Exploration of ML algorithms towards anomaly detection for the Arronax accelerator operation

Lead: Freddy Poirier (Arronax), Diana Mateus (LS2N)
+ 2024 Work from: F.Basbous
Previous work from M2 students: O.Cooper, C.Lassalle, J.Rioux

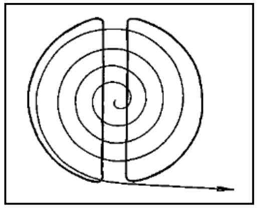
ARRONAX: Accelerator for Research in Radiochemistry and Oncology at Nantes Atlantique.

M4CAST- Nov 2024

ARRONAX Cyclotron Activities

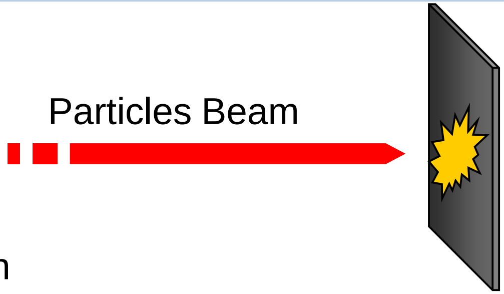
- A cyclotron to produce radionuclides for research in nuclear medecine (50%):
  - Imaging: β+ radioelements for PET (ex: ^{82}Sr/^{82}Rb, $^{44m/44}$Sc, ^{52}Fe, ^{64}Cu ...)
  - Therapy: α immunotherapy (^{211}At → preclinical phase), β^- radioelements : ^{64}Cu (preclinical phase), ^{47}Sc
- Also a tool for R&D on (50% of op):
  - Physics, cross section measurements, radiolysis, radiobiology studies (eg flash biology), archeology
  - Detectors developments (PEPITES, DIAMMONI, Gafchromics, space detectors,…), machine
- A tool for training and education

Radioelement for therapy



Acceleration

Particles Beam



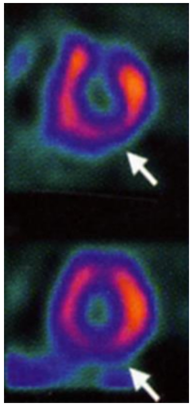
Target irradiation

Production of radioelements
In the target by transmutation

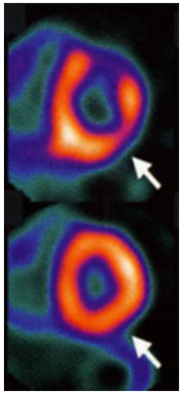
Chemical extraction
Separation
Purification
Marking

Radioactive medical element (experimental)

