

# Critical Parameter Identification for Safety Events in Commercial Aviation Using Machine Learning

ARRIS Yanis    TEHAMI Safia

Master 2 AMIS

Janvier 2026

## Article authors :

HyunKi Lee, Sasha Madar, Santusht Sairam,  
Tejas G. Puranik, Michelle Kirby, Olivia J. Pinon,  
Dimitri N. Mavris, Alexia P. Payan

*Aerospace Systems Design Laboratory,  
Georgia Institute of Technology  
MDPI — Aerospace, 2020*

# Plan de la présentation

- 1 Contexte et problématique
- 2 Objectifs
- 3 Données FOQA et Machine Learning
- 4 Méthodologie SAFE
- 5 Étude de cas
- 6 Discussion des limites et perspectives

- La sécurité aérienne s'appuie aujourd'hui sur l'analyse des **données issues des vols de routine (FOQA)**.
- La généralisation des **capteurs embarqués** entraîne une **accumulation massive de données**, atteignant des volumes de l'ordre du téraoctet à l'échelle d'une flotte.
- En parallèle, le trafic aérien connaît une forte croissance :
  - environ 38 millions de vols en 2018, avec une projection à 90 millions d'ici 2040.
- L'enjeu principal est donc de transformer cette masse de données en **connaissance exploitable** pour améliorer la **prévention des risques**.

- En sécurité aérienne, l'enjeu n'est pas uniquement d'analyser les accidents, mais de comprendre **comment ils se construisent progressivement**.
- Cette compréhension repose sur :
  - l'analyse de la **chaîne causale** des événements (*pathogenesis*),
  - l'identification des **facteurs sous-jacents** (*etiology*).
- La sécurité aérienne peut être décrite comme un **continuum** :

Précurseurs → Événements initiateurs → Incidents → Accidents
- **Objectif** : détecter et anticiper le passage d'un vol **nominal** vers une situation à **risque**, **avant** la survenue d'un incident.

- Développer un cadre méthodologique SAFE (Safety Analysis of Flight Events)
- Exploiter les données FOQA pour :
  - Identifier vols **à risque** vs vols **normaux** à partir de **précurseurs**
  - Comprendre la transition vers un **état dangereux**
  - Intervenir **avant** la survenue d'un incident

# Pourquoi les approches classiques atteignent leurs limites ?

- Approche souvent **réactive** : on analyse après l'événement.
- Détection basée sur des **seuils fixes** définis par des experts (**SMEs**).

## Limites principales :

- Les seuils ne capturent pas bien les **dégradations progressives**. (Vision **binaire** : normal / événement)
- Les interactions entre paramètres (effets combinés) sont peu exploitées.
- L'analyse devient difficile à généraliser quand les données sont **massives** et **hétérogènes**.

- Données issues des **vols de routine**, appelées **FOQA** (*Flight Operational Quality Assurance*).
- Enregistrements sous forme d'**observations successives dans le temps** couvrant l'ensemble du vol.
- Plusieurs centaines à milliers de paramètres, échantillonnés jusqu'à **16 Hz**.
- Paramètres provenant de systèmes variés :
  - moteurs, atmosphère, dynamique de vol
  - navigation, commandes de vol, configuration avion

# Pourquoi du Machine Learning ?

- Objectif : passer d'une analyse **rétrospective** à une analyse plus **prospective**.
- Le ML permet de :
  - Exploiter l'**observability-in-depth**
  - Capturer **patterns complexes et non linéaires**
  - Identifier **précurseurs faibles** avant dépassement des seuils
- Finalité : **anticiper** les événements de sécurité, plutôt que les analyser uniquement a posteriori



La méthodologie SAFE (*Safety Analysis of Flight Events*) est structurée en **cinq étapes successives** visant à transformer les données FOQA en information exploitable pour la sécurité aérienne.

- ➊ Prétraitement des données
- ➋ Réduction de la dimensionnalité
- ➌ Génération des vecteurs de caractéristiques
- ➍ Classification
- ➎ Post-traitement et interprétation

# Étape 1 : Prétraitement des données

Objectif : nettoyer les données FOQA brutes afin d'obtenir un jeu exploitable.

- **suppression** : suppression des paramètres trop incomplets
- **Remplacement** : remplacement des valeurs manquantes pour conserver les vols

Exemples (inspirés de l'article) :

suppression :

Temps	A	B	C	⇒	Temps	A	B
10	0	0.05	DNE		10	0	0.05
11	0	0.02	DNE		11	0	0.02
12	1	0.01	DNE		12	1	0.01

Remplacement :

Temps	A	B	⇒	Temps	A	B
13	1	Empty		13	1	0.0000
14	1	Empty		14	1	0.0000

*Les paramètres trop incomplets sont supprimés, les valeurs manquantes ponctuelles sont imputées.*

## Étape 2 : Réduction de la dimensionnalité

Les données FOQA présentent de nombreuses redondances dues à des corrélations physiques entre paramètres.

