

LE RENDEZ-VOUS INTERNATIONAL
CONFÉRENCES | ATELIERS | EXPOSITION

THE INTERNATIONAL MEETING
CONFERENCES | WORKSHOPS | EXHIBITION

HPC
SIMULATION
BIG DATA

Forum Teratec

Unlock the future!

11 & 12 JUIN
JUNE 11 & 12 2019

ÉCOLE POLYTECHNIQUE
PALAISEAU-FRANCE

PLATINUM SPONSORS

Atos

DELL
Technologies

Hewlett Packard
Enterprise

intel

GOLD SPONSORS

BECHTLE

cea

CRAY

DDN
STORAGE

Mellanox
TECHNOLOGIES

SILVER SPONSORS

arm

ATEMPO

crisi

FUJITSU

GENCI

Microsoft

NVIDIA

Western Digital.

XILINX.

PARTENAIRE
CAFÉ EUROPÉEN
DE LA RECHERCHE

AVEC LE SOUTIEN DE

Inria
investisseurs du monde numérique

NUMERIC
EUROPEAN
RESEARCH
*IledeFrance

Predictive maintenance solution without additional sensors



Christophe BIERNACKI, Head of MODAL research team , INRIA

Margot CORREARD, co-founder of DiagRAMS Technologies (start-up INRIA)



RESEARCH TEAMS

Applied Mathematics, Computation & Simulation

Machine Learning, Statistical Methods

Algorithmics, Programming, Software & Architecture

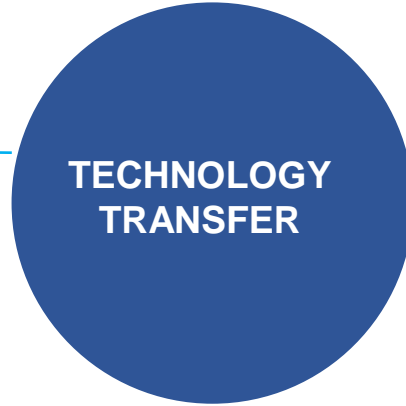
Security, confidentiality, cryptography,

Networks, Distributed Computing

Distributed Systems & middleware, software engineering

Perception, Cognition, Interaction

Data & Knowledge Representation & processing, Visualization



TECHNOLOGY TRANSFER

Joint laboratories

(joint labs, innovation labs, labcoms)

R&D partnerships

(collaborative projects)

Technology transfers

(software and patents)

Transfer of knowledge/ know-how

(expertise, mobility)

