



IMT Lille Douai
École Mines-Télécom
IMT-Université de Lille

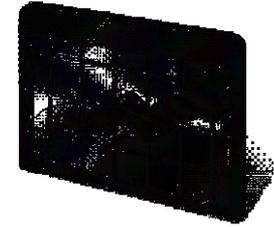
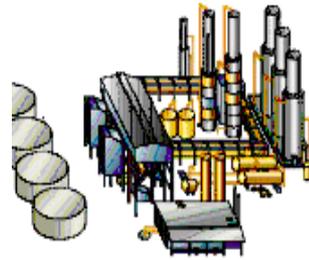
SYSTÈMES DE PRODUCTION DU FUTUR

ANALYSE DE DONNÉES ET PERFORMANCE INDUSTRIELLE

S. LECOEUICHE, A. FLEURY
*INFORMATIQUE ET AUTOMATIQUE
IMT LILLE DOUAI*

B. HERVE
E.L.M. LEBLANC





Optimisation de processus

Réduction de coûts, énergie, matière, consommables

Amélioration et Optimisation des performances

« Faire le mieux avec ce que nous avons »

Réduction de la variabilité de la « Qualité »

Réduction des pertes et minimisation des risques

Adaptation au contexte, Usine Agile

Analyse et suivi de la performance

Monitoring des indicateurs de performances – mise à jour en continue

Maintenance prédictive

Analyse durée de vie

Traçabilité



IA
Informatique
et Automatique

Acquisition des données venant d'automates, des machines outils, de réseaux industriels, d'IOT des systèmes de GPAO, des ERP...

Mise en œuvre de protocoles générique (OPC) ou propriétaire

Datation et qualification des données

Fiabilisation des données

Gestion de sécurité de haut niveau

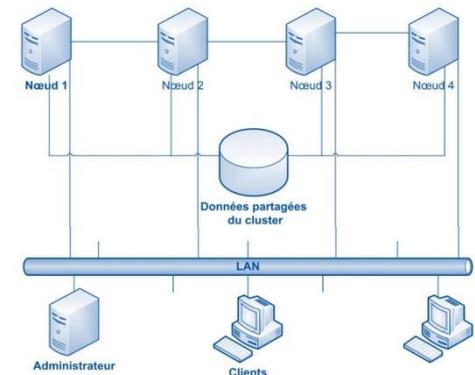
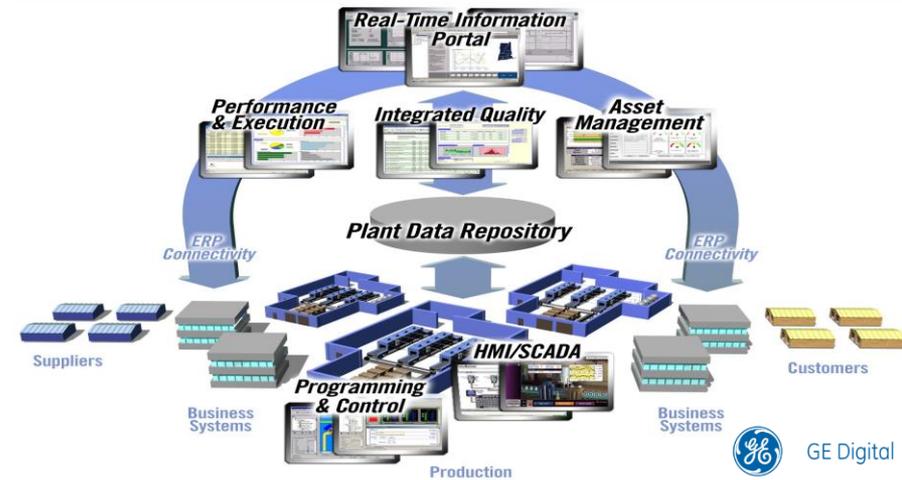
Architectures d'archivage et de traitement

