

CRAN CENTRE DE RECHERCHE EN AUTOMATIQUE DE NANCY

ADAPTATION DE MODÈLE EN PRÉSENCE D'UNE DÉRIVE SUR LE PROCESSUS MODÉLISÉ

NOYEL M., THOMAS P., THOMAS A., CHARPENTIER P.
RECONFIGURATION PROCESS FOR NEURONAL CLASSIFICATION MODELS: APPLICATION TO A QUALITY MONITORING PROBLEM
COMPUTERS IN INDUSTRY 83 (2016) PP 78-91

P. Thomas

INS2I
INSIS

AM2I

cnrs UMR 7039

UNIVERSITÉ DE LORRAINE

Institut de Cancérologie de Lorraine
Alexis Vazirin
Ensemble, construisons l'avenir

1

PLAN

- Introduction
- Machine learning
- Monitoring du modèle
- Sélection des données pour réapprentissage
- Applications
 - Exemple de simulation
 - Monitoring de la qualité (ACTA mobilier)
 - Prédiction de la pollution (Anaximen)
- Conclusion

CRAN

IMS2-26 novembre 2021

2 / 23

P. Thomas

ISET

2

INTRODUCTION

- **Machine learning**
 - Exploitation de l'historique de données
 - Modèle statique
 - Classification - régression

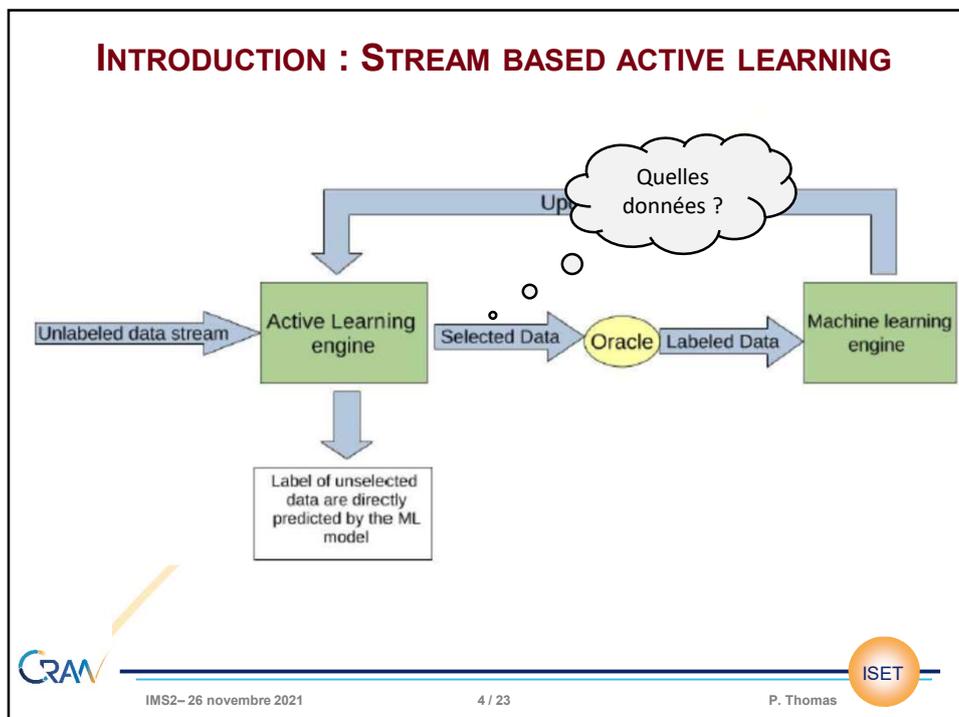
- **Systèmes industriels**
 - Evolutifs
 - Changement brusque (remplacement de machine ...)
 - Dérive continue (encrassement de filtre...)
 - Besoin d'adapter le modèle au changement

- **Active learning**
 - Pool-based active learning
 - Une base de données ; Certaines sont sélectionnées pour être labélisées
 - Membership query synthesis based active learning
 - Données générées artificiellement puis labélisées par un oracle
 - Stream based active learning

