

# Détection et caractérisation d'un événement anormal par l'analyse des données d'usage

*Application à l'hélicoptère*

**Présentation FIMA**



thinking without limits

**Pierre Bect**

Zineb Simeu-Abazi

Pierre-Loic Maisonneuve

# Déroulement de la présentation

- Contexte industriel
- Présentation générale de l'approche
- Construction d'un modèle de comportement normal
- Exploitation du modèle dans un objectif de détection
- Perspectives de développement

# Contexte industriel

## — La maintenance à Eurocopter

- Segmenter par sous-systèmes
  - Hélicoptère = système de systèmes
  - Compétences variées
- Politiques de maintenance très différentes selon les sous-systèmes
  - Guidées par :
    - La sécurité et les autorités de certifications
    - Les contrats de maintenance du client

## — Besoins

- Centraliser les informations de maintenance
- Connaître l'état de santé global de la machine
- Identifier un comportement déviant de l'appareil

# Contexte industriel

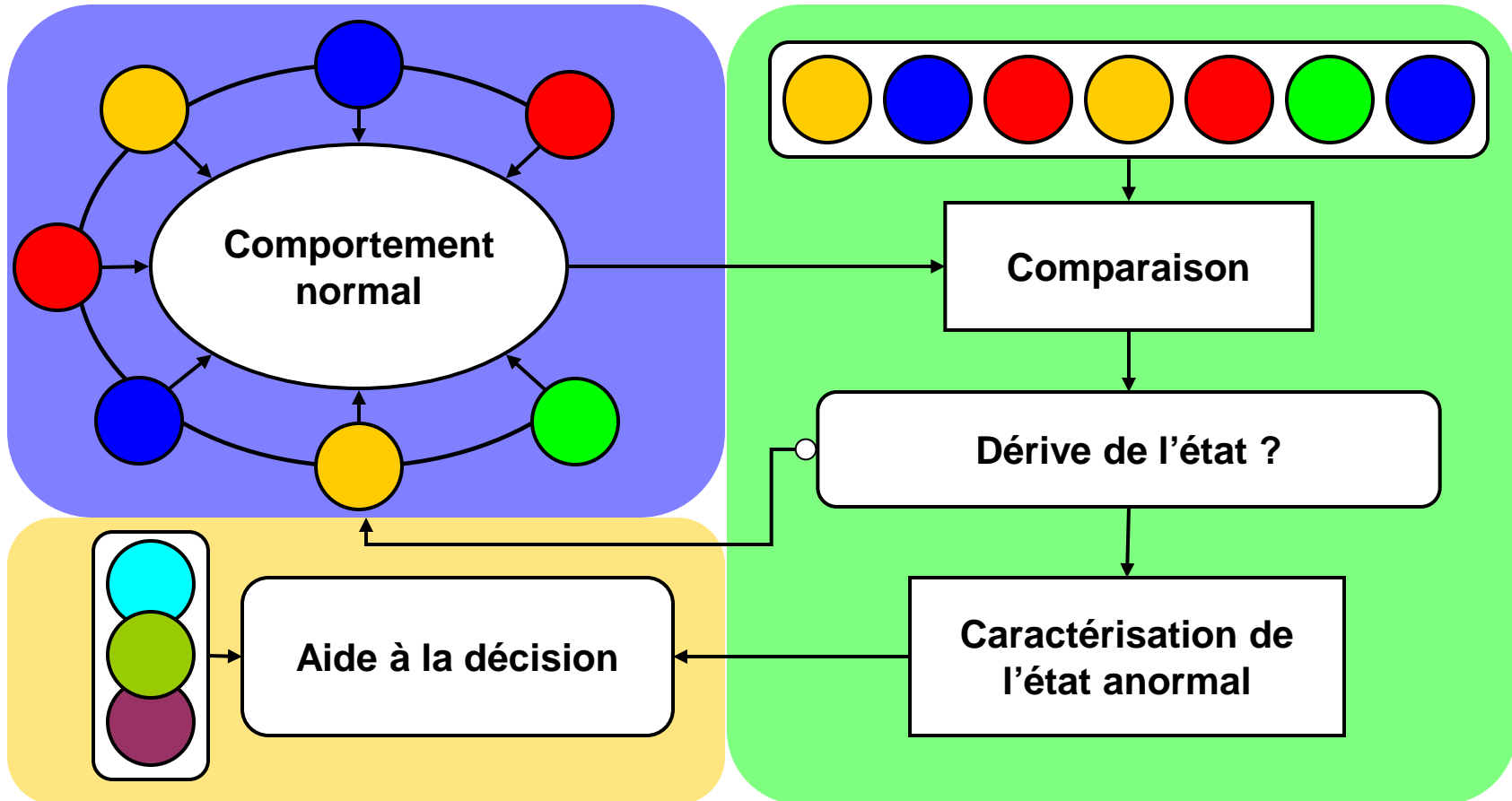
## — Opportunités

- Mise en place d'une infrastructure de traitements performantes
- Retour d'expérience
  - Volumétrie importante (jusqu'à 5 ans de donnée sur certains appareils)
  - Diversité des données
  - Expertise des ensembles dynamiques
- Méthode de Data Mining

## — Problématique dégagée

- **Comment extraire d'un historique de données, un modèle représentatif du comportement normal du système ?**
- **Comment détecter et caractériser un phénomène anormal à partir d'un modèle du comportement normal du système ?**

# Présentation générale de l'approche



# Présentation générale de l'approche

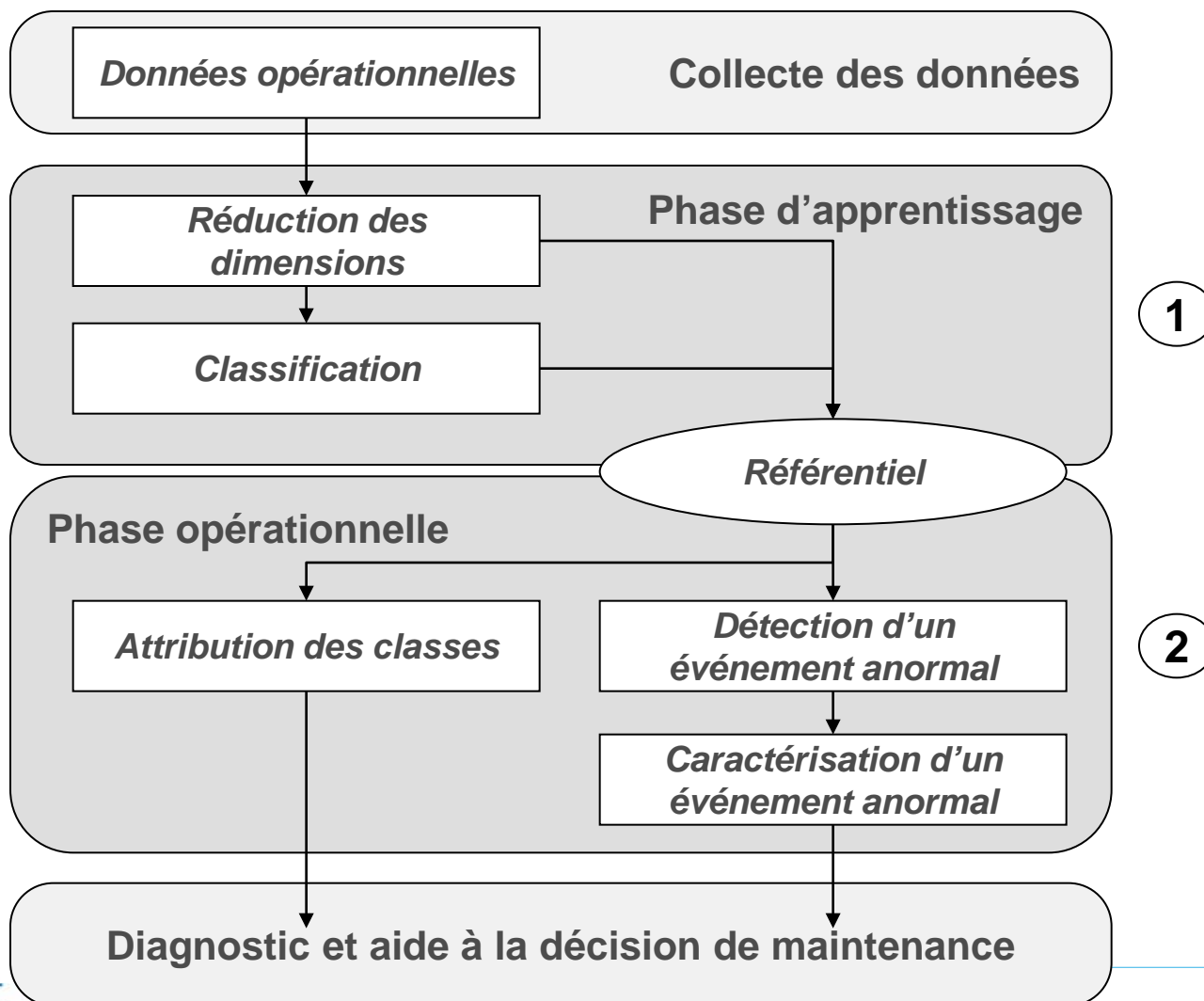
## — Les données d'usage :