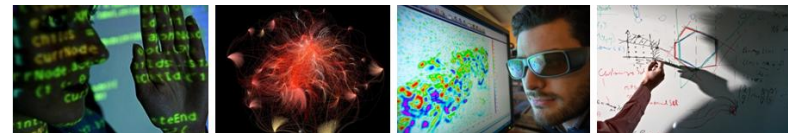


30 years
OF CREATING
START-UPS

120 start-ups

including 75% in activity
or bought out

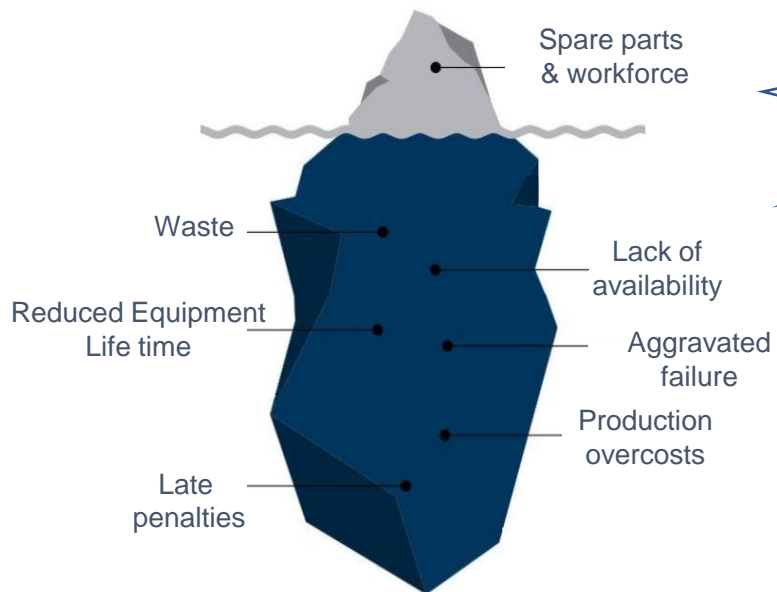
3,000 jobs created



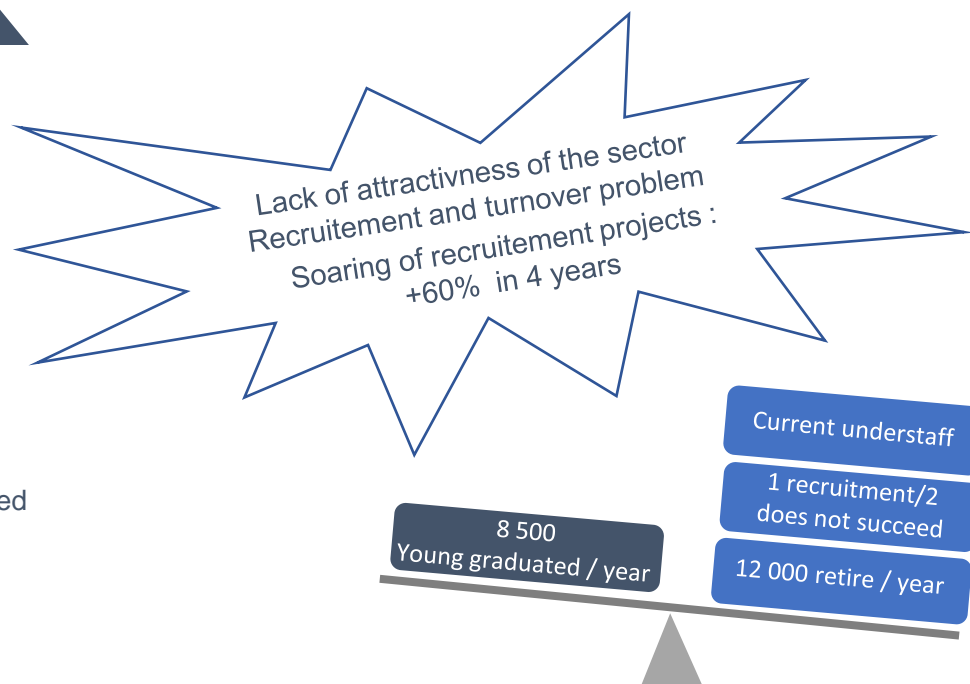
THE RISING COST OF DOWNTIME

\$260 000

Average Cost per Hour of Downtime*



RECRUITMENT PROBLEMS



* The Aberdeen Group, 2016

OUR UNIQUE VALUE PROPOSITION



Data acquisition without
new sensors



✓ Reduced implementation cost



Monitoring of overall
equipements



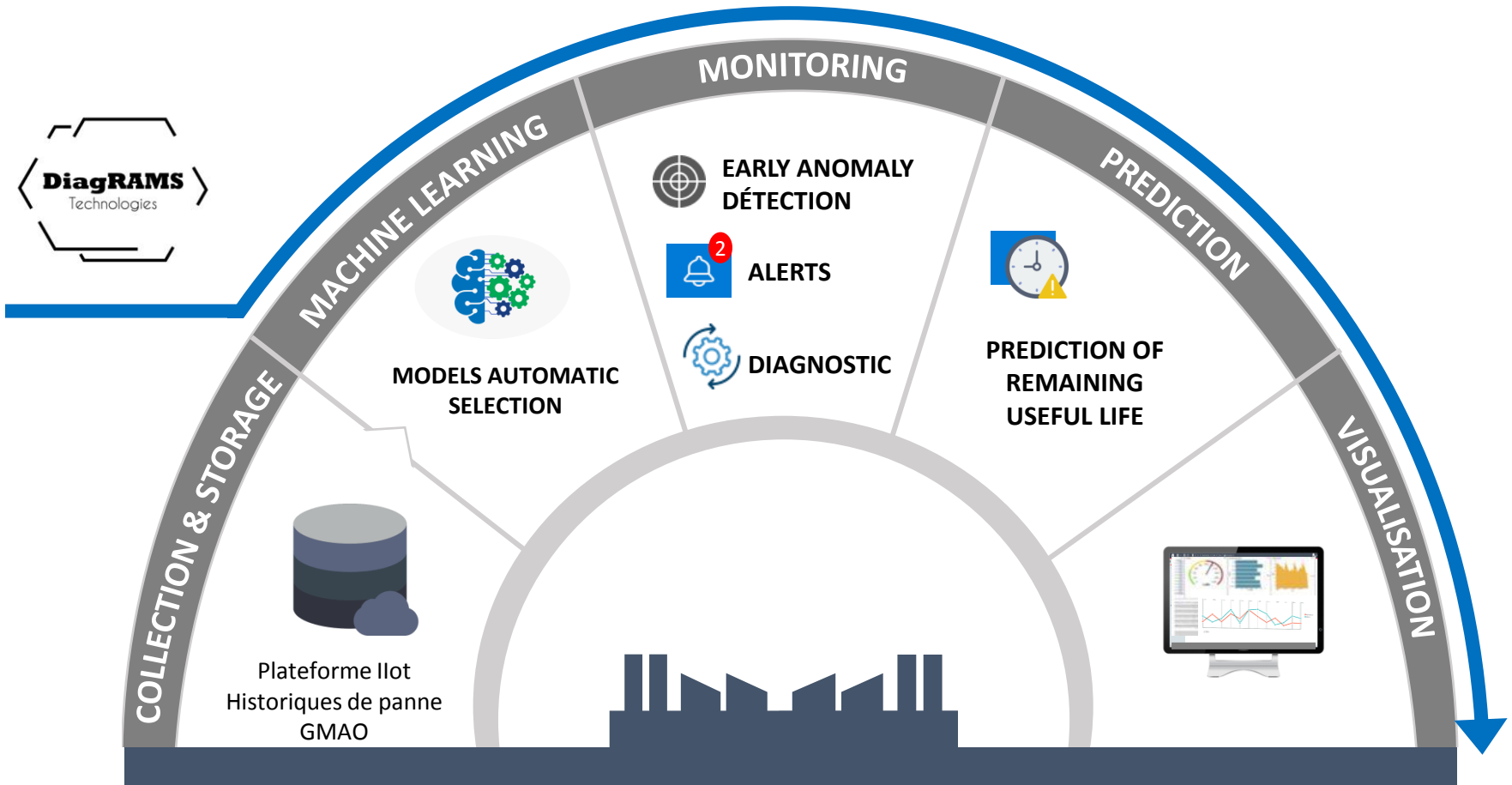
✓ Only one tool



Turnkey solution for
field experts



✓ Acceleration of technician
competence acquisition

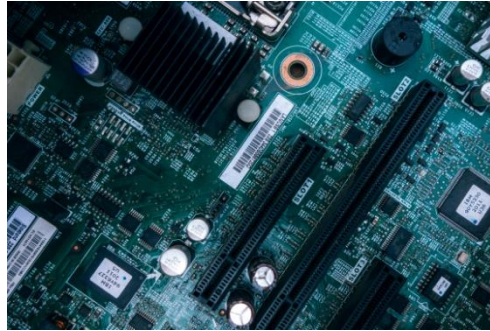




NOKIA



ALSTOM



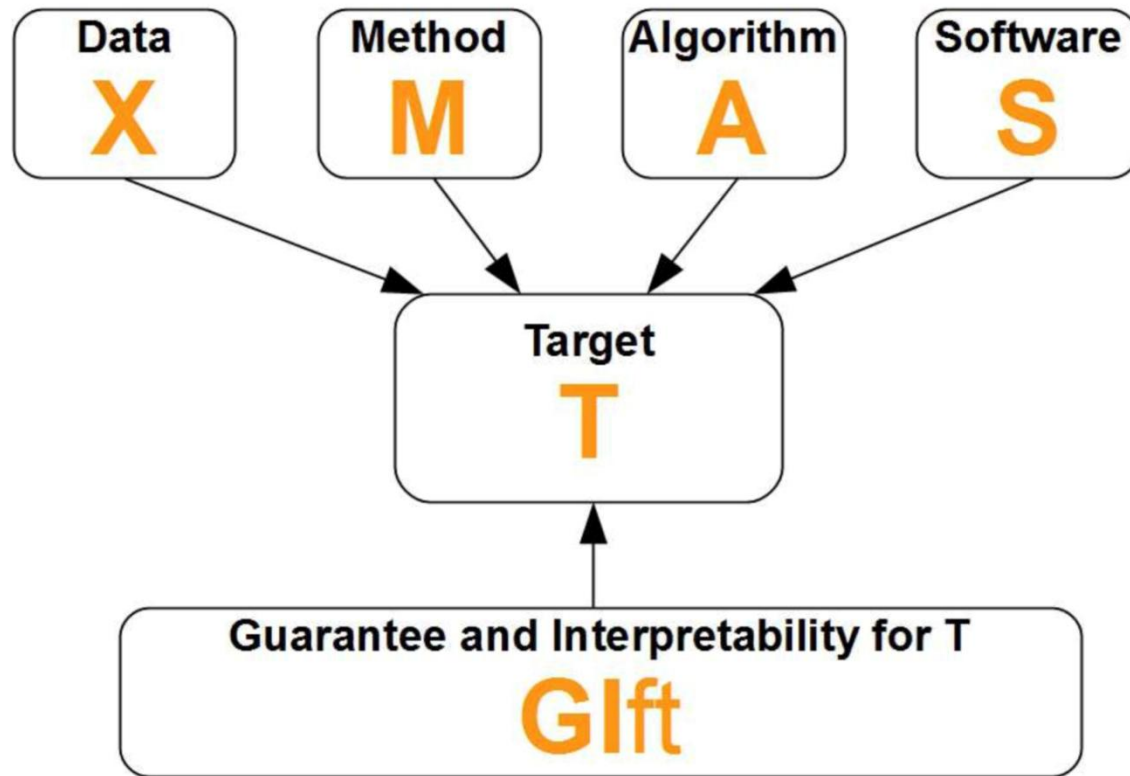