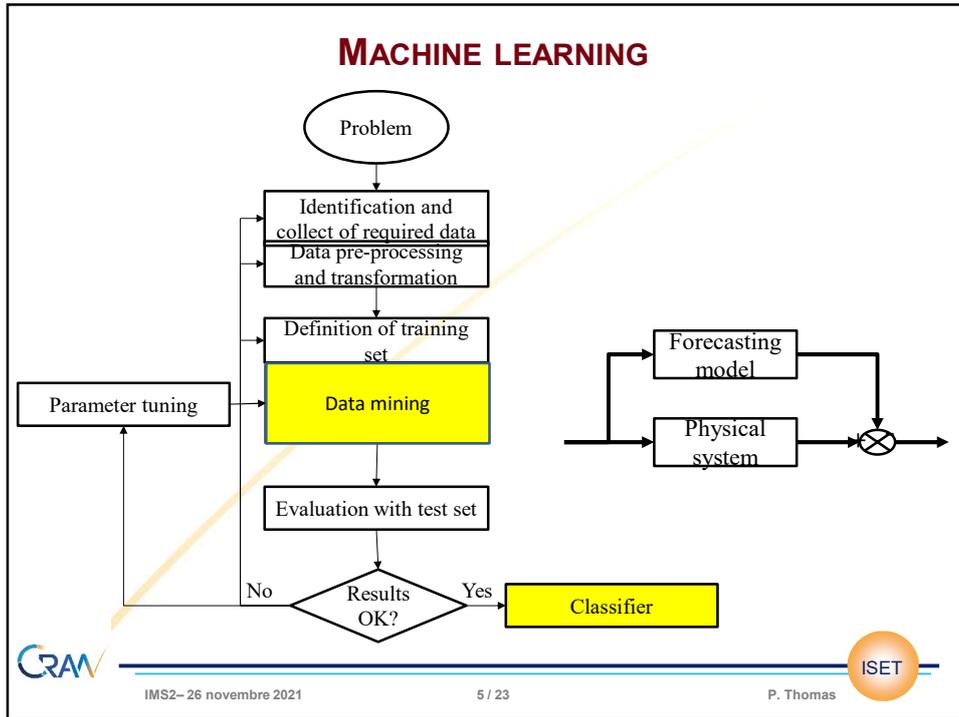
 

IMS2-26 novembre 2021 3 / 23 P. Thomas

3



4



5

MACHINE LEARNING : PERCEPTRON MULTICOUCHES

- **NN feedforward**
 - Apprentissage supervisé
 - Problèmes de classification ou de régression
 - Jeux de données d'apprentissage et de validation

$$z = g_2 \left(\sum_{i=1}^{n_1} w_i^2 \cdot g_1 \left(\sum_{h=1}^{n_0} w_{ih}^1 \cdot x_h^0 + b_i^1 \right) + b \right)$$

- **Etapes de l'apprentissage**
 - Initialisation [Nguyen and Widrow, 1990]
 - Apprentissage (robust Levenberg-Marquard algorithm [Thomas et al., 1999])
 - Détermination de la structure (Pruning [Setiono and Leow, 2000])

IMS2-26 novembre 2021 6 / 23 P. Thomas ISET

6

MONITORING DU MODÈLE : BESOIN

- Le système évolue
 - Modification du process
 - Choix de tête de laquage
 - Déplacement de capteurs
 - ...
- L'environnement évolue

- Contrôle des paramètres environnementaux (climatisation)
- Réglage robuste des paramètres
- Amélioration du modèle en fonction des conditions nouvelles

IMS2-26 novembre 2021 7 / 23 P. Thomas

7

MONITORING DU MODÈLE : PRINCIPE

IMS2-26 novembre 2021 8 / 23 P. Thomas

8

MONITORING DU MODÈLE : CARTE DE CONTRÔLE

- Probabilité de 0.27% qu'une donnée sorte des bornes 3-sigma (Pareto).
- Objectif : Déterminer statistiquement si la variation est « sous contrôle » ou s'il s'agit d'une dérive.

○ Chaque point : valeur du critère (MSSE, taux de mauvaise classification) sur un ensemble de k données

○ 2 paires de bornes

$$\left\{ \begin{array}{l} LCL_{95\%} = p - 1.96 \sqrt{\frac{p(1-p)}{k}} \\ UCL_{95\%} = p + 1.96 \sqrt{\frac{p(1-p)}{k}} \end{array} \right. \quad \left\{ \begin{array}{l} LCL_{99.8\%} = p - 3 \sqrt{\frac{p(1-p)}{k}} \\ UCL_{99.8\%} = p + 3 \sqrt{\frac{p(1-p)}{k}} \end{array} \right.$$

