

# Evaluation de performances

## Cas de l'étude Dataset-M (disponible sur Arche)

- 1 Introduction
- 2 Présentation du cas de l'étude
- 3 Implantation d'un processus de pronostic
- 4 Résultats attendus

Phuc Do  
phuc.do@univ-lorraine.fr



## Objectif:

- Modéliser le processus de dégradation de trains de filtrage d'eaux industriel
- Estimer l'évolution du son indicateur de santé
- Estimer sa fiabilité et sa durée de vie résiduelle

## Données et outils:

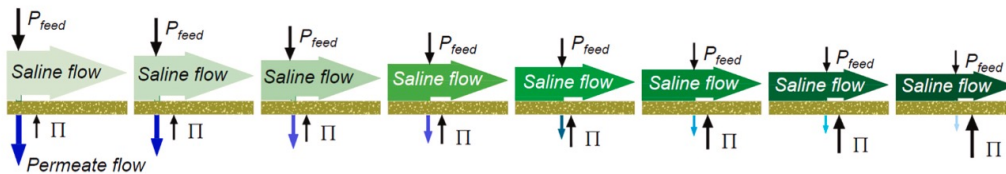
- Un jeu de données disponible sur Arche
  - 14 trains donc chaque train est composé de 8 filtres élémentaires
  - Différentes actions de maintenance (permutation)

## Modalité de l'évaluation:

- Travail en groupe ou seul(e)
- Compte rendu à déposer sur Arche pour le **21/05/2024**

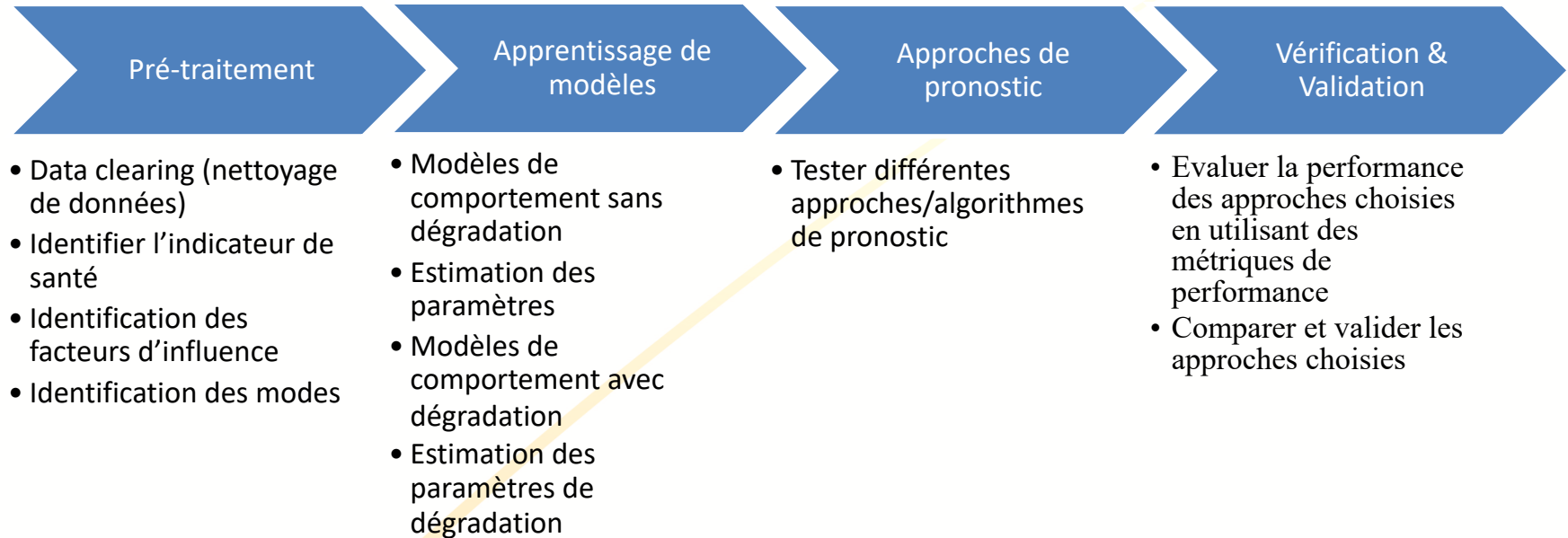
## Mesures/indicateurs pour chaque train:

Date	Da	wee	Train Status Code	Normalized DP (bars)	Feed Volume (m <sup>3</sup> )	Brine Volume (m <sup>3</sup> )	Product Volume (m <sup>3</sup> )	Recovery (%)	K	Train Status
07/11/2015	1	0,14	5	0,780	21243	10375	10867	51,2	1,0	Operation
08/11/2015	2	0,29	5	0,776	30060	14709	15351	51,1	1,0	Operation
09/11/2015	3	0,43	5	0,776	30065	14712	15353	51,1	1,0	Operation
10/11/2015	4	0,57	5	0,775	29762	14562	15201	51,1	1,0	Operation
11/11/2015	5	0,71	5	0,773	29371	14352	15019	51,1	1,0	Operation
12/11/2015	6	0,86	5	0,778	29563	14443	15120	51,1	1,0	Operation
13/11/2015	7	1,00	5	0,779	11067	5411	5656	51,1	1,0	Operation
14/11/2015	8	1,14	5	0,763	16893	8256	8637	51,0	1,0	Operation
15/11/2015	9	1,29	5	0,769	29278	14038	15240	52,1	1,0	Operation
16/11/2015	10	1,43	5	0,769	29215	14002	15214	52,1	1,0	Operation
17/11/2015	11	1,57	5	0,773	29235	14001	15234	52,1	1,0	Operation
18/11/2015	12	1,71	5	0,775	29761	14251	15510	52,1	1,0	Operation



# Implantation d'un processus de pronostic de défaillance

## Etapes principales:



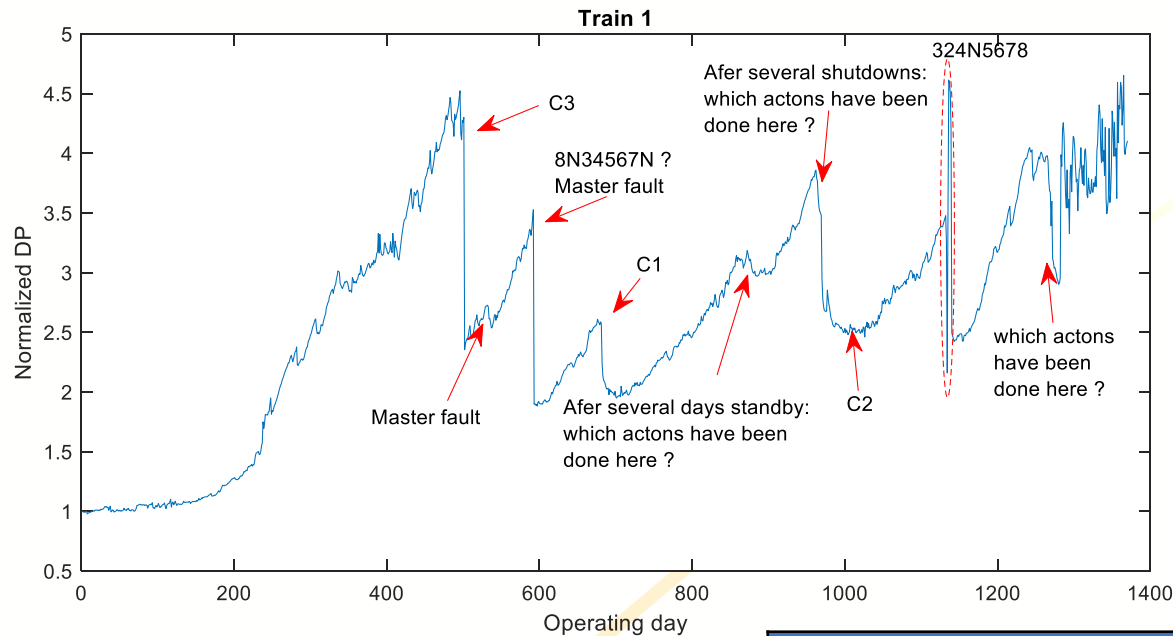
# Implantation d'un processus de pronostic de défaillance