Matrice de corrélation

	A	B	C	D
A	1	0.996	0.994	0.707
B	0.996	1	0.997	0.991
C	0.994	0.997	1	0.947
D	0.707	0.991	0.947	1



Regroupement des paramètres

Paramètre	Paramètres corrélés
A	[B, C]
B	[C, D]
C	[]
D	[E]
E	[]
F	[]

*Les corrélations  $> 0.99$  indiquent des paramètres redondants.*

*Un seul paramètre est conservé par groupe de forte corrélation (seuil : 0.99), sur la base de l'expertise métier.*

## Étape 3 : Génération des vecteurs de caractéristiques

Les observations successives dans le temps sont transformées pour être compatibles avec la classification.

- Chaque vol est représenté par un **vecteur de caractéristiques**

**Stratégies temporelles proposées :**

- sélection d'un instant unique
- sélection de timestamps consécutifs avant l'événement
- sélection de timestamps échelonnés

# Étape 3 : Génération des vecteurs de caractéristiques (un seul timestamp)

Données temporelles par vol

Flight ID : 123123

<i>C</i>	<i>E</i>	<i>F</i>
10	0	0.0534
11	0	0.0276
12	1	0.0129



Feature Vector Matrix

Flight ID	<i>C</i>	<i>E</i>	<i>F</i>
123123	10	0	0.0534
123126	16	0	0.0276
148279	12	0	0.0386
213219	8	0	0.1108
459876	21	1	0.0041

Flight ID : 123126

<i>C</i>	<i>E</i>	<i>F</i>
16	0	0.0276
17	1	0.0129

*Un instant représentatif est sélectionné par vol.*

# Étape 3 : Génération des vecteurs de caractéristiques (timestamps multiples)

## Données temporelles par vol

Flight ID : 123123

<i>C</i>	<i>E</i>	<i>F</i>
10	0	0.0534
11	0	0.0276
12	1	0.0129

Flight ID : 123126

<i>C</i>	<i>E</i>	<i>F</i>
16	0	0.0276
17	1	0.0129

## Feature Vector Matrix

Flight ID	<i>C</i> <sub>1</sub>	<i>E</i> <sub>1</sub>	<i>F</i> <sub>1</sub>	<i>C</i> <sub>2</sub>	<i>E</i> <sub>2</sub>	<i>F</i> <sub>2</sub>
123123	10	0	0.0534	11	0	0.0276
123126	16	0	0.0276	17	1	0.0129
148279	12	0	0.0386	13	1	0.0021
213219	8	0	0.1108	9	1	0.0098
459876	21	1	0.0041	22	1	0.0017

*Plusieurs instants consécutifs sont concaténés pour capturer la dynamique temporelle.*

## Étape 4 : Classification et critères de sélection

La sélection de l'algorithme de classification repose sur plusieurs critères définis dans l'article :

- Capacité à détecter correctement les événements de sécurité (accuracy)
- Robustesse face aux paramètres corrélés
- Capacité à traiter des données de grande dimension
- Robustesse face au surapprentissage
- Capacité à gérer des paramètres hétérogènes
- Simplicité du réglage des hyperparamètres

Ces critères sont utilisés pour comparer plusieurs algorithmes de classification.

# Comparaison des algorithmes de classification

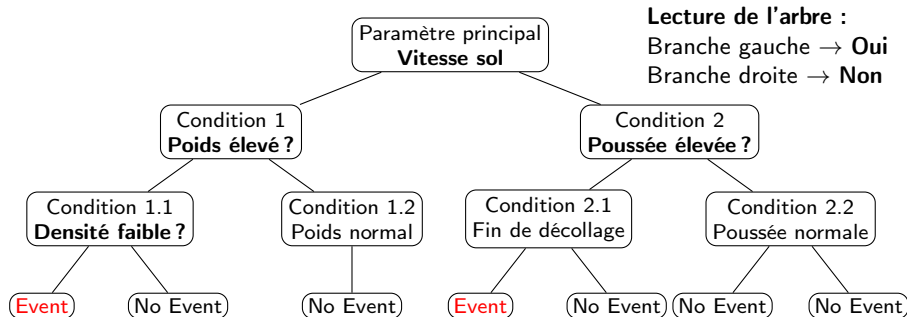
Algorithme	Acc.	Corr.	Dim.	Overf.	Robust.	Tuning
Boosting Ensemble	X	X	X	X		
Decision Tree	X	X				
K-NN		X	X		X	
Bayes Class.			X		X	X
<b>Random Forest</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>	<b>X</b>	