Archivage de données qualifiées et datées,

Archivage sécurisé en vue d'utilisation normative

Mémorisation des anomalies

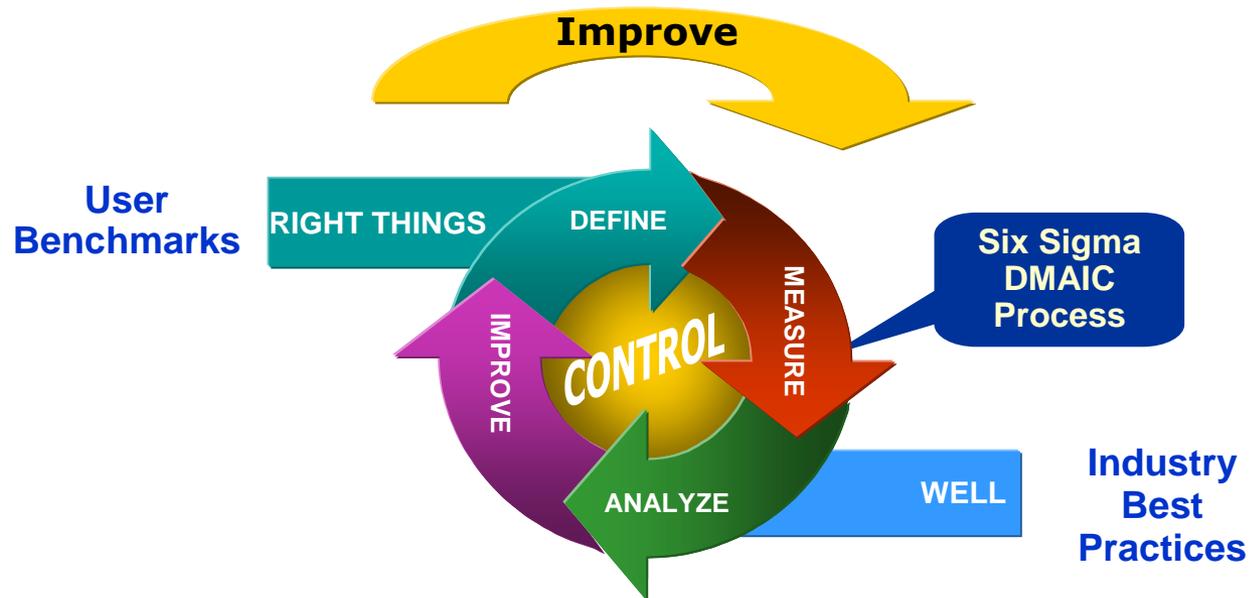
Archivage compatible espace-projet Big Data



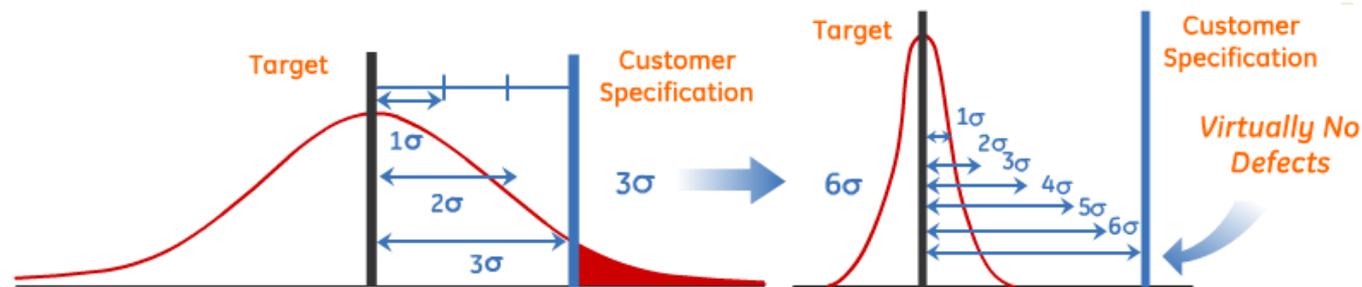
Analyse de données pour l'amélioration des procédés industriels

Inspiré de la démarche du 6 σ

Définir, Mesurer, Analyser, Améliorer, Contrôler ...