APSYS
AN AIRBUS COMPANY

MODAL

Generic learning from complex data



Modal: guarantee & interpretability



Probabilistic guidelines (\mathbf{X} random...)

- **Model-based**: define a model directly on data \mathbf{X}

$$p(\mathbf{X}; \theta) \longrightarrow \hat{\mathbf{T}}_{\mathbf{XMAS}}$$

- Model design should be guided by both space-adequacy and user-meaning
- Then use model estimation and model selection provided by theoretical statistics

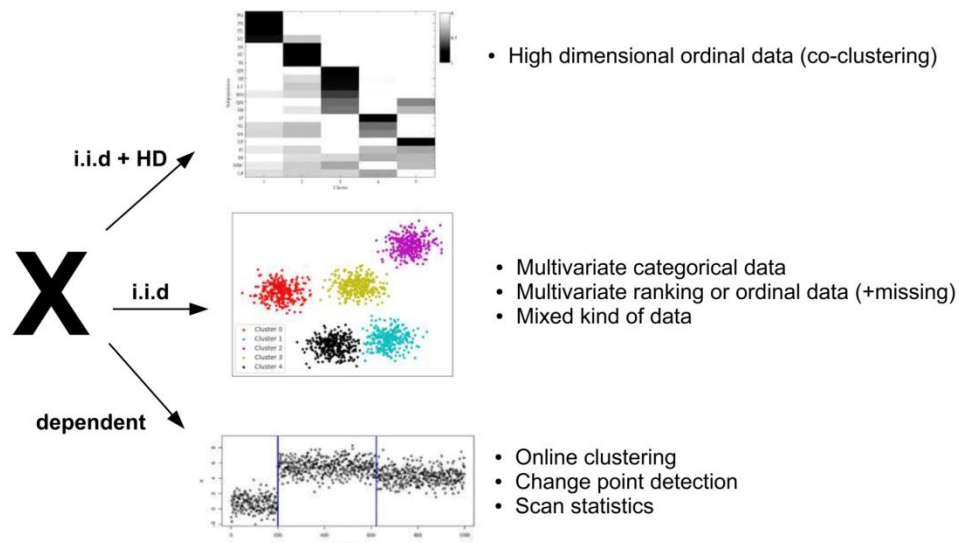
- **Model-free**: define probabilistic bounds on empirical (\mathbf{X} -dependent) quantities

$$p \left(\text{loss}_{\hat{\mathbf{T}}_{\mathbf{XMAS}}} \leq \widehat{\text{loss}}_{\hat{\mathbf{T}}_{\mathbf{XMAS}}} + \text{residual}(\delta, \dots) \right) \geq 1 - \delta$$