K : TAILLE DU DATASET
P : VALEUR CENTRALE À ESTIMER → VALEUR DU CRITÈRE OBTENUE SUR LE JEU DE VALIDATION

Avantages

- Solution visuelle
- Particulièrement adaptée au contrôle dynamique de séries de données temporelles (Tague, 2004)

Limites

- Pas d'estimation de l'instant de début de dérive

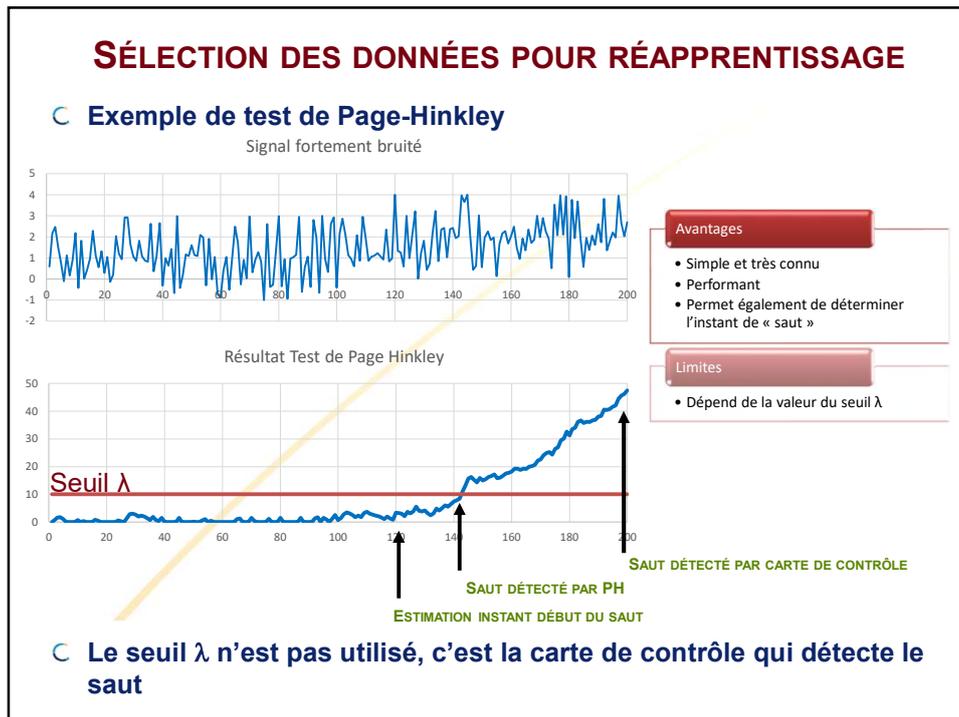
9

SÉLECTION DES DONNÉES POUR RÉAPPRENTISSAGE

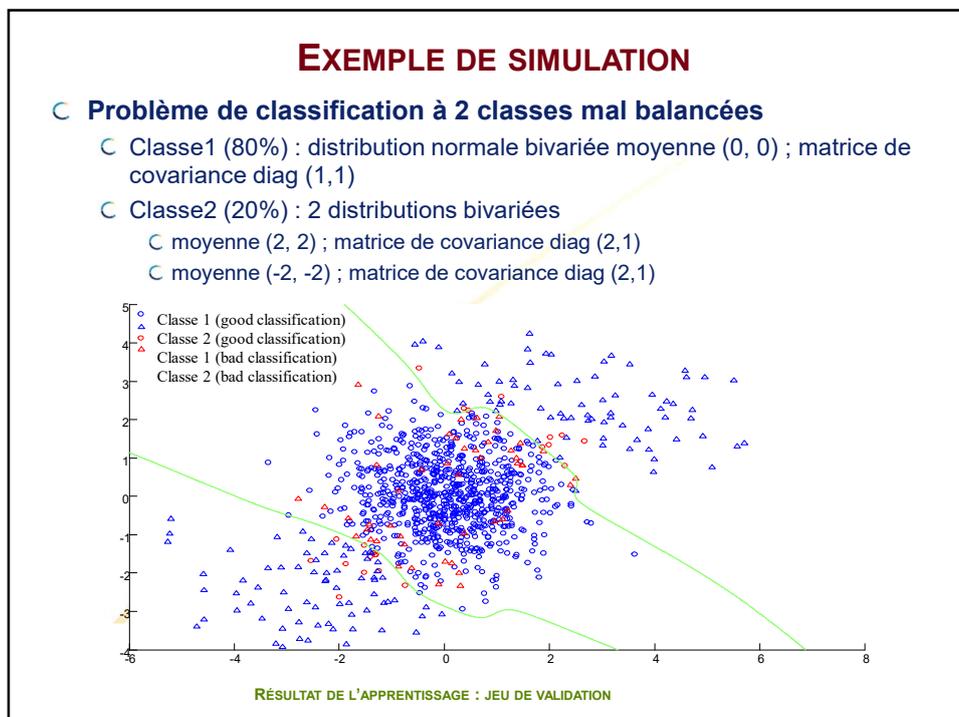
- **Dépend de la cause de la dérive**
 - Défaillance du système → c'est le système qui doit être corrigé
 - Nouveau point de fonctionnement → reprendre toutes les données
 - Changement du système → reprendre les données depuis le changement
- **Détermination du dataset pour réapprentissage**
 - Prendre les k dernières données
 - Prendre toutes les données depuis le dernier apprentissage
 - Détecter le début du changement : test de Page-Hinkley
 - Technique d'analyse séquentielle utilisée pour détecter des sauts de moyenne dans un signal constant pollué par un bruit blanc (Hinkley, 1971)
 - Test d'hypothèse :