Tomographic imaging of the heart



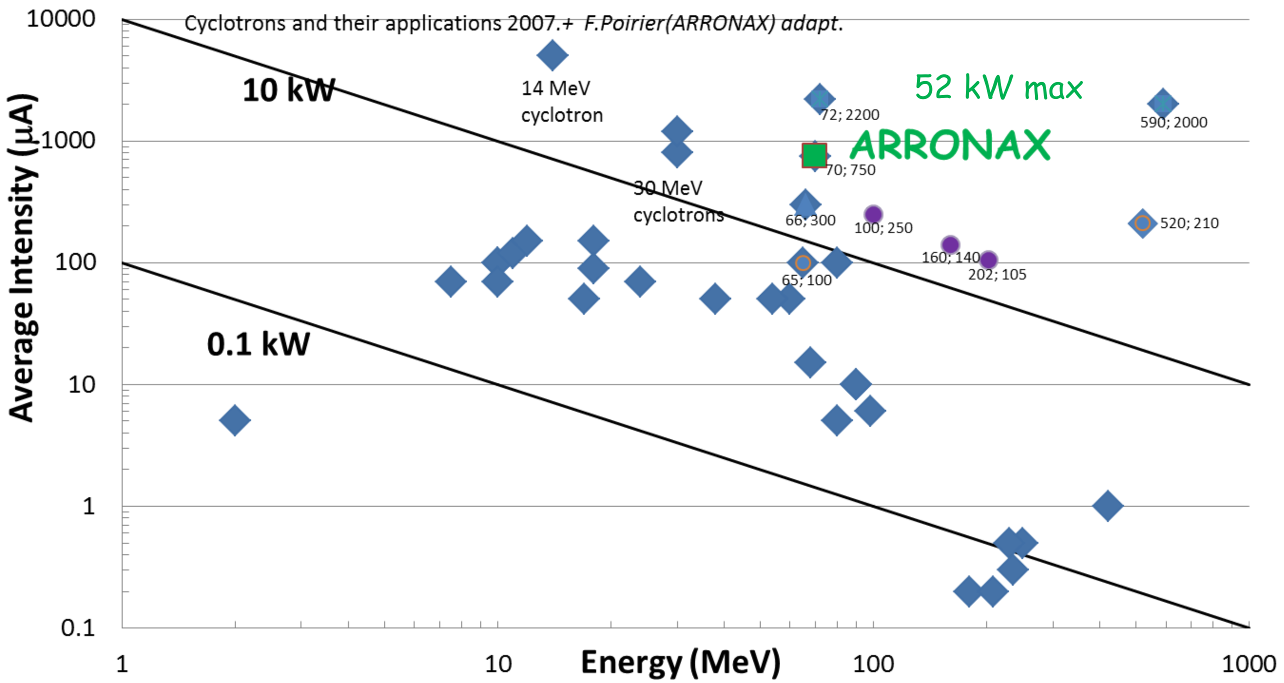
^{99}Tc-MIBI SPECT



^{82}Rb-PET

D. Le Guludec et al, Eur J Nucl Med Mol Imaging 2008; 35: 1709-24

ARRONAX cyclotron among other machines



Extracted Particles	Energy range (MeV)	current (µAe)
H+	35 - 70	375 x 2
He2+	70	70
HH+	35	50
D+	15 - 35	85

- High power machine for fixed target.
- Several similar machine (proton) are being used (or constructed) throughout the world for radionucleides production or even injectors for isotopes separation

<u>At Arronax:</u>

- Arronax beam time access to operation, experiments, are available via collaboration.
- Accelerator data (+experimental/environmental data), scada and curation codes can also be available via collaboration.
- Arronax foresee within the next 3 years to open run time to dedicated ML studies
  - For a certain number of hours/year
  - Through a selective committee
  - With the collaboration of local team (R&D group + phD student)