- Principle of concentration inequalities
- Context of the probably approximately correct (PAC) paradigm

Axis 1: unsupervised setting

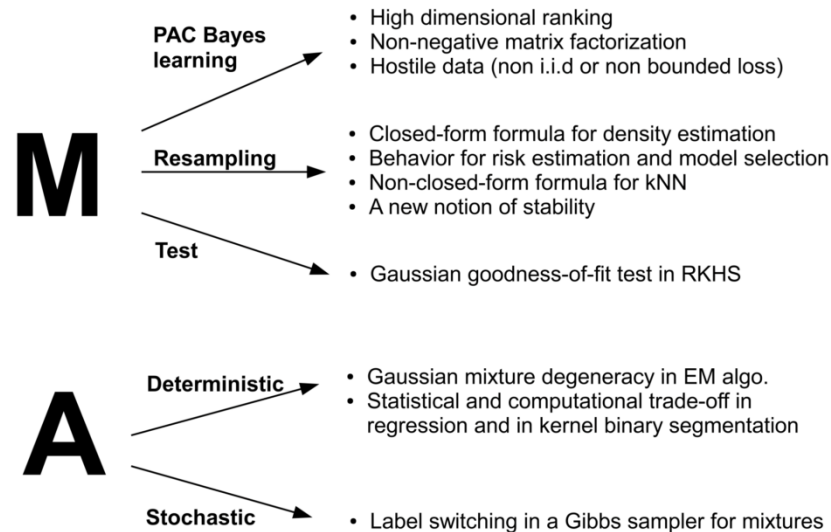
The larger the data set, the stronger the motivation for **discovering knowledge**



Mixture models, RKHS, PAC-Bayes, copula, . . .

Axis 2: performance assessment

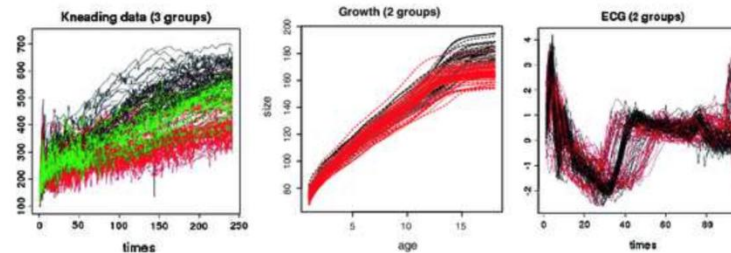
Human activities rely increasingly on statistical outputs: **need to have guarantees**



EM-like, gradient descent, Gibbs sampler, kNN, regression, ranking, NMF...

Axis 3: functional data

Time series are now **widespread** because of technological advances



T

Clustering

X

- Categorical functional data
- Multivariate functional data
- Spatial functional data

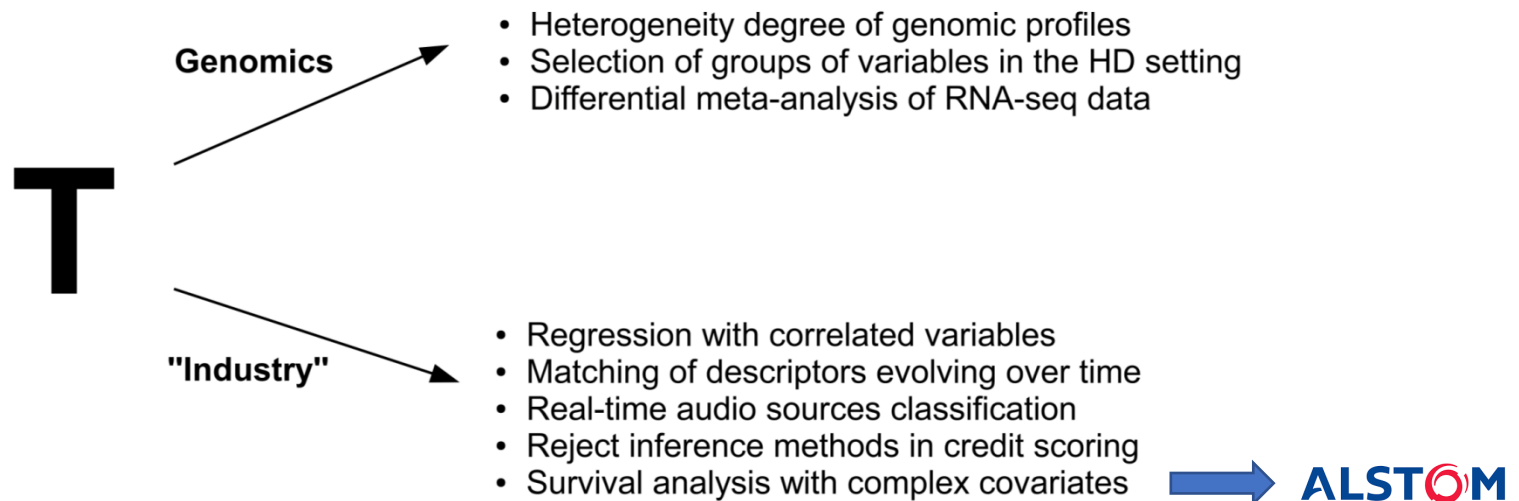


Prediction

- Spatial prediction of solar energy
- Regression with auto-correlated data
- Binary linear models for stratified samples
- Functional linear spatial autoregressive models
- Application to hydrological or environmental data
- Auto-associative models