$$\left\{ \begin{array}{l} H_0 : e_i \sim N(m_0, \sigma^2), \quad P(e_i) = P_0(e_i) \quad i = 1, \dots, l \quad (H_0 : \text{pas de saut}) \\ H_1 : \begin{array}{l} e_i \sim N(m_0, \sigma^2), \quad P(e_i) = P_0(e_i) \quad i = 1, \dots, r-1 \\ e_i \sim N(m_1, \sigma^2), \quad P(e_i) = P_1(e_i) \quad i = r, \dots, l \end{array} \quad (H_1 : \text{saut}) \end{array} \right.$$

10



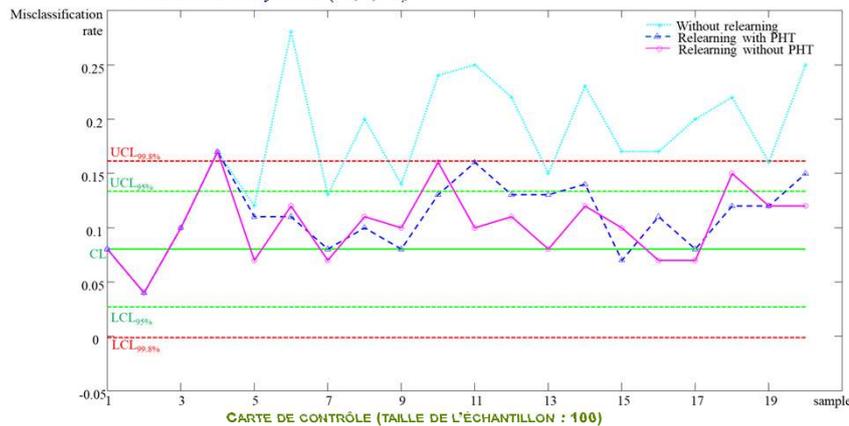
11



12

EXEMPLE DE SIMULATION : CONCEPT SHIFT (SAUT)

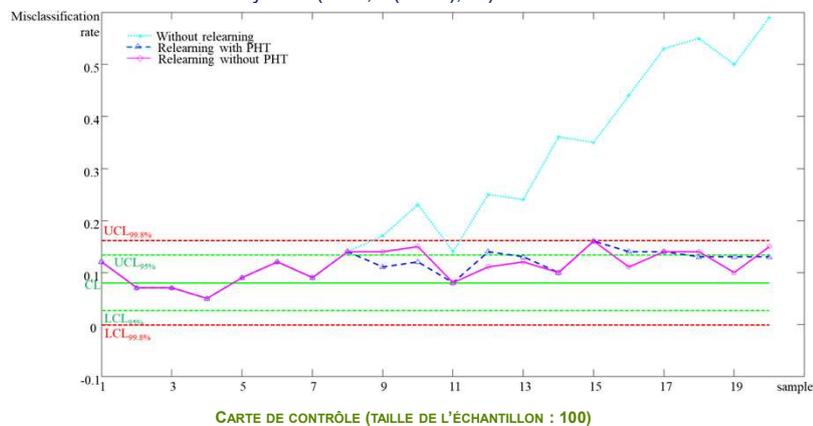
- Jeu de donnée supplémentaire : 2000 (même distribution)
- Saut de moyenne à l'instant $t=300$
 - Classe1 : nouvelle moyenne (1,5, 0)
 - Classe2 : 2 distributions bivariées
 - Nouvelle moyenne (3,5, 2)
 - Nouvelle moyenne (-0,5, -2)