Irradiations

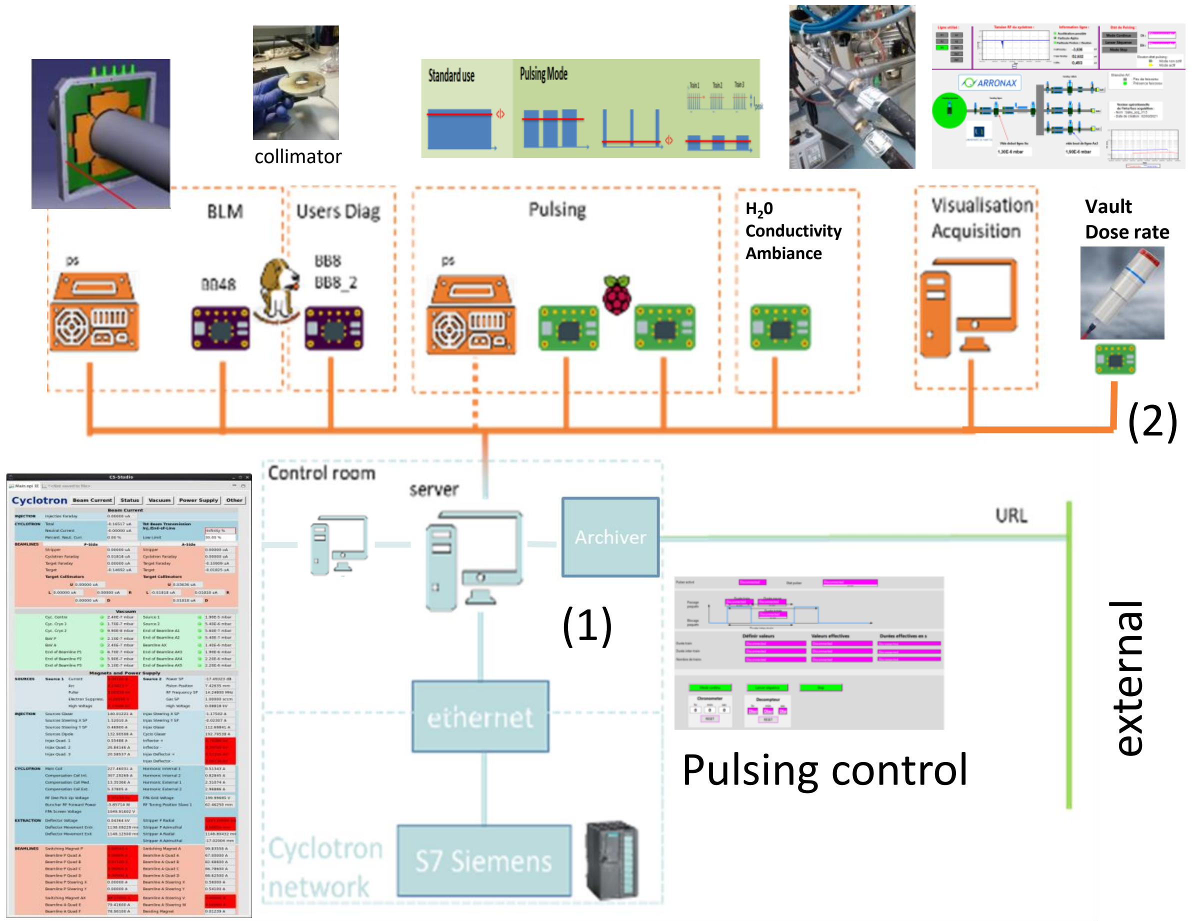
- A production run can last few weeks and leads to a yield in mCi/uAh (a rather experimental standard value)
  - This number can fluctuate depending on the condition due to the machine and beam itself, the following chemistry processes (treatment and purification, losses,…)
- A run is:
  - Mainly on a single beamline, at a time (now become dual beams).
  - Operation protocol is basically to keep the intensity on target, dimension on collimators, and beam transmission from the injection to the target (~43%) and avoid losses
  - Various parameters are being followed:
    - Machine settings and diagnostics and environmental parameters
    - Ex: cooling water average conductivity evolves during the runs and follow a typical progression but can also undergo timed increase:
      - » Adding deionised water due to losses
      - » Released species in the water from aging pipes or damaged target
  - In addition: various breakdowns can happen during the runs: short and long, wanted or unwanted breakdowns of the machine, planned stops
  - Ambiant online dosimetry with ionization chamber in the vault is an indirect indication of the evolution of a run

Data aquisition

- Starting in 2016, we installed an acquisition system based on EPICS (Experimental Physics and Industrial Control System):
  - Parallel to the existing proprietary control software
  - Part of a global and coherent plan to acquire a certain amount of data on the cyclotron, not used for main cyclotron control
  - Allowed additional data from future diagnostics and In-house developments
    - Beam: BLM and other diag (tests in experimental area, collimators,…)
    - Technical: on the environment of the cyclotron (water, gaz, various controls,…)
  - Synchronisation and keep the history on data
  - Allows access to the data remotely
- Initially cosylab installed EPICS on a central server based on an internal Ethernet network:
  - 800 data at 1Hz to 5 Hz on selected parameters/data
- Since then the network has expanded:
  - More data on the cyclotron
  - More systems, more visualisation panel,…
- In 2022-2023, new decentralised servers and an extended archive system has been introduced:
  - Keeping work on cleaning and upgrade
- 2024: updates on the HMI (Cstudio to Phoebius): not done yet (tight schedule)

