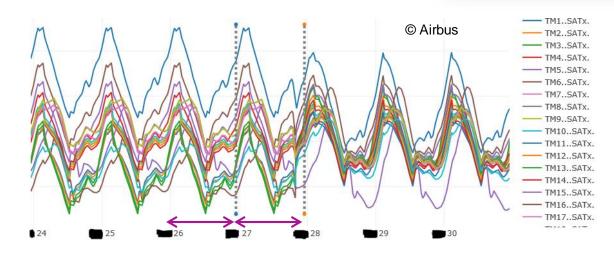
Analyse multivariée (multi-signaux)

Calcul de la matrice de covariance entre les signaux

Test statistique : comparaison des matrices de covariances

Test rejeté : événement détecté Idée: comparer les covariances dans un groupe de télémesures pour deux orbites consécutives





L'analyse multivariée permet de retrouver les événements majeurs à bord de



Analyse d'images

Objectif:

Offrir des services en plus des images livrées à nos clients :

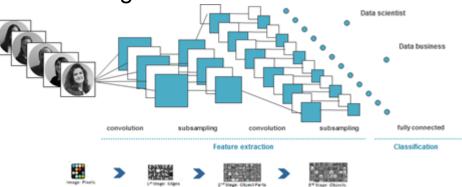
- Détection automatique des nuages sur les images gage de qualité de l'image.
- Détection de changements entre deux passes au dessus d'un même lieu.
- Génération de carte automatique à partir d'une image.
- Comptage d'avions, voitures dans un aéroport/parking...

Comment?

En utilisant des méthodes de type deep learning / réseaux de neurones convolutionnels

• Très performants pour analyser les contenus des images

• Echantillons de données labellisées





Analyse d'images



Erreur < 3%!

Les résultats sont souvent meilleurs que l'œil humain!

Source: https://www.youtube.com/watch?v=xasA9EUop_w https://hal.archives-ouvertes.fr/hal-01783857

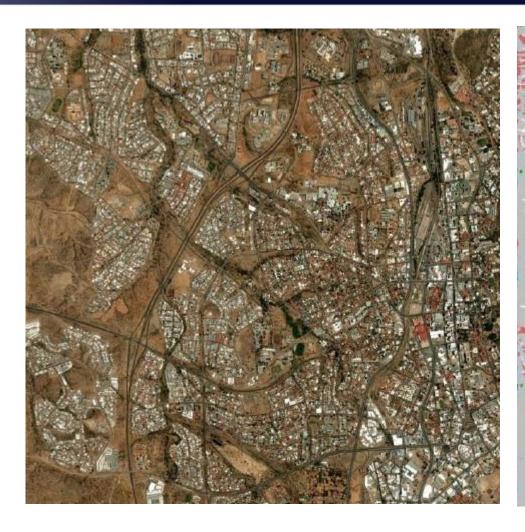


Enjeux & ApplicationsAnalyse d'images





Analyse d'images





La Data Science à Airbus Defence and Space Airbus Confidential

Une approche de type Big Data







Un atelier d'analytics complet et facile à utiliser

Outils pour la data science



















Des bonnes pratiques pour capitaliser du code et de la connaissance











Librairies collaboratives

De l'exploration de données au déploiement d'algorithmes

Exploration de données



Entrainement et validations de modèles



Déploiement et orchestration d'algorithmes

Un datalake centralisé et sécurisé

Télémesures et télécommandes en temps réel





Ensemble des données du satellite



Données marketing et données de design

Pour créer des analyses à valeur ajoutée

- Limiter les impacts des pannes
- Apprentissage tout au long de la vie du satellite : comportement et tendances
- **Optimisation:** Améliorer la conception et l'utilisation des satellites
- Etendre le périmètre: Corrélation multi satellites et corrélations