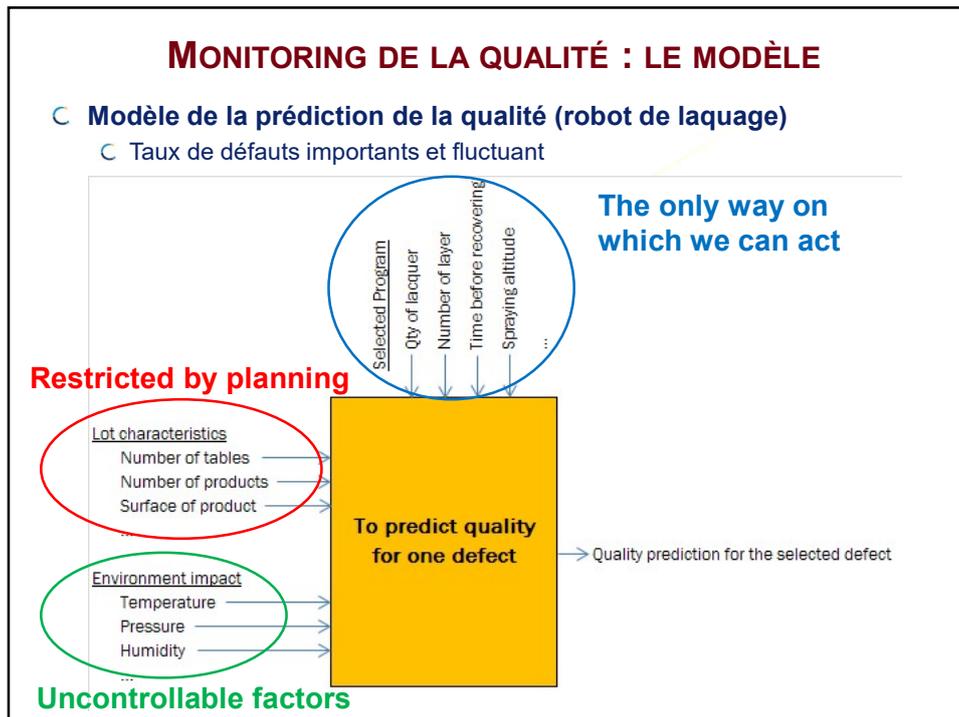
13

EXEMPLE DE SIMULATION : CONCEPT DRIFT (DÉRIVE)

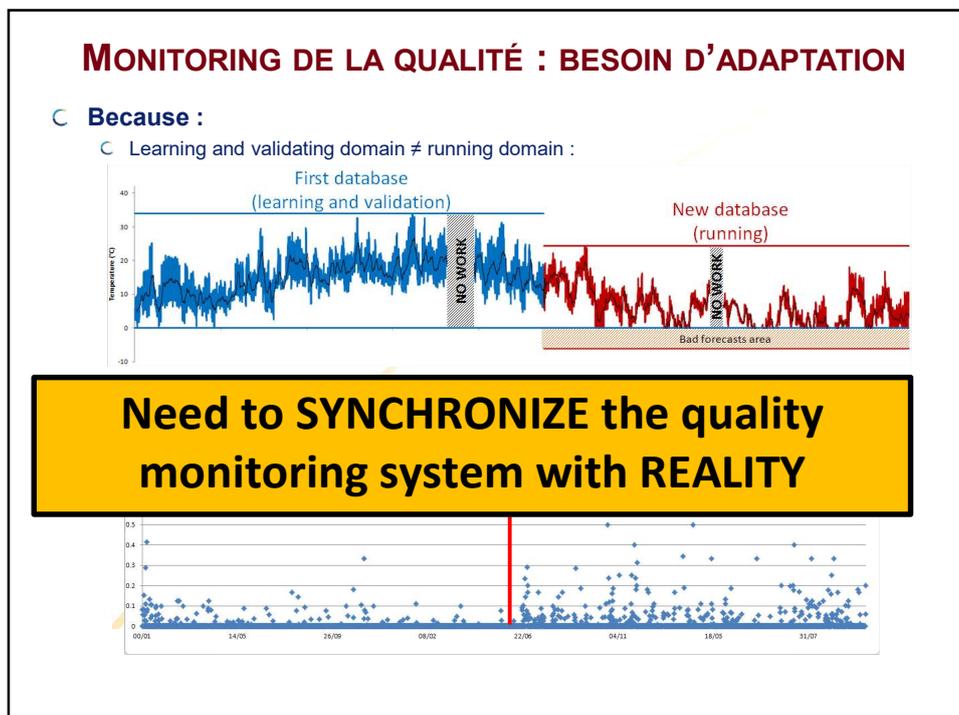
- Jeu de donnée supplémentaire : 2000 (même distribution)
- Dérive à parti de l'instant $t=300$
 - Classe1 : nouvelle moyenne $(0,02*(t-300), 0)$
 - Classe2 : 2 distributions bivariées
 - Nouvelle moyenne $(2+0,2*(t-300), 2)$
 - Nouvelle moyenne $(-2+0,2*(t-300), -2)$