EPICS Network

- Distributed over the controlled and technical area:
  - Quite a few simple CPUs (BB, Raspb.)
- Centralised on a server
  - Archiver (data secured)
  - Independent PC
- First part of the upgrade:
  - in-house dvt (1)
  - New Tech Diag (2)



Main window

Some questions for our production runs

- What question can we ask using data mining or (non) supervised machine learning? (Sometimes it is just a question of starting and see)
- Can we explore algorithms and tools to tell us something about our productions?
- Are there algorithms adapted to the time evolution of data analysis?
- Can we identify operational anomalies during a run (time wise) or unexpected final results?
  - That we know Or not?
  - Associated to operational modification of the accelerator or from specific parameters?
- With automatic learning, can we identify parameters which increase the production?
- Can we steer the algorithms towards parameters choice and more supervision?

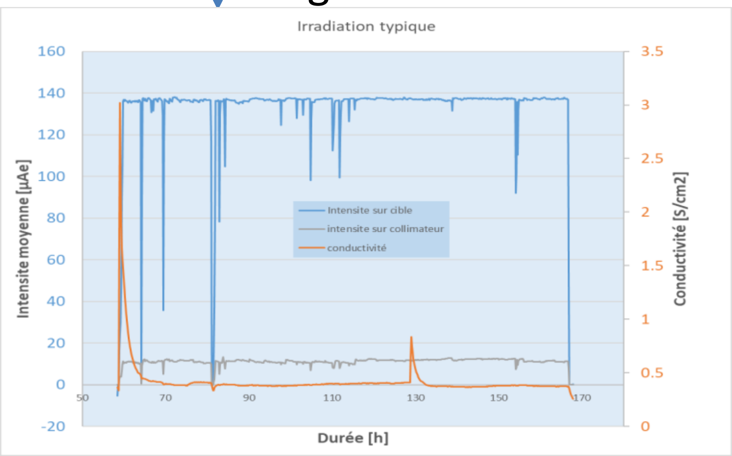
Data assembling and approach of ML

- Inter-services work to assemble data (physics, chemistry, quality, risk prevention group)
- Python codes to retrieve data from EPICS
- Perform analysis with Pandas and scikit-learn and other sofwares and implementation of tests (towards robustness)

Data studies: two extreme approaches

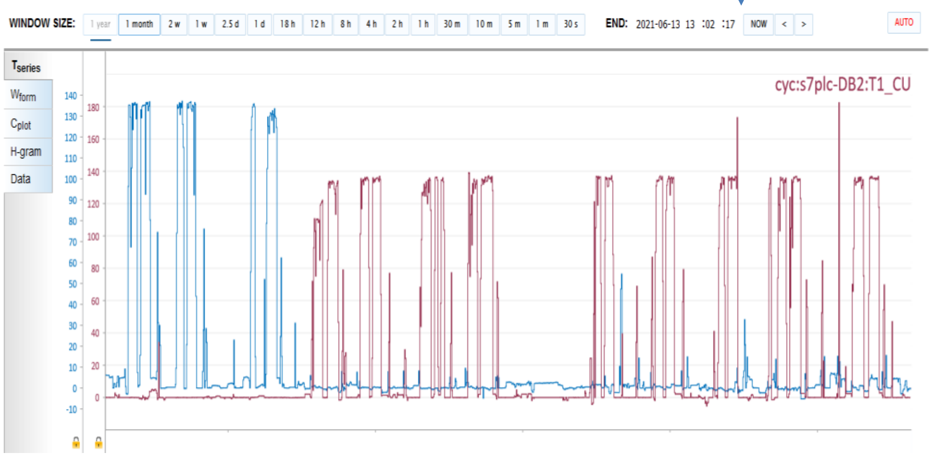
This talk

Single run



Irradiation example over 5 days:, $\langle I \rangle_{cible}$=130µA, $\langle I \rangle_{colli}$=10µA
- Water conductivity
-Source and RF breakdowns = intensity diminution