- Echantillonnées à 2Hz à partir de l'allumage d'un des deux moteurs
- Informations enregistrées :
  - Positions de commande
  - Positions de l'appareil dans l'espace
  - Vitesses et accélérations
  - Données moteurs (couples, vitesses, consommation, températures)
  - Données environnementales (température, pression, vent)
  - Paramètre initiaux de la machine (poids, quantités de carburant)

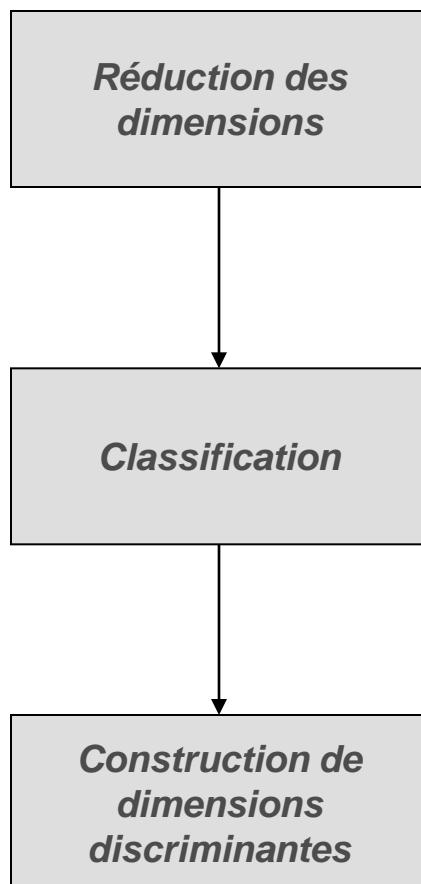
## — Beaucoup de comportements différents sur un hélicoptère

- Découpage des données en phases de vol (24000 phases de vols regroupées en 13 familles)
- 2 familles de phases de vols récurrentes :
  - Vol en palier (un tiers du temps) -> régime très stable
  - Phase de vol au sol (un peu moins du tiers du temps) -> régime variable