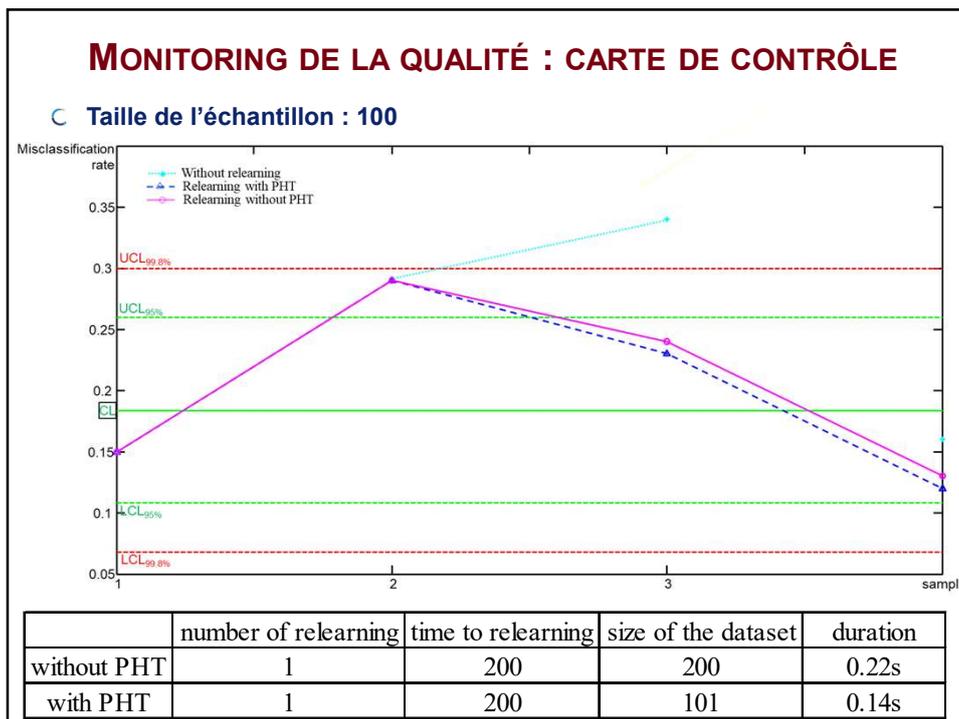
14



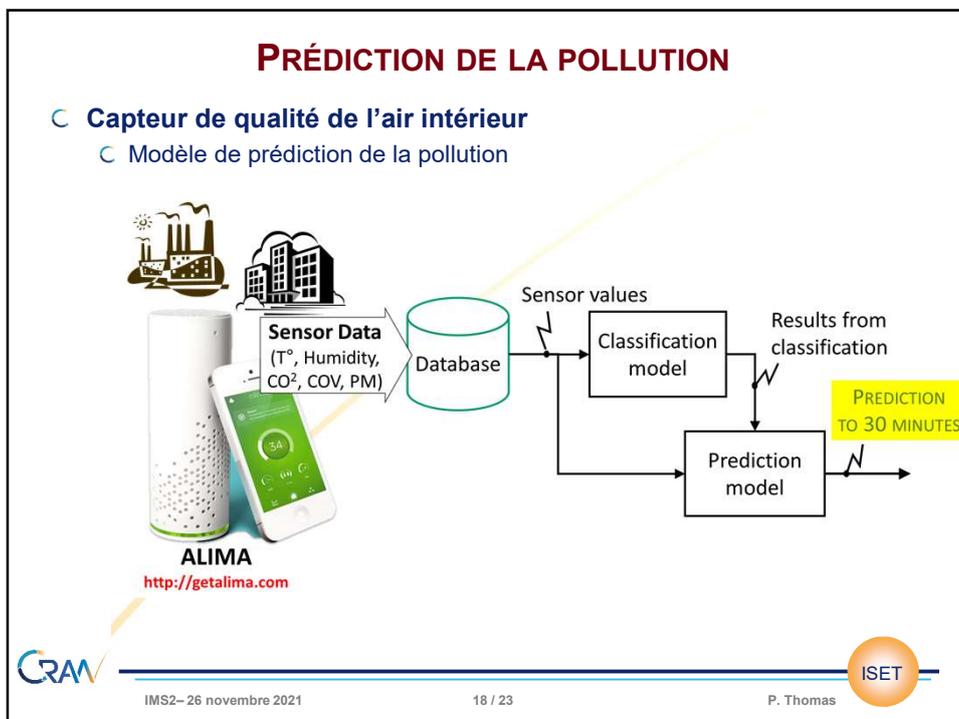
15



16



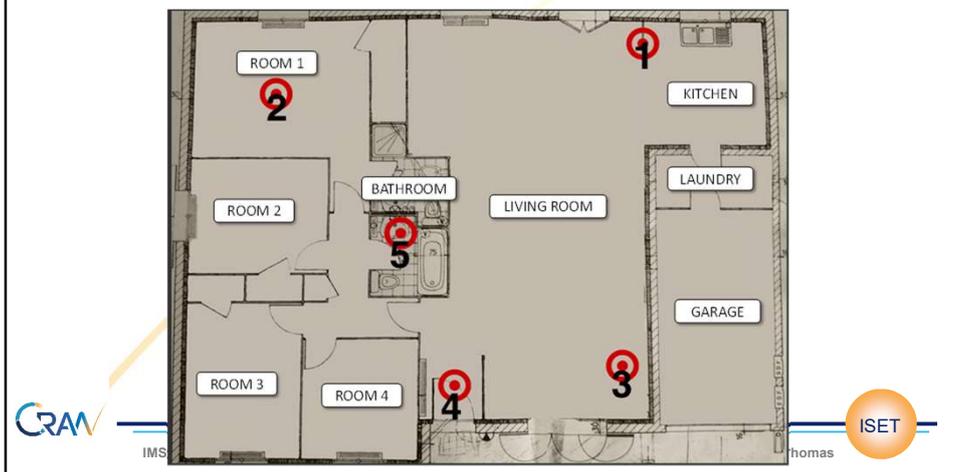
17



18

PRÉDICTION DE LA POLLUTION : BESOIN