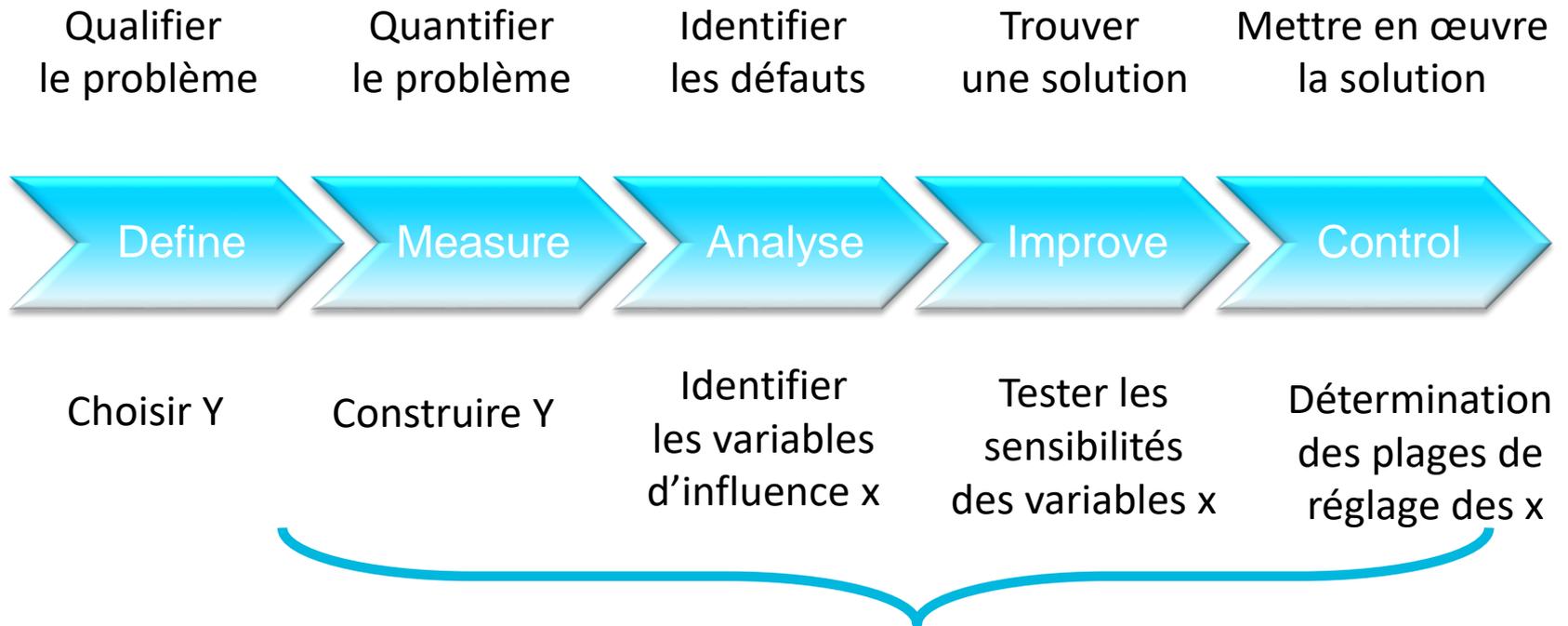
Garantir une variabilité dans une plage très acceptable



Enjeu : Mise au point de procédés plus rapide et garantie dans le temps

**Adapter rapidement les réglages pour obtenir des points de fonctionnement idéaux selon la qualité des produits à réaliser,
Réduire les délais de réglage, ajuster les réglages,
Adapter le contrôle ou intervenir sur le procédé...**

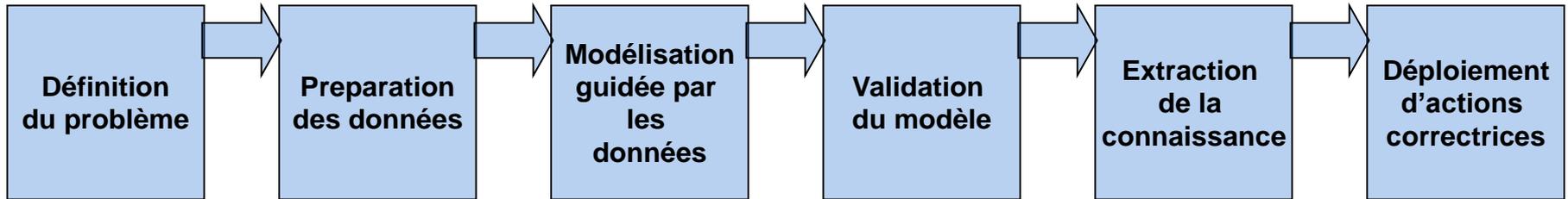
Inspirée du « DMAIC Improvement Process »