Axis 4: applications motivating researches

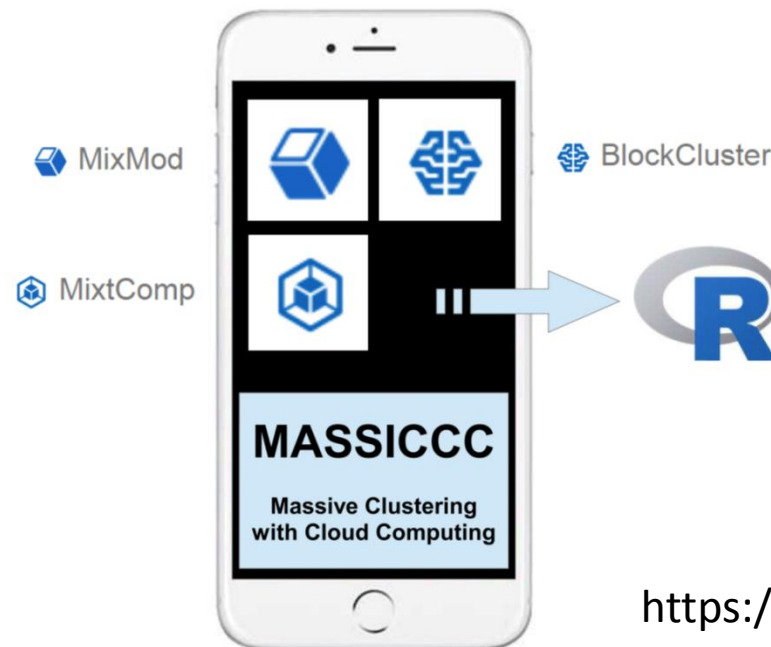
Locks in practical activities may reveal sensible **applied research** questions



Software as an output

Platforms

The key for integrating in a user-friendly way the most mature packages/software



<https://massiccc.lille.inria.fr>



Alstom collaboration n°1

Automated diagnosis for maintenance

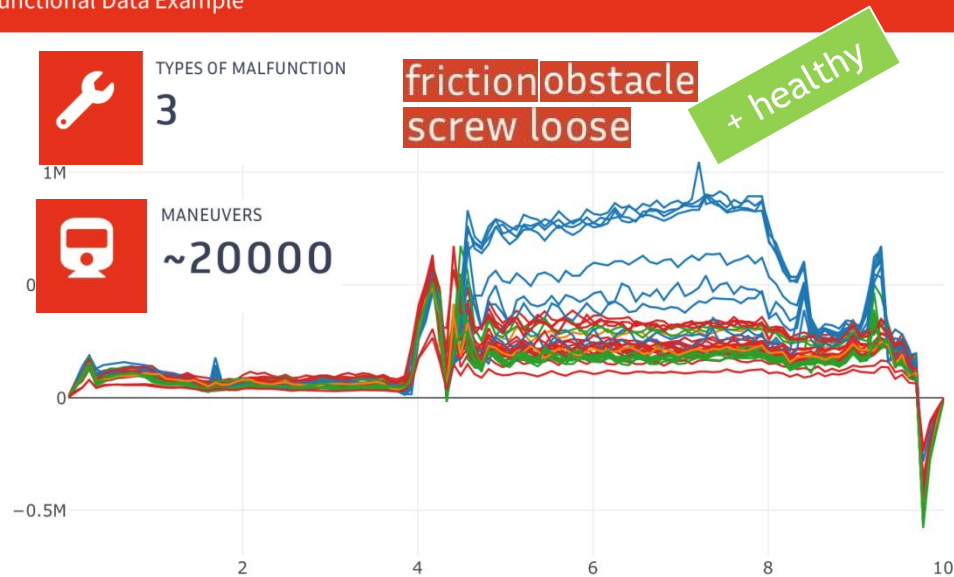
Detection & diagnosis of rail switch malfunctions

power consumption of the needle motor of a rail switch

power consumption evolves along the movement (10s approx)

Functional Data Example

Data

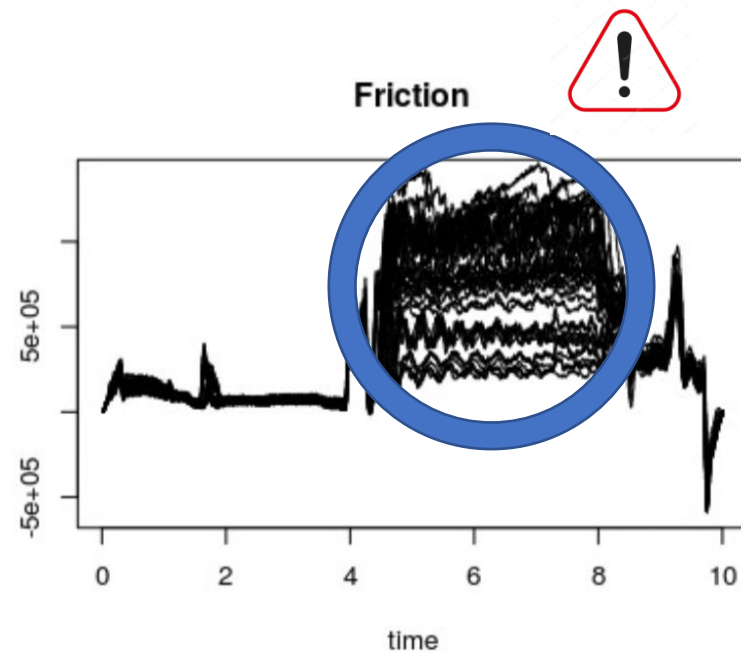


Detection & diagnosis of rail switch malfunctions

1) Functional data are complex data

2) Heterogeneity of malfunction types

Locks



Model-based functional data

Idea of the model implemented in MixtComp

A time series/functional datum is composed of L polynomial segments (+ some noise)