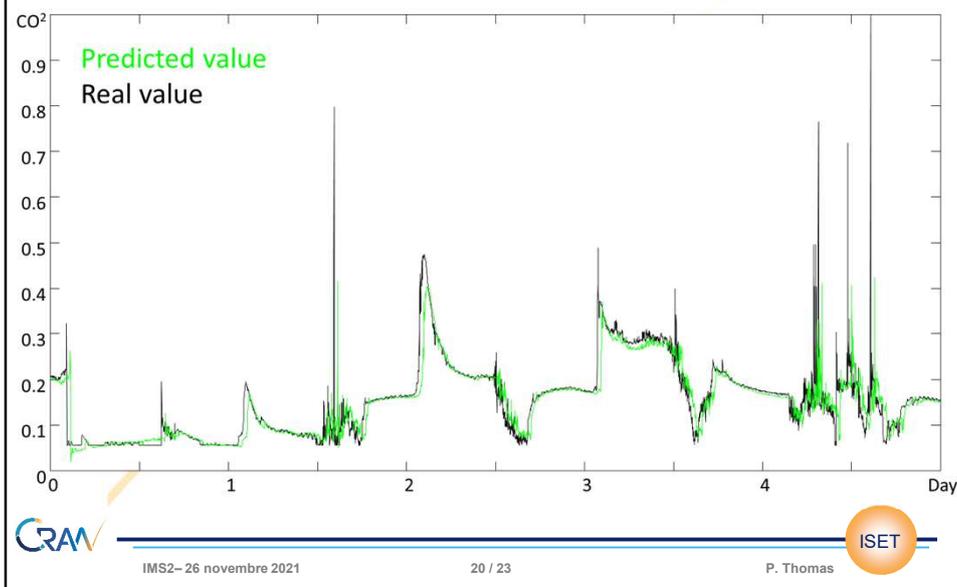
- Acquisition de données dans une habitation réelle Durant 1 mois
- 5 capteurs différents
- Peut on utiliser le modèle développer sur 1 capteur pour 1 autre capteur ?