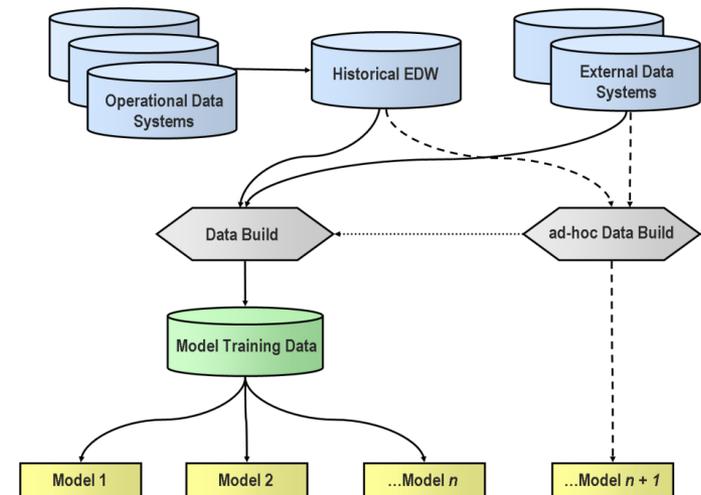
$$Y = f(x) + \epsilon$$

ETAPES GÉNÉRALES D'UN PROJET ADAPI

8



- Choix de la structure du problème
- Sélection de variables (voire transformation)
- Choix de la méthode de modélisation
- Réglages de paramètres
- Validation
- Généralisation du modèle



Besoins

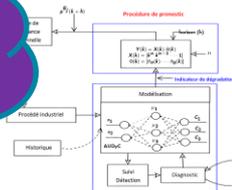
Comprendre les phénomènes et les dépendances complexes
Modéliser et caractériser les comportements hybrides

Caractériser les modes de fonctionnement et identifier les informations discriminantes
Exploiter des modèles mis à jour, pour une meilleure planification et par des stratégies de pilotage autonome

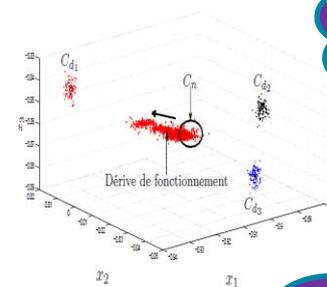
Détecter et suivre les changements de comportement
Pronostiquer les modes / comportements futurs
Anticiper en termes de maintenance et de résilience

Intelligence Artificielle et Automatique

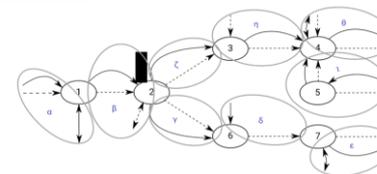
Modélisation des systèmes dynamiques et hybrides



Classification dynamique



Décision et Optimisation multi-critère



Projet de R&D en partenariat avec l'entreprise e.l.m. Leblanc

Faisabilité de la prédiction d'un retour produit, à partir des données de la chaîne de production

Phase du projet

Découverte du processus, des ateliers de fabrication, et des données

Appropriation du vocabulaire, des champs par type de bases fournies et des codifications utilisées

Mise en place d'un espace projet Big Data TERALAB

Compréhension des processus par les données disponibles

Recherche d'attributs et de modèles permettant une prédiction

Evaluation des modèles

Recommandation en vue de l'implémentation des actions d'amélioration



e.l.m. leblanc

Des fichiers issus des plusieurs systèmes d'information de 2010 à 2017
Planification de la production, méthodes, SAV, Audit qualité, Logistique...