*Acc. : capacité de prédiction    Corr. : résistance aux paramètres corrélés    Dim. : gestion de données de grande dimension    Overf. : résistance au surapprentissage    Robust. : capacité à gérer des paramètres hétérogènes    Tuning : facilité de réglage des hyperparamètres*

**Conclusion :** Random Forest présente le meilleur compromis global, en particulier grâce à sa résistance au surapprentissage et à sa capacité d'interprétation.



# Random Forest : exemple concret (Tire Speed Event)



## Principe de Random Forest

- Chaque **arbre de décision** apprend une suite de règles simples basées sur des seuils de paramètres.
- Les arbres sont construits à partir de **sous-ensembles différents** des données et des paramètres.
- Chaque arbre produit une décision *Event* / *No Event*.
- La décision finale est obtenue par **vote majoritaire** de l'ensemble des arbres.

## Étape 5 : Évaluation et interprétation

L'évaluation du modèle repose sur :

- la **matrice de confusion**
- le **F1-score**, particulièrement adapté aux données déséquilibrées

Matrice de confusion :

	Event réel	Non-Event réel
Event prédit	<i>TP</i>	<i>FP</i>
Non-Event prédit	<i>FN</i>	<i>TN</i>

Interprétation :

- **TP (True Positive)** : événement correctement détecté par le modèle
- **FP (False Positive)** : fausse alerte (événement prédit mais absent)
- **FN (False Negative)** : événement manqué par le modèle (cas critique en sécurité)
- **TN (True Negative)** : vol sans événement de sécurité, correctement identifié

# F1-score : formules et exemple

Formules utilisées :

$$Precision = \frac{TP}{TP + FP} \quad Recall = \frac{TP}{TP + FN}$$

$$F1 = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}$$

**Exemple (Tire Speed Event, article) :**

- $TP = 55$ ,  $FP = 1$ ,  $FN = 0$

$$Precision \approx 0.982 \quad Recall = 1 \quad F1 \approx 0.991$$

# Étude de cas : Tire Speed Event

**Définition :** Le **Tire Speed Event** correspond au dépassement de la vitesse certifiée des pneus lors du décollage.

**Résultats principaux :**

- Performances très élevées du modèle
- **F1-score proche de 1**

**Paramètres critiques identifiés par SAFE :**

- Vitesse sol
- Poids de l'avion
- Densité de l'air
- Poussée moteur (thrust)

**Apport du Machine Learning :** La poussée moteur n'est pas toujours mise en avant dans les descriptions classiques, mais l'analyse ML montre qu'elle devient **significant** lorsqu'elle est considérée **en interaction** avec le poids de l'avion et la densité de l'air.

# Bilan des performances, limites et pistes d'amélioration

- Le cadre SAFE obtient de bonnes performances globales sur plusieurs types d'événements de sécurité.
- Les scores F1 restent élevés, entre **0,85 et 0,99**.
- Le cadre est implémentable en pratique et transférable à d'autres contextes.
- Son efficacité dépend fortement de la qualité des définitions d'événements.

- Certains événements restent difficiles à analyser, notamment le Roll Event et le Landing Distance Event.
- Ces difficultés proviennent de définitions imprécises ou non standardisées.
- Une tension existe entre définitions fonctionnelles et définitions physiques.
- Les algorithmes statistiques peuvent confondre corrélation et causalité, introduisant de l'incertitude dans la chaîne causale.

- Le Machine Learning seul ne suffit pas pour capturer la dynamique réelle des événements.
- L'analyse algorithmique doit être complétée par l'expertise métier (SME).
- Les experts jouent un rôle clé pour contextualiser les résultats et valider leur cohérence physique.
- Une interaction continue entre ML et SME est nécessaire pour garantir la validité opérationnelle des résultats.



- Adapter la conception des événements en réduisant l'usage de seuils fixes et l'agrégation excessive de sous-événements.
- Adapter les algorithmes afin de limiter les facteurs corrélés mais non causaux.
- Explorer des alternatives aux modèles basés sur des arbres de décision.
- Le ML ouvre la voie à une analyse plus prospective des événements, à condition d'une utilisation critique et raisonnée.

# Merci de votre attention

*À vos questions*