Multiple runs



Several Irradiations over 1 year: 2 beamlines here

ARRONAX
afaq ISO 9001 Qualité AFNOR CERTIFICATION

Conclusion (intermediary)

- **Preliminary results = Beginning of the exploration of the algorithms**
  - Experimenting on the whole process from data to root cause
- **Result:**
  - **Several algorithms are being employed for the analysis** → towards a selection of best suited algorithms?
    - Shows some interesting results and can help the identification/definition of the outliers
- **Precautions:**
  - The events associated to anomalies have to be connected to real parameters (ex machines): This is not direct or takes time → and here requires in addition **robustness method** (indices, SHAP, evaluation metrics) to eliminate the aboundance of outcomes or unstabilities (sampling strategies).
  - Probably need to dig further in the algorithms (+ more data, multivariates, limits): We are at the beginning (!!!)
- **Future:**
  - Time series:
    - On short term: reconsider/rework some of the present algorithms
    - Temporal Flux to be analysed with more understood algorithms and new algo (?)
    - Define what could be dedicated operational runs and specific to anomaly detection
  - Global studies (statistics):
    - Impact of the results of the rejection of anomalies on the final yields of the analysis (mCi/uAh)
    - Taking into account more information and « impression » from the members of the team

Approach: time series (from 2023 on)

Which algorithm?

- Algorithm exploration for anomaly detection on time series:
  - Rolling windows
  - Pre-processing (raw, scaling,interquartile,…)
  - Multivariate data
- Application of several algorithm:
  - Clustering algorithms :
    - Cluster-based local outlier factor
    - One-class SVM (OCSVM)
  - Density estimation algorithms :
    - Isolation forest (Iforest)
    - Gaussian mixture model (GMM)
    - Local outlier factor (LOF)
  - Distance algorithms :
    - Nearest Neighbours

Which Evaluation Metrics?

- Receiver Operating characteristic curve (ROC curve)
  - Plot of the true positive rate against the false positive rate, at various threshold settings
- F1-score
  - Harmonic mean of the precision and recall
- Precision Recall curve (PR curve)
  - Plot of the precision against the recall, at various threshold settings
- Mass volume (MV*) and Excess mass (EM)
  - scaled to subsets

Testing out Algorithms and Metrics

* Goix, Nicolas (2016). How to Evaluate the Quality of Unsupervised Anomaly Detection Algorithms? arXiv: 1607.01152 [stat.ML]

Choice of algorithm and testing

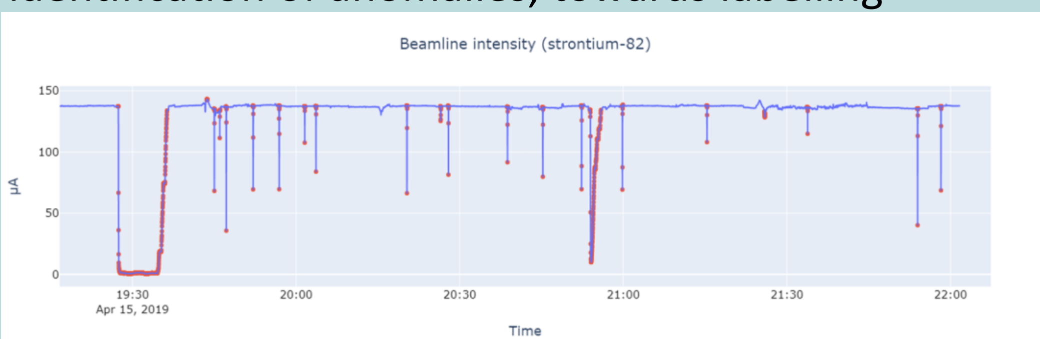
Algorithm and metrics studies:

- Metrics point to some algorithms that should perform better: eg CBLOF model
- Not all metrics are relevant here