# Présentation générale de l'approche



# Construction d'un modèle de comportement normal

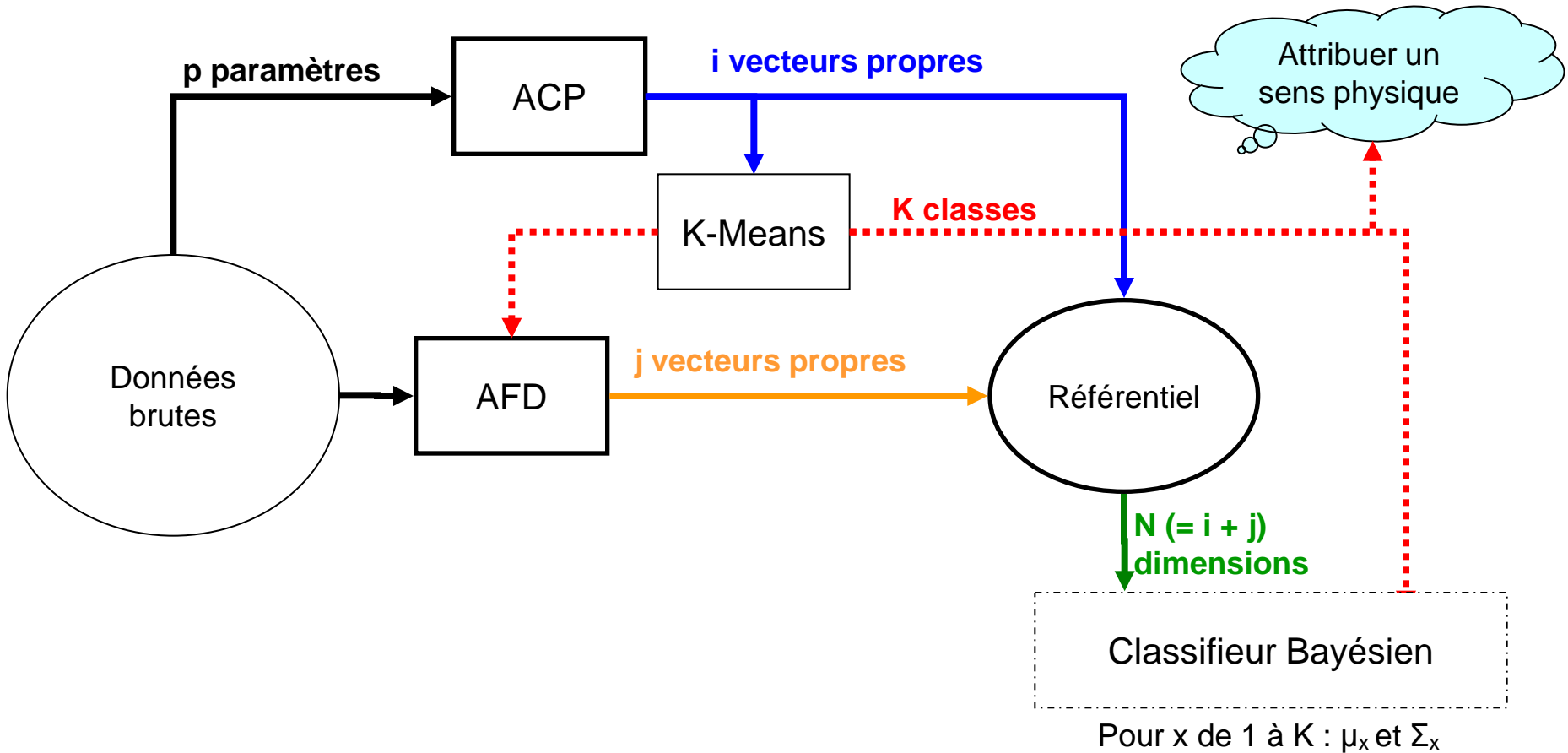


- **Objectif** : Construire un espace de dimensions réduites indépendant des vols.
- **Outils** : analyse en composantes principales (ACP)
- **Particularités** : les vecteurs propres sont construits sur l'ensemble de la base d'apprentissage par le biais d'un traitement vol à vol.
  
- **Objectif** : Mettre en avant les particularités des projections (phénomène de tâches), identification des différents nuages de points.
- **Outils** : clustering K-mean
- **Particularités** : le clustering est effectué sur la totalité des données de la base d'apprentissage. Le traitement ne se fait pas par vol mais sur l'ensemble des données.
  
- **Objectif** : Déterminer quels sont les dimensions prépondérantes pour segmenter les nuages de points.
- **Outils** : analyse factorielle discriminante (AFD).



# Construction d'un modèle de comportement normal

Apprentissage données « sans défaillance »



# Construction d'un modèle de comportement normal

## — Analyse en composantes principales

- Chaque vol est indépendant et possède ses propres caractéristiques
- ACP effectuée vol à vol
  - Nécessité d'avoir un référentiel commun, donc des vecteurs de projections communs
- On propose de mitiger les vecteurs de projections
  - Identifier les vecteurs de projections semblables
    - On utilise la méthode des moindres carrés pour identifier la proximité entre les vecteurs de projections.
    - Les vecteurs les plus proches seront associés selon la formule ci-dessous.
  - Faire intervenir des notions de temps