- Interpretability
- No feature extraction
- Model selection

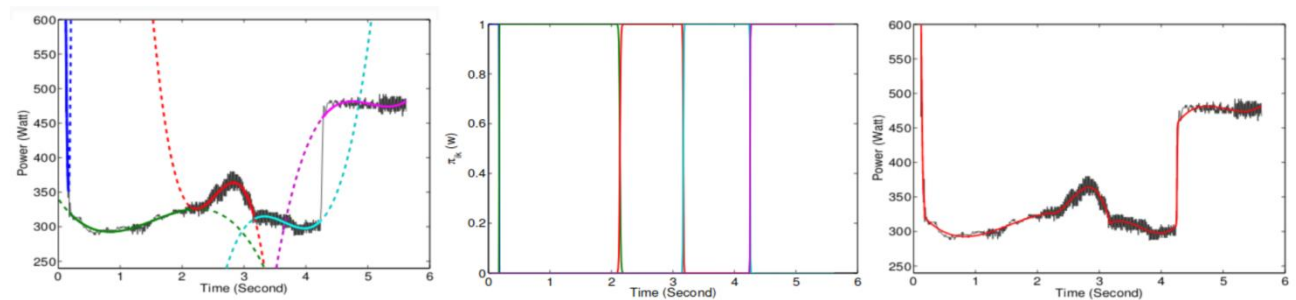


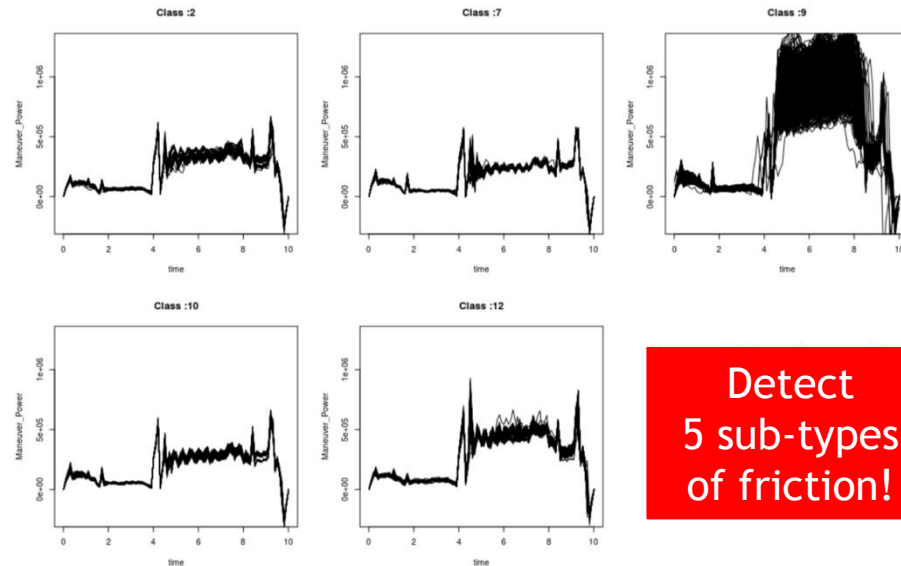
Figure: From left to right: decomposition in 5 polynomial segments, the probabilities to belong to a segment, the automatic segmented signal.

$$f(y|t; \Theta) = \prod_{j=1}^m \sum_{l=1}^L \kappa_l(t_j; \alpha) \mathcal{N}(y_j; \beta_l^T T_j, \sigma_l^2)$$

Model-based prediction of malfunction types

- 1) Probabilistic modeling of curves for guarantees
- 2) Unsupervised classification intra-malfunctions

Learning



Detect
5 sub-types
of friction!

Model-based prediction of malfunction types

1) Diagnosis ability with risk assessment

2) More knowledge on malfunctions (sub-types)

Results

Confusion Matrix


% CORRECT PREDICTIONS
85.2

	Prediction: Friction	Prediction: Healthy	Prediction: Obstacle	Prediction: Screws loose
Friction	584	0	0	0
Healthy	0	192	40	47
Obstacle	25	47	1368	203
Screws loose	29	61	164	1402

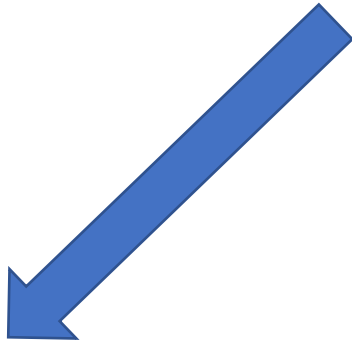
To be compared to 80.7% obtained with expert manual feature extraction...

From diagnosis to prognosis...


$$P(T > t|\mathbf{x}) = \sum_{k=1}^K P(T > t|\mathbf{x}, k)P(k|\mathbf{x}) = \sum_{k=1}^K \underbrace{P(T > t|\mathbf{x}, k)}_{\text{key assumption}} P(k|\mathbf{x})$$



Result is high flexible and interpretable pronostic



Use any usual pronostic method (comparison...)