Model	ROC	PR	F1	EM	MV
CBLOF	0,996	0,991	0,948	0,007	24582
GMM	0,993	0,986	0,960	0,005	24582
IForest	0,988	0,975	0,585	0,005	24582
KNN	0,961	0,960	0,300	0,005	382
LOF	0,907	0,894	0,299	0,072	46
OCSVM	0,987	0,985	0,300	0,005	382

Classification results. The best algorithms are in green and the worst algorithms are in red

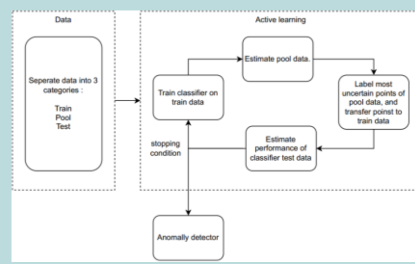
Output example of typical CBLOF model identification of anomalies, towards labelling



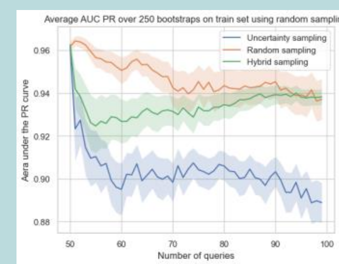
Semi-Active learning:

- Algorithm is trained on an initial train dataset then evaluates a pool dataset and selects the best observations to query
- Iterative process where sample is selected within the unlabelled data pool:
  - Labbeled is performed and sample included in new training data
  - Model is updated
  - 50 queries before terminating the active learning loop
- Used of various strategy for the sampling (basically are you sure that it is an anomaly?)