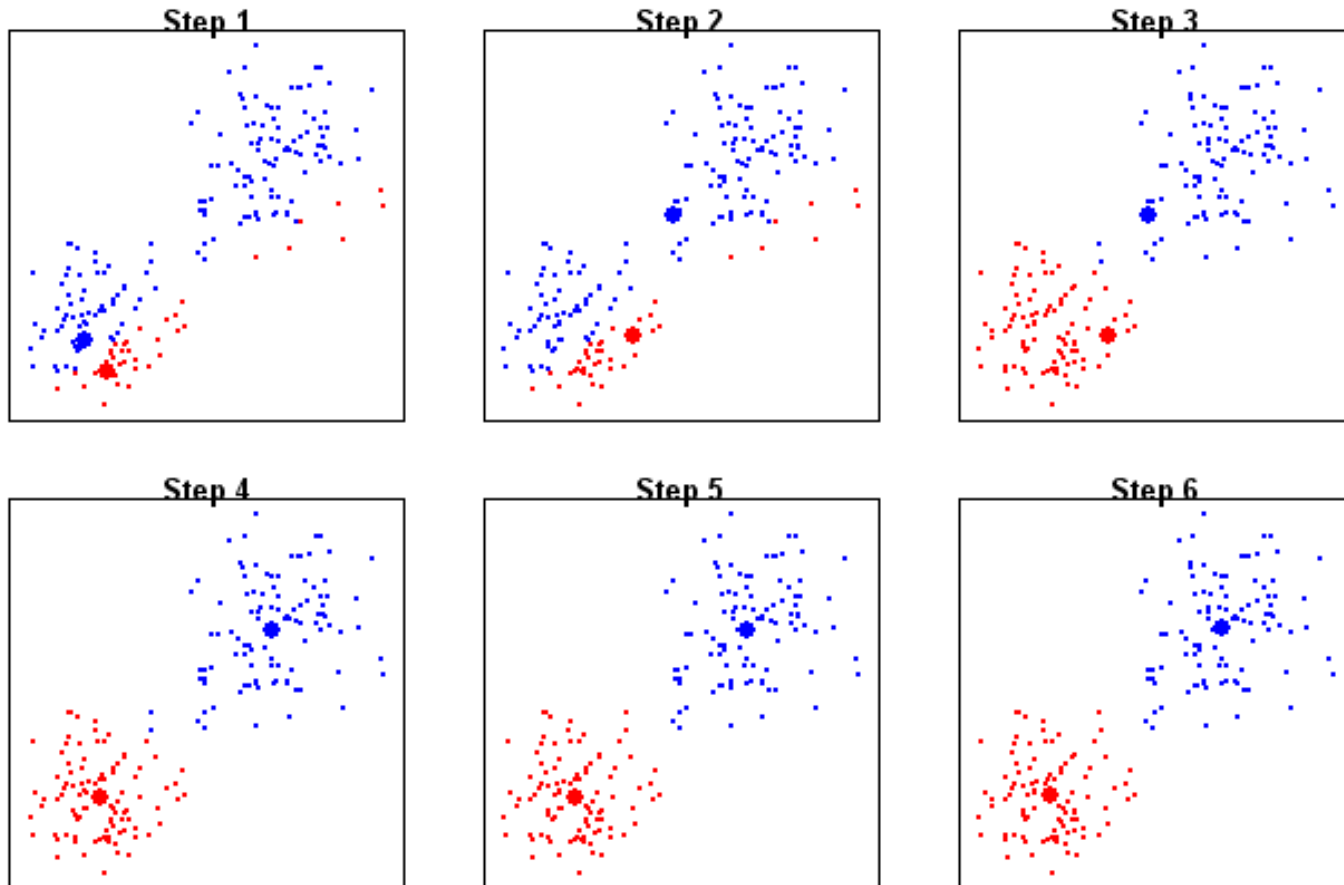
$$PCA_{gen} = \frac{\sum_{x=1}^n PCA_x \cdot Nfra_x}{\sum_{x=1}^n Nfra_x}$$