Variété importante de produits rencontrés

602 933 numéros de série différents

164 000 numéros de série communs à plusieurs fichiers

Nombre de chaudières en retour SAV

161928 chaudières n'ayant jamais sans retour

9807 chaudières concernées

2804 chaudières revenues en SAV dans les 2 ans

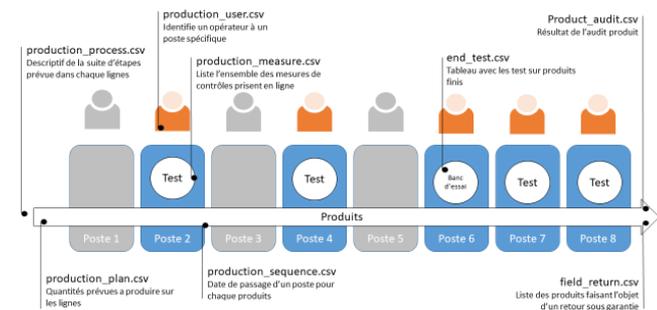
6745 avec un numéro de série valide

Une large éventail de variables

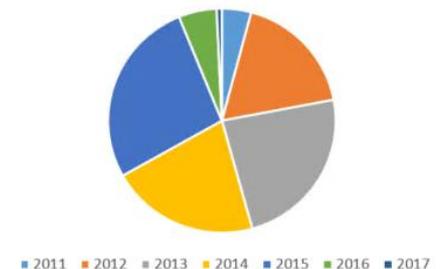
94 attributs potentiellement utilisables comme variables explicatives

2 comme variables à expliquer :

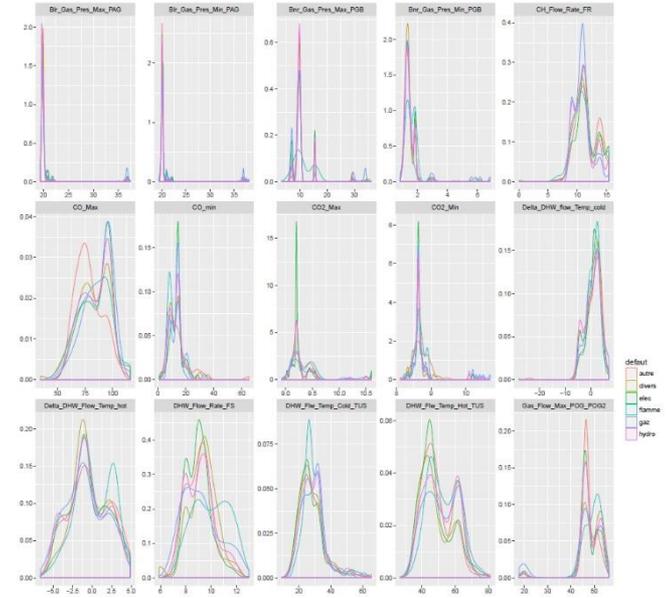
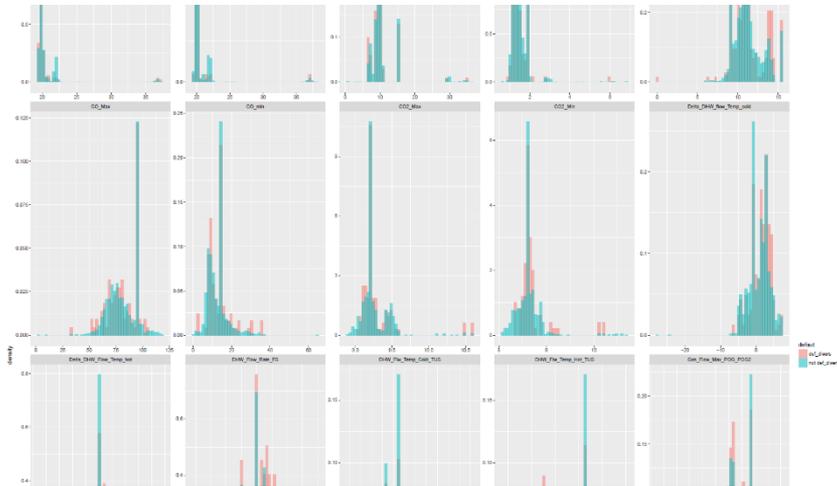
le retour SAV et le type de défaut associé