Active learning loops methodology



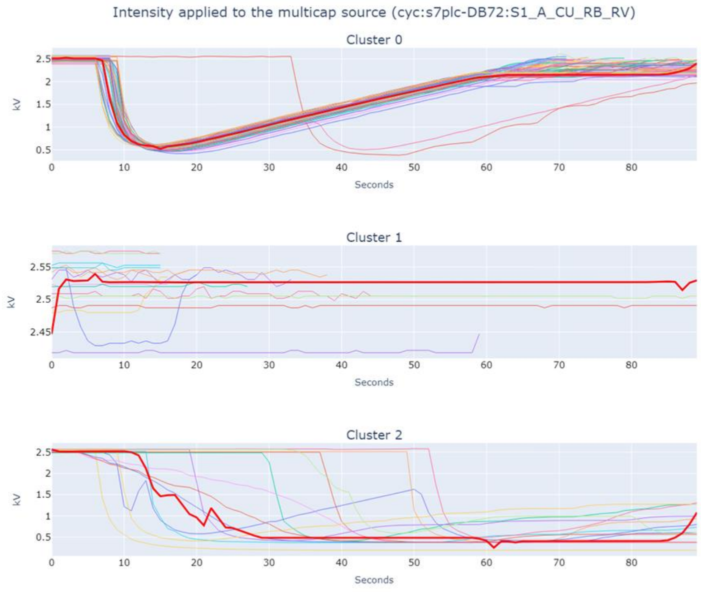
PR curve after 50 active learning loops



Rather stable results but time consuming

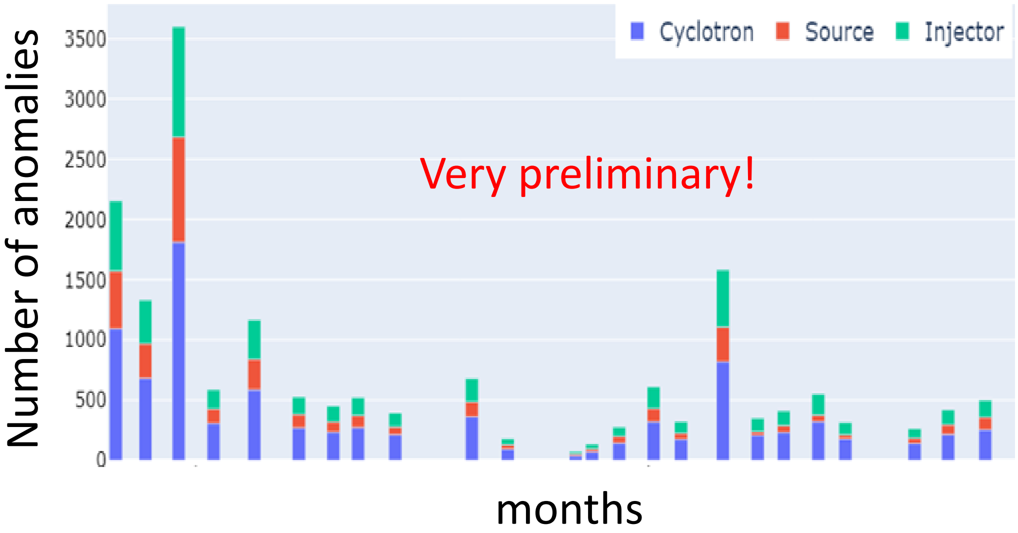
Root cause: towards identification

- Clusters of anomalies



Specific anomalies have been indentified. Here only 3 clusters selected

- Counting the anomalies

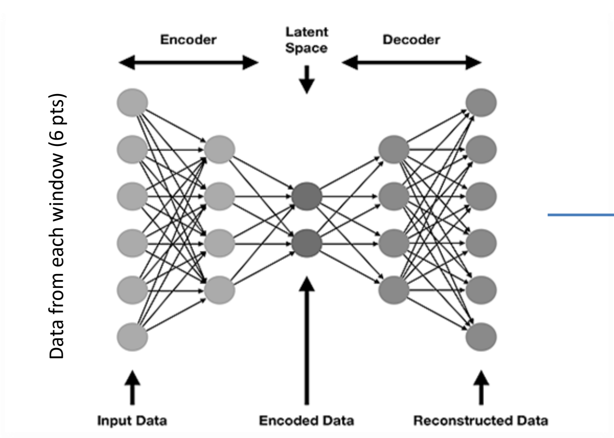


The present method: Statistical and traditional ML. What about the neuronal network?

- Work from PhD F.Basbous:
  - Redefinition of the anomalies in a context of a simplification for labellisation and check of training biaises
  - Application of neural networks model
    - To be compared with other “traditional” methods
    - Can be applicable to multivariate data

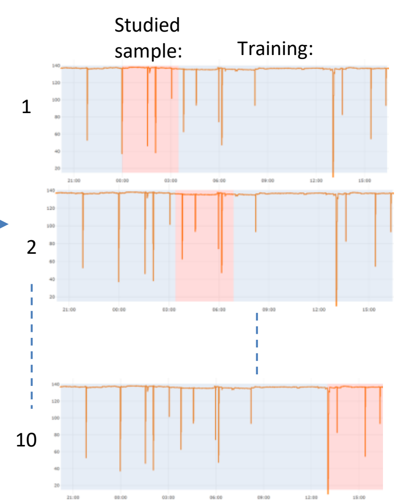
Autoencoders

- Extension of the work to neural network such as autoencoders:
  - “Said” to be particularly well adapted to anomaly detection due to their ability to learn data representations and reconstruct inputs
  - Constituted of:
    - An encoder that compress and store the data input in several latent variables
    - A decoder that reconstruct (decompress) as reliable as possible.
    - A number of layer correspond to a number of latent variable



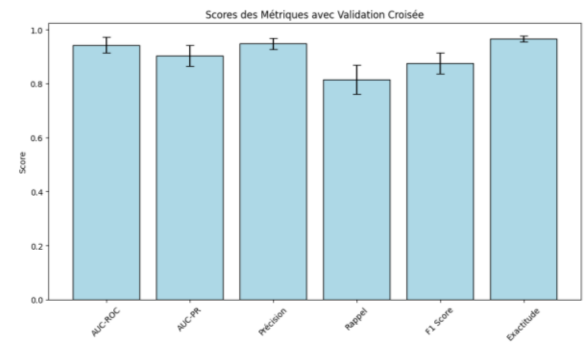
This middle layer does the extraction of the data characteristics (hyperparameters)