# Construction d'un modèle de comportement normal

## Clustering K-Means

- Chaque classe n'apparaissant pas forcément dans chaque vol, la classification doit s'effectuer sur l'ensemble de la base de projection.
- Il faut déterminer le nombre optimal de classes à identifier.
  - Plusieurs méthodes existent, mais nous avons utilisé le critère du MML (Minimum Message Length).
    - Fixer un nombre maximal de clusters (suffisamment grand pour ne pas rater le nombre optimal, mais pas trop non plus... le but est de faire de regrouper)
    - L'algorithme K-Means est lancé plusieurs fois et conserve le K qui conduit à une minimisation du critère.
- On place ensuite les centres initiaux de chaque cluster
- Phase de convergence...
- Critère d'arrêt

# Construction d'un modèle de comportement normal



# Construction d'un modèle de comportement normal

## — Analyse Factorielle Discriminante

- Du fait que chaque classe n'apparaissant pas forcément dans chaque vol, l'AFD se doit être implémenter sur l'ensemble de la base d'apprentissage.
- Très similaire à l'ACP sauf l'AFD cherche à maximiser l'écart entre les classes.
- Nécessite la construction de matrice de variance covariance intra et inter classe
  - L'objectif étant de favoriser la variance inter classe au profit de la variance intra classe.
- L'analyse des paramètres prépondérant dans chacun de ces axes va ainsi nous permettre de connaître qu'elles sont les facteurs discriminants entre les classes.

# Construction d'un modèle de comportement normal

De là, on dispose :

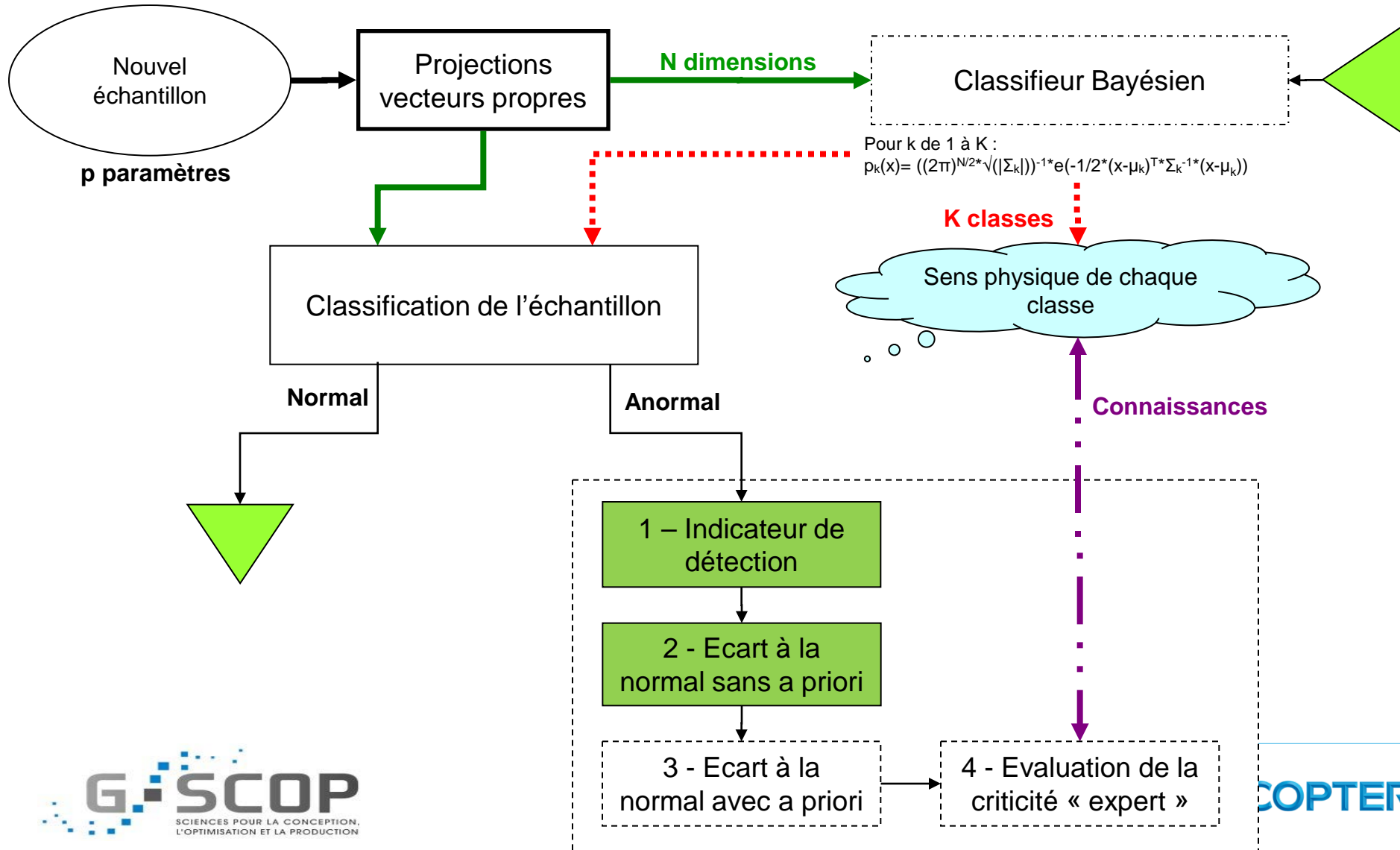
- D'un repère de projection commun à l'ensemble des vols.
- D'une décomposition de l'information en regroupement représentatif d'un comportement particulier.
- D'une quantité de projections représentative du comportement normal de l'appareil dans 2 phases de vol très différentes.