## Prétraitement

- ❑ Data clearing (nettoyage de données)
  - Eliminer les données inutiles (périodes d'arrêt, points autour de redémarrage, ...)
  - Filtrer les bruits
- ❑ Identifier l'indicateur de performance: *paramètres en sortie*
  - Variation de pression (DP)
- ❑ Identification des facteurs d'influence: *paramètres en entrée*

# Présentation du cas de l'étude

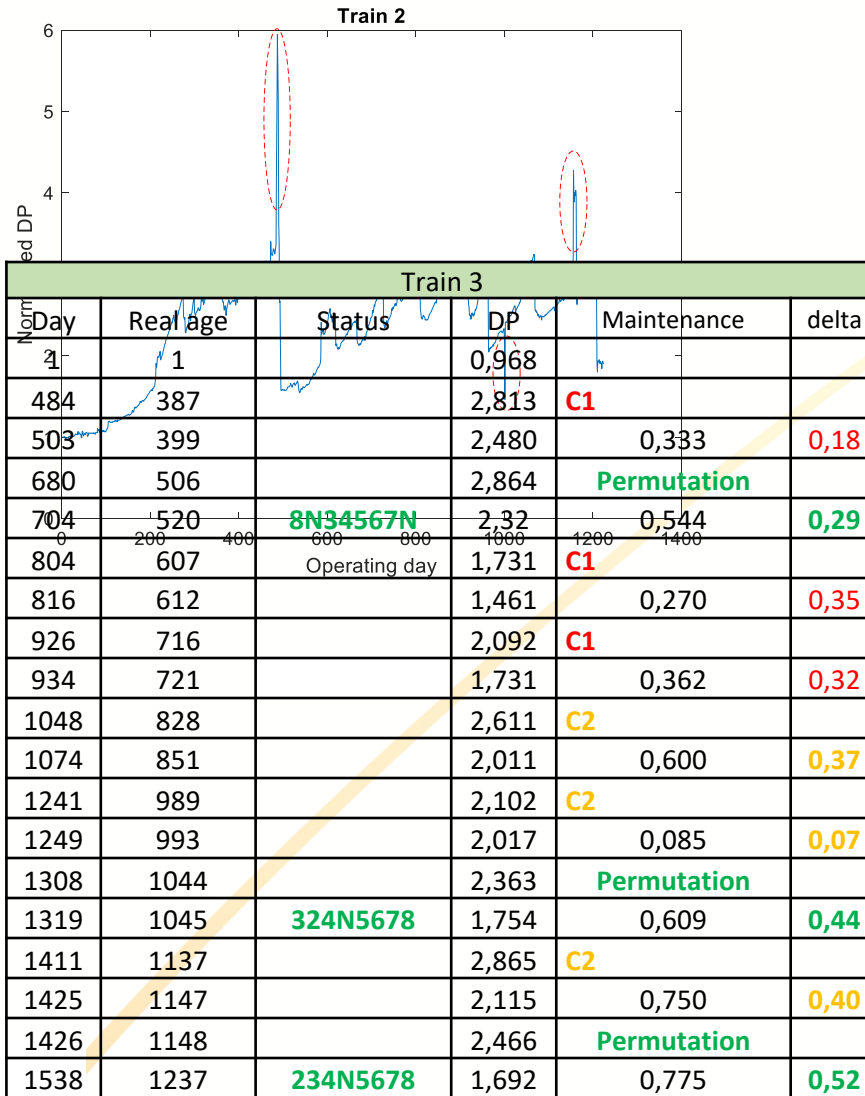
## Quelques exemples sur des données après nettoyage