Répartition des chaudières en retour SAV



Représentation des distributions selon la classe de défaut

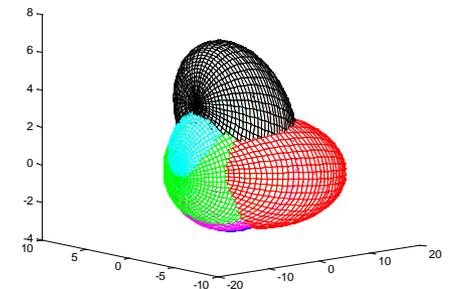


Analyse de la variance selon le type de défaut

Projection suivant 14 principales variables explicatives

Fort chevauchement entre les 6 modes de défaut

Test MANOVA en échec



Construction d'un modèle de prédiction, cherchant à estimer en sortie d'atelier, un éventuel retour en SAV

Exploitation d'attributs plus complets, construits à partir des données et des mesures de la chaîne de production et sélectionnés en fonction de leur influence sur les performances du modèle

Etude de l'influence des attributs construits (46 après suppression des données manquantes), suivant leur capacité à discriminer les défauts SAV

Rendu possible grâce à la performance de calcul de TERALAB
3 différentes méthodes de sélection ont été testées

Une limitation : incertitude sur la labélisation des chaudières conformes, utilisés comme références pour le modèle.

Manque de « Golden Batches » exploitables

Modèles d'apprentissage avec des méthodes de reconnaissances non linéaires (Machines à Vecteurs Supports).

Attributs	Méthode 1	Méthode 2	Méthode 3
Temps sur les postes	2, 3, 4, 14	1, 2, 6, 7, 8, 11, 12, 15	5, 6, 7, 8
Temps passé sur des mesures	<u>debit_gaz</u> , <u>pression_sanitaire</u> , fuite_sanitaire	<u>debit_gaz</u> , pression_WB6, courant_masse, tension_dielectrique, courant_dielectrique, <u>pression_sanitaire</u> , pression_caisson	/
Nombre de mesures	serrage_force	courant_masse, courant_dielectrique	pression_WB6 et pression_caisson
Valeurs moyennes de mesures	/	Tous les types disponibles (tous sauf mesure_diametre)	debit_WB6 et fuite_sanitaire
Pourcentage mesures échouées	/	pression_caisson	/
End Test	DHW_Flow_Rate_FS, DHW_Flw_Temp_Hot_TU S, Gas_Flow_Max_POG_PO G2, Blr_Gas_Pres_Min_PAG, Blr_Gas_Pres_Max_PAG, CO_min	Toutes	DHW_Flow_Rate_FS, CO_min
Total	14	46	10

En entrées – 46 attributs

En sorties – 6 classes (conforme et 5 défauts)

Nombre d'individus

120734 conformes, 69 df1, 64 df2, 333 df3, 584 def4, 1302 def5

Deux modèles appris

V1 – Basé sur les éléments disponibles

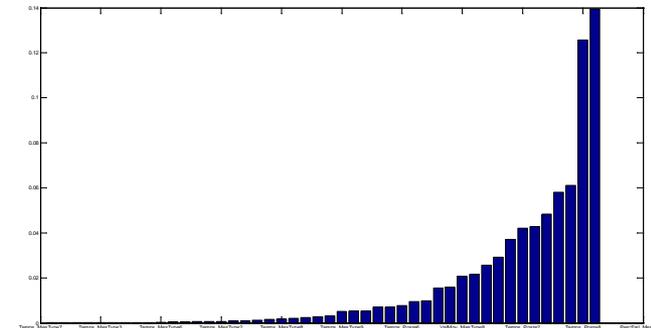
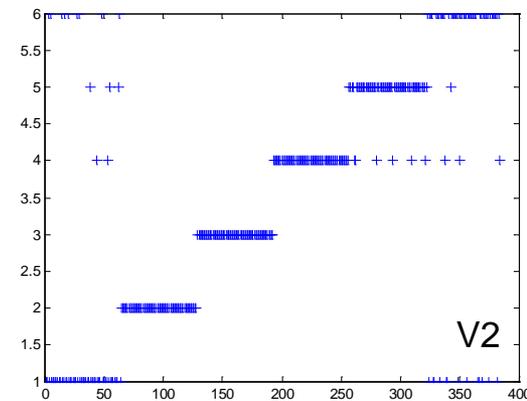
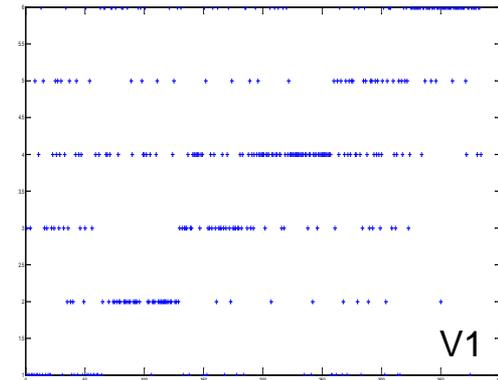
V2 - En construisant une base d'apprentissage corrigée