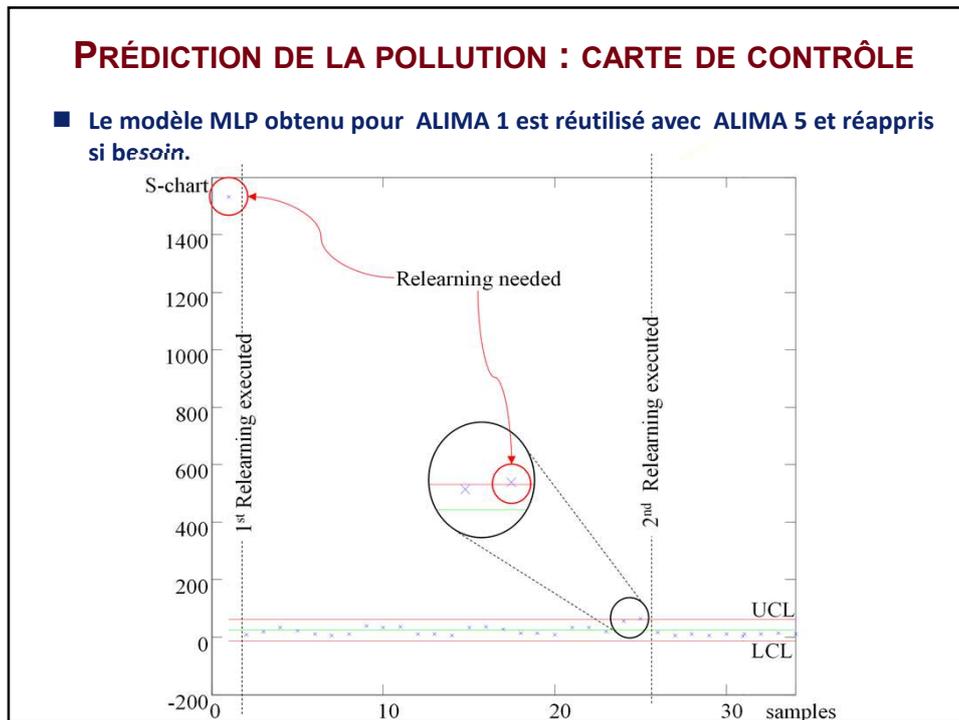
19

PRÉDICTION DE LA POLLUTION : MODÈLE ALIMA 1

- CO₂ prediction (t+30min) - validation



20



21

CONCLUSION

- **Machine learning**
 - Permet de modéliser des problèmes de régression et classification
 - Nécessite des phases de réadaptation au changement
- **Adapter le modèle**
 - Détecter le besoin de réapprendre
 - Déterminer le dataset sur lequel réapprendre
 - réapprendre
- **Carte de contrôle**
 - Permet de déterminer le besoin de réapprentissage
- **Test de Page Hinkley**
 - Permet de construire le dataset sur lequel réapprendre

CRAN

IMS2-26 novembre 2021

22 / 23

P. Thomas

ISSET

22



MERCI POUR VOTRE ATTENTION

ADAPTATION DE MODÈLE EN PRÉSENCE D'UNE DÉRIVE SUR LE PROCESSUS MODÉLISÉ

NOYEL M., THOMAS P., THOMAS A., CHARPENTIER P.
RECONFIGURATION PROCESS FOR NEURONAL CLASSIFICATION MODELS: APPLICATION TO A QUALITY MONITORING PROBLEM
COMPUTERS IN INDUSTRY 83 (2016) PP 78-91

P. Thomas

INS2I
INSIS

AM2I



UMR 7039



UNIVERSITÉ DE LORRAINE



Institut de Cancérologie de Lorraine
Alexis Vazirin
Ensemble, construisons l'avenir