Detect more refined malfunctions types



Alstom collaboration n°2

Free text maintenance database

Free text maintenance database

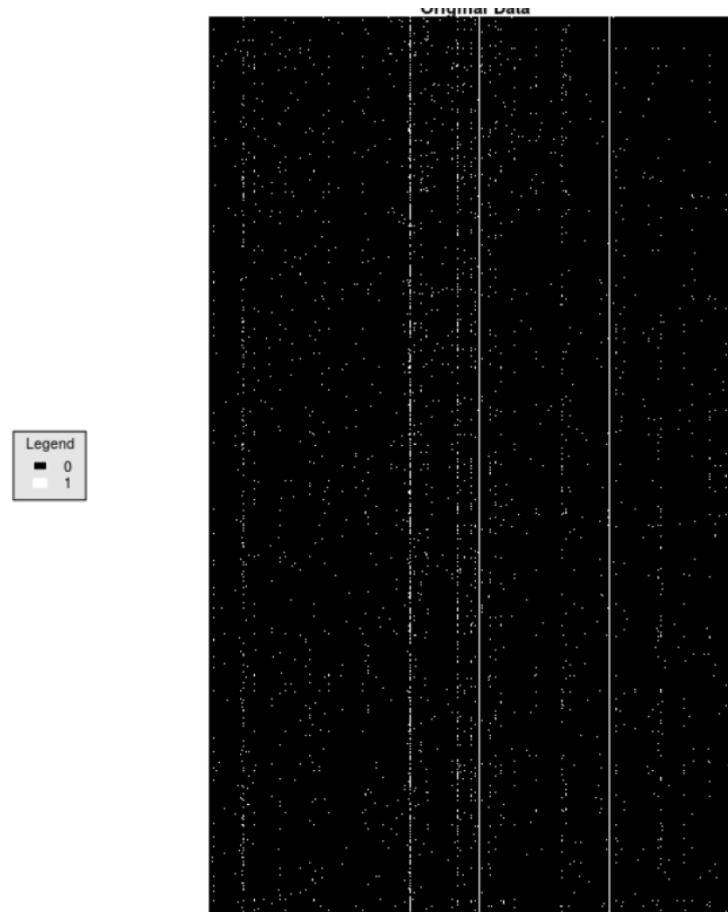
- **Data:** 88,190 maintenance notes
- **Aim:** group similar maintenance events
- **Lock:** free text
- **Key:** synonyms (Alstom) + co-clustering (MODAL)

D1_ref	data_cleaned
213	hi 1435 sri 0274 fire whilst in service class 175 incident summary and initial findings provided by uk mainline engineering manager c english 26 08 09 reported fire on 175109 at prestatyn station
286	propulsion inverter electronics agate d1 impact on service and maintenace activities due to remove number of pie from train failed pie causes train to be withdrawn from service d5 completed
1422	scrubber break up in operation lost of one part of the brake shoe system securite perturbation du service commercial d5 completed d6 completed
1536	coupler damages on front coupler support damages on coupler support side panels gets broken broken supports might create problems at coupling process d5 completed d6 completed
1570	cracks found on the cross beam connection to the side beam cracks found on the cross beam connection to the side beam reduced lifetime d5 completed complete rework history book has beer
1572	wheel bearings with current flow electrical marks in june 2008 the customer sl reported possible signs of current flow across wheel set bearings at axle 7 of the train sets x6007and x6023 first vis
1689	roof fairing fixing found not tightened on train roof fairing fixing found not tightened on train cis006 safety of service root causes identified d6 completed
1835	wcml pim new build poor quality of frames and all associated sub assemble from suppliers why missing d5 completed d6 completed
1943	out of round wheels soft supplier lucchini have confirmed that the r8t wheel steel issue has been confined to a single batch of wheels that have resulted in early wheelset changes there is no fur
2042	customer complaint late delivery of spares high levels of rejections at customer inspection at preston and in hong kong concerns regarding preston process controls and preston supplier controls
2099	axle box interior grease leakage discovered a leakage of grease the bearing unit used is a tbu reference 1639609 aa 02 mileage about 260 000km maintenance interval suggested by skf 800 000k
2101	toshiba vcb late delivery of sets 33 47 initial order for 32 off 50kv vcbs for toshiba locomotives is virtually complete after many issues with design type test supplier quality areva interrupters and
2126	cracks on trailer bogie frame of roma 1 tram cracks appeared on tram roma 1 trailer bogie frames welding between the longeron and the transom these frames has been already retrofitted on h

Synonyms dictionary (Alstom)

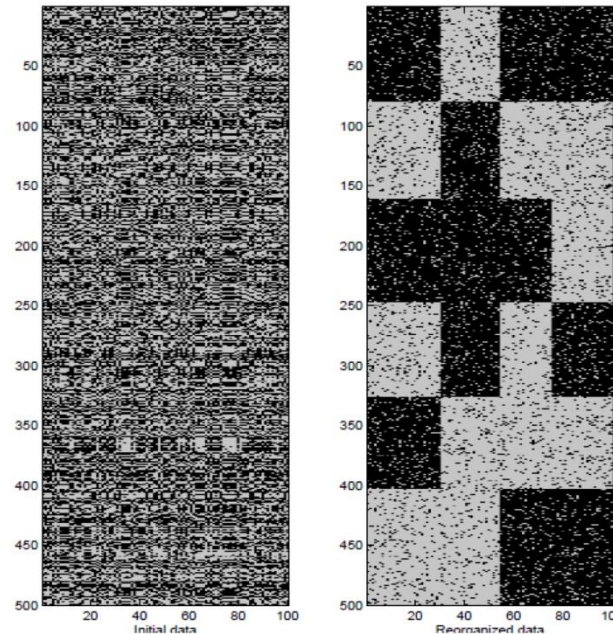
D	synonyms (1)	synonyms (2)	synonyms (3)	synonyms (4)	synonyms (5)	synonyms (6)	synonyms (7)	synon
1	HVAC	HVAC	Air conditioning	Ventilateur	Ventilator	clapet valve	Compresseur	Comp
2	Assemblages	Catapultage	detachment	detached	Fissure	crack	Vis	bolt
3	Bogie	Bogie	Drehgestell	Bogie Supporting Structures	Chassis	essieux	axles	Suspe
4	power supply	transformer	pantograph	collector shoes	traction battery	energy storage	storage	drive u
5	driving desk	pupitre	control apparatus	driving control	braking control	master controller	Traction Control Unit	Syster
6	Auxiliaires	Auxiliary Operation equipment	Auxiliary Converter	Convertisseur	Auxiliary battery	Emergency Power Generator	Generation	Coolir
7	Safety equipment	Safety device	Data transmission	data communication	Fire protection	extincteurs	equipement incendie	
8	lighting	lights	eclairage	lampe	ampoule	phare		
9	Portes	doors	step	marche				
10	panneau d'affichage	screen	information facilities	display	PIS	PAPIS	haut parleurs	speake
11	Intercirculation	Intercar	couloir	gangway	coupler	buffer	coupleur	
12	cablage	harnais	electrical connexion	earthing	cables	harness	wire	junctio
13	GAUGE	gabarit statique	EN 15273	gabarit cinématique	UIC 505	gabarit dynamique	obstacle	obstac
14	sanding	sable						
15	flange lubricating	graisseur de boudins						
16	Passenger Counting Device	comptage passager						
17	ticketing	ticket fare						
18	Brake	frein						