# Exploitation du modèle dans un objectif de détection

## Objectifs

- Identifier l'apparition d'une donnée anormale
- Identifier l'apparition d'un événement anormal
  - Notion de « quantité de projections anormales »
- Quantifier la distance d'une donnée anormale vis-à-vis des projections considérées comme normales
- Quantifier la distance d'une densité de projections anormales vis-à-vis des projections considérées comme normales

# Exploitation du modèle dans un objectif de détection





# Exploitation du modèle dans un objectif de détection

## — Détection d'un phénomène anormal

- Franchissement de l'enveloppe normale
  - Enveloppe concave
  - Enveloppe convexe

-> Il faut se ramener à des projections 2D
- Des solutions dans l'imagerie existent aussi
- Détection d'un échantillon isolé
  - On considère qu'un échantillon est isolé si aucune projection normale n'existe dans un périmètre défini.

-> Solution adoptée

# Exploitation du modèle dans un objectif de détection

## — Détection d'un phénomène anormal

- Une donnée anormale peut-être simplement due à un défaut d'enregistrement, on va chercher à savoir la proportion de points anormaux sur une période paramétrable :

$$Ind_{\text{detection}} = \frac{N_{\text{abnormal}}}{N_{\text{total}}}$$

- Nous travaillons dans un espace multidimensionnel... Ce n'est donc pas parce que 2 projections sont anormales qu'elles le sont de « la même manière ».

# Exploitation du modèle dans un objectif de détection

## — Caractérisation d'un phénomène anormal

- Identifier la distance d'une projection par rapport aux projections considérées comme normales
  - On travaille dans un référentiel purement abstrait, la distance que l'on mesure n'a donc aucune signification physique.
  - On aimerait avoir un indicateur de distance normalisé (compris entre 0 et 1) afin de s'interfacer plus aisément avec des travaux en cours à Eurocopter.

En cours de développement...

- On propose de calculer la probabilité d'appartenance d'un nouvel échantillon anormal à l'ensemble des projections normales.
- Identifier la direction d'une projection par rapport aux projections considérées comme normales
  - En se ramenant à des coordonnées complexes, on peut déterminer un angle entre la projection anormale et le centre des données normales.

# Exploitation du modèle dans un objectif de détection

## — Caractérisation d'un phénomène anormal

- Identifier les groupes de projections anormales
  - Carte de Kohonen
    - ✓ Identifier des projections anormales récurrentes
    - ✓ Graphiques -> présentation des résultats efficace
  - K-Mean
    - ✓ Déjà utilisé
    - ✓ Paramétrage assez pénible
  - ...

# Perspectives de développement

- Compléter les phases de tests pour valider ou invalider les objectifs et éventuellement reprendre tout ou partie du travail
- Faire le lien avec d'autre type de données :
  - En cours avec la vibration
  - De nombreuses possibilités avec l'avionique
- Pour le moment, on construit un modèle par hélicoptère. Il serait intéressant d'en avoir un par flotte.
- Les fonctions HFDM (Helicopter Flight Data Monitoring) sont en plein développement et ces types de traitements présentés peuvent être adaptés à leurs besoins.

— ...

Merci pour votre attention...