Sur la base globale, moins de 0,5% d'erreur sur les chaudières sans retour et 8% d'erreur sur les chaudières avec retour

Analyse de sensibilité du modèle V2

Les 13 attributs les plus influents

'CO_Max'; 'CO_min'; 'Gas_Flow_Max_POG...'; 'Bnr_Gas_Pres_Max...';
'CO2_Max'; 'Blr_Gas_Pres_Max...'; 'Gas_Temp_TAG';
'Gas_Flow_Min_POG...'; 'CO2_Min'; 'Safety_Time';
'Blr_Gas_Pres_Min...'; 'CH_Flow_Rate_FR'; 'Bnr_Gas_Pres_Min...'



Plusieurs modèles disponibles sous R, facilement transposables dans un service logiciel pouvant être interfacé avec la supervision d'atelier

Principales difficultés

Faible nombre de retours en proportion & Variété importante des produits

Evolution des gammes

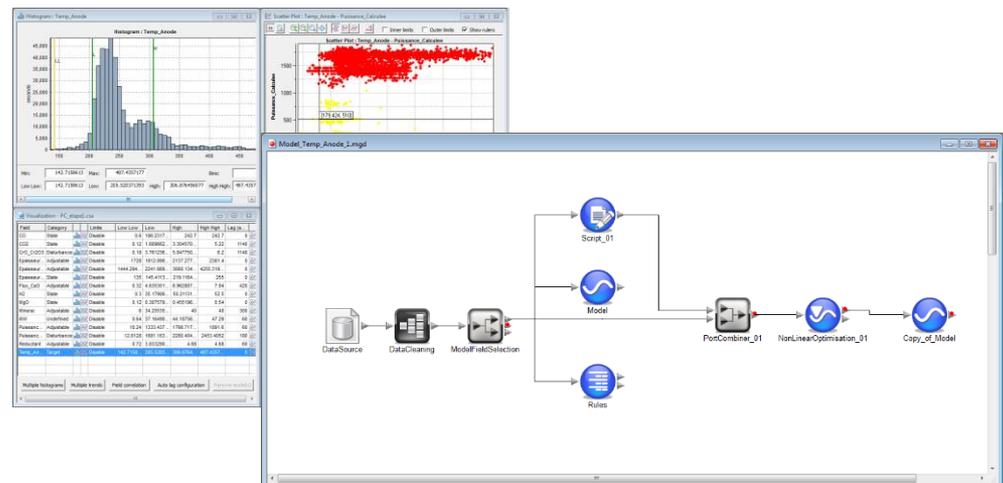
Modifications sur lignes de production ; Etalonnage des systèmes de mesures

Recommandations

Mettre en place un entrepôt de données dédié,

Normaliser les mesures des différentes lignes de production

Reprendre régulièrement les modèles étudiés en les validant et les adaptant aux futures données disponibles



ADAPI : Analyse de Données pour l'Amélioration des Processus Industriels
**À l'intersection entre l'informatique industrielle, le traitement du signal,
l'automatique et les statistiques**

Suivre et améliorer l'existant (gestion courante)

Améliorer le fonctionnement du procédé,
Optimiser les flux,
Améliorer la qualité et la performance, ...

Détecter les défauts au plus vite (gestion des aléas)

Diagnostiquer et prédire l'état du processus,
Pronostiquer une durée de vie restante (RUL), ...

Conçue autour de techniques originales à la frontière entre Automatique et IA

Basée sur des outils industriels, ...

**Supervision et MES, Troubleshooting DataAnalytics
et environnement : Python, ScikitLearn, R, Matlab,
SAS, Hadoop, Spark...**

et la plate-forme nationale TERALAB
dédiée au Big Data