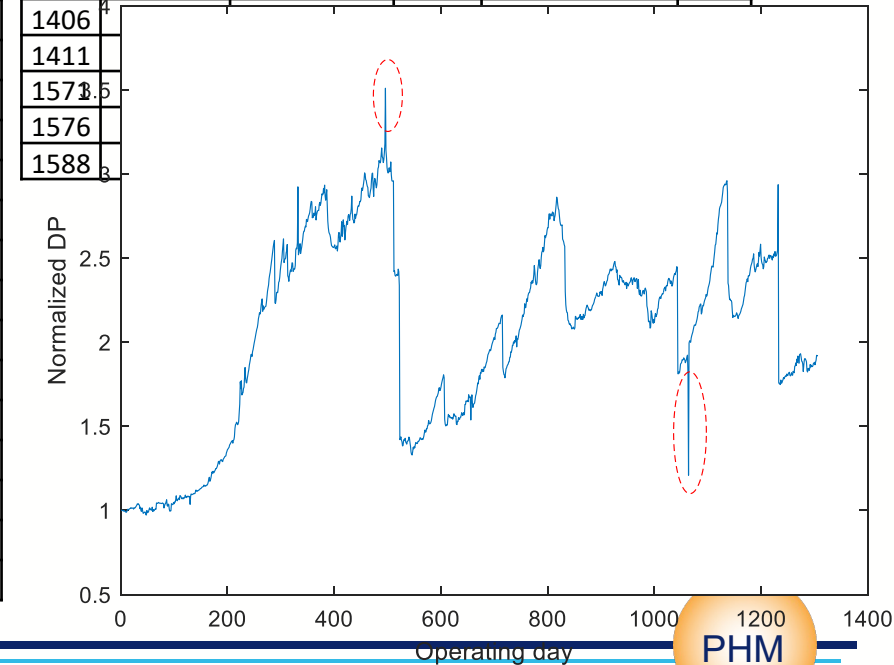
Train 1					
Day	Real age	Status	DP	Maintenance	delta
1	1		0,78		
610	497		3,526	<b>C3</b>	
620	503	234567N	1,835	1,691	0,62
835	682		2,008	<b>C1</b>	
858	701		1,527	0,481	<b>0,39</b>
1156	981		2,106	<b>C2</b>	
1188	999		1,948	0,158	<b>0,12</b>
1189	1000		2,745	<b>Permutation</b>	
1349	1141	<b>324N5678</b>	1,897	0,847323784	<b>0,43</b>

# Présentation du cas de l'étude

## Quelques exemples sur des données après nettoyage



Train 2					
Day	Real age	Status	DP	Maintenance	delt a
1	1		0,732		
465	362		2,154	<b>C1</b>	
491	372		1,826	0,328	<b>0,23</b>
640	492		2,479	<b>Permutation</b>	
647	495	<b>8N34567N</b>	1,497	0,982	<b>0,56</b>
933	667		1,790	<b>C1</b>	
942	672		1,558	0,232	<b>0,22</b>
1037	765		2,007	<b>C1</b>	
1045	768		1,952	0,055	<b>0,04</b>
1290	96		2,131	<b>Permutation</b>	
1298	966	<b>324N5678</b>	1,497	0,634	<b>0,45</b>



# Implantation d'un processus de pronostic de défaillance

Apprentissage  
de modèles

## Principe de base:

- 70% données pour la phase d'apprentissage
- 30% données restant pour le test

- Apprendre le modèle de comportement sur l'évolution de l'indicateur de santé identifié en fonction des facteurs d'influence identifiés
  - ✓ Modèle de régression: linéaire ou non-linéaire, machine learning
  - ✓ Estimation des paramètres sur la période [0 70%]
- Tester les paramètres estimés sur la période restant, e.g., [70% 100%]
  - ✓ R2 (coefficient de détermination)
  - ✓ RMSE (Écart quadratique moyen)



## Choisir deux approches/modèles à tester

- 1 ou 2 modèles de comportements (linéaire, non-linéaire)
- 1 ou 2 modèles de dégradation (eg., regression, processus de stochastique, ...)
- Autres

# Implantation d'un processus de pronostic de défaillance

## Vérification & Validation

1. Vérifier la cohérence les résultats de prédictions donnés par les approches choisies
2. Evaluation la performance de deux approches choisies
  - Tester avec des métriques performance (R2, RMSE, etc.)
  - Classer les deux approches choisies selon les critères utilisés

1. Identifier clairement un ou plusieurs indicateurs de santé à suivre ainsi que leur seuil de défaillance
2. Proposer deux approches permettant de modéliser les comportements avec et sans dégradation du système
3. Estimation des paramètres des modèles proposés
4. Interprétation des résultats finaux donnés par la meilleure approche identifiée