Transform word dataset into binary dataset



Co-clustering (MODAL)

Blockcluster
software

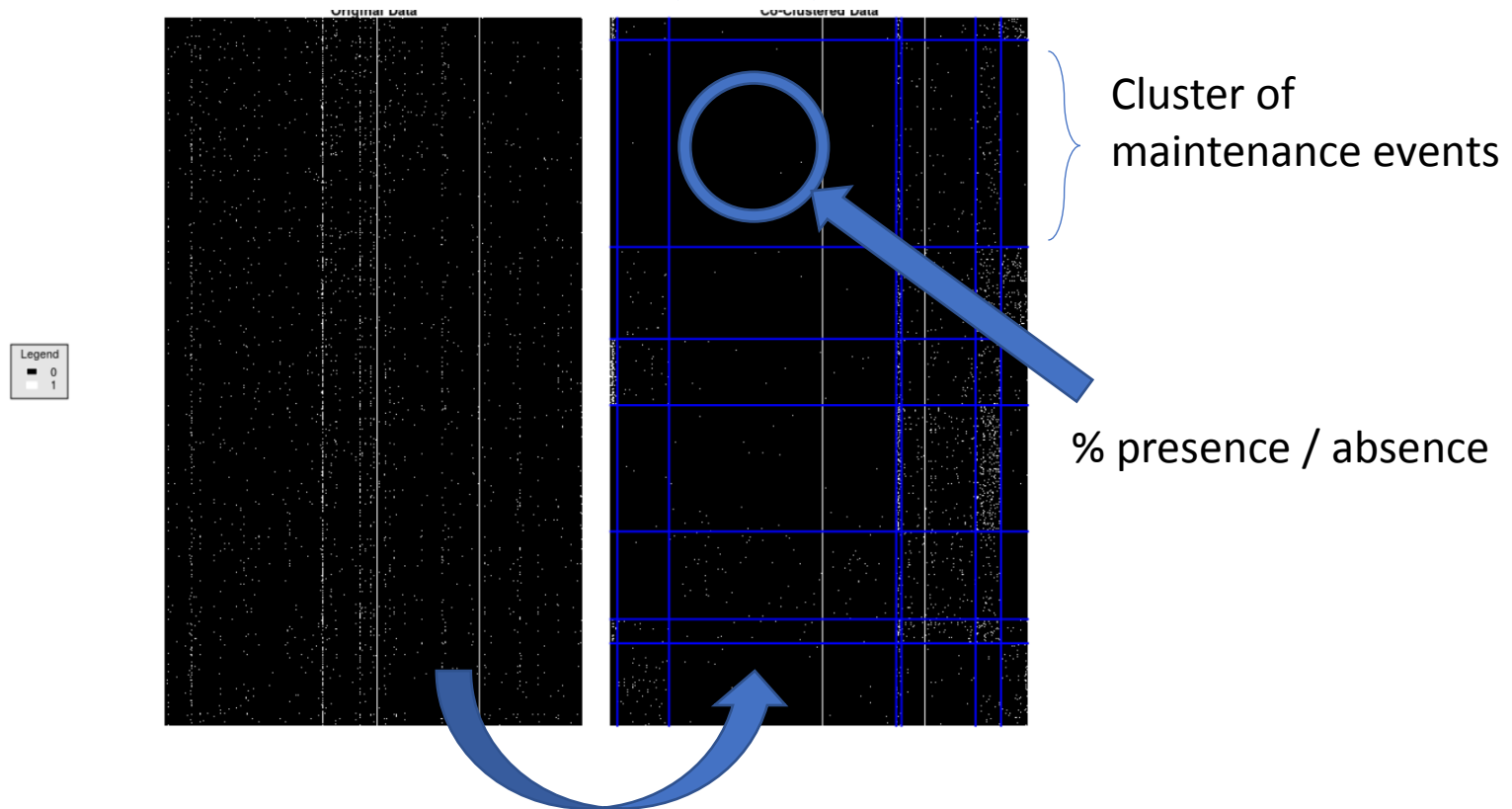


Extension of the traditional mixture model-based clustering

$$p(\mathbf{x}; \theta) = \sum_{(z, \mathbf{w}) \in \mathcal{Z} \times \mathcal{W}} \prod_{i, j} \pi_{z_i} \rho_{w_j} p(x_{ij}; \alpha_{z_i w_j})$$

Cluster maintenance events **and** synonyms

Cluster of synonyms
for interpretability



Conclusion

MODAL

- Generic learning for complex data
- Related software: <https://massiccc.lille.inria.fr>
- Multiple applications, including maintenance
- Complex data in maintenance

DiagRAMS

- Partnership with MODAL
- Partnership with industry
- Best of both worlds...



MERCI

christophe.biernacki@inria.fr

margot.correard@inria.fr