Data are divided in subsets (training & data)



Cross validation: mixture of training and sampling

Algorithm performance



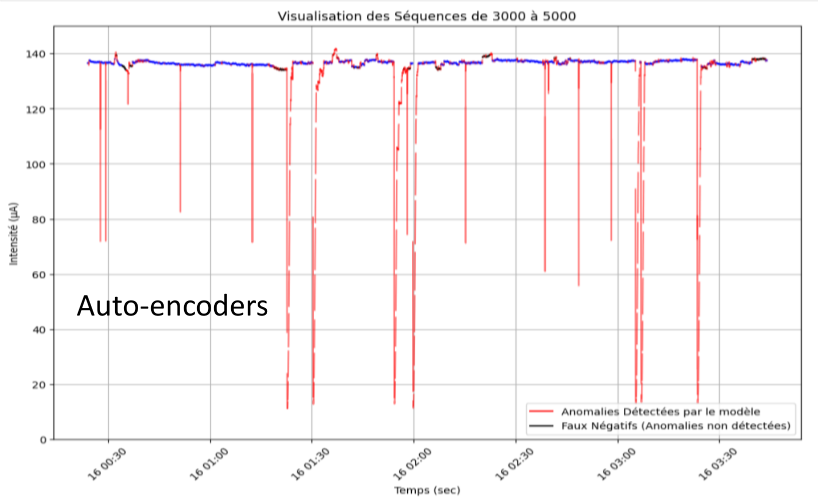
Metrics

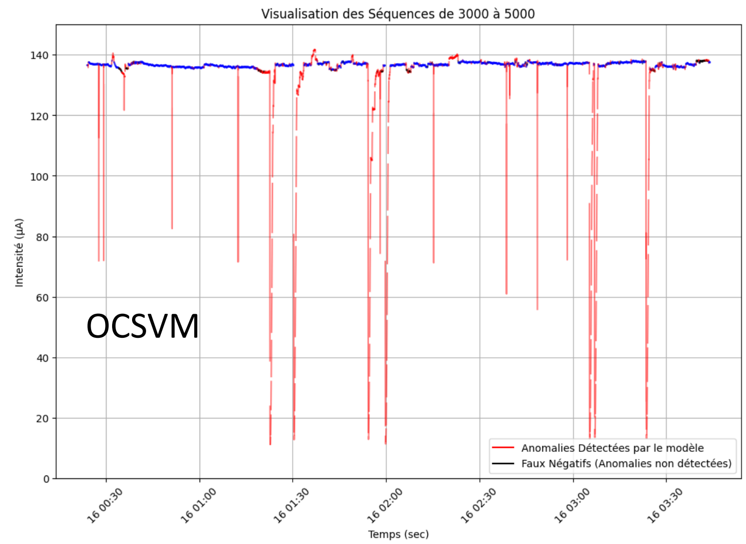
To make “more” sure we are less dependent on data/training

Found Anomalies

- As a visual check:

preliminaries





Auto-encoders

Métriques	AUC-ROC	AUC-PR	Précision	Rappel	F1-score	Exactitude
Score	0.9429 ± 0.0295	0.9036 ± 0.0394	0.9480 ± 0.0202	0.8148 ± 0.0548	0.8758 ± 0.0385	0.9655±0.0100

OCSVM

Métriques	AUC-ROC	AUC-PR	Précision	Rappel	F1-score	Exactitude
Score	0.9532 ± 0.0033	0.8705 ± 0.0054	0.8311 ± 0.0083	0.8000 ± 0.0168	0.8151 ± 0.0084	0.9453 ± 0.0029

More correct prediction (i.e. an anomaly is really an anomaly)

Visually: not such a big difference
But details counts: there are distinct outcomes

This needs to be better understood

Our commitment (Road Map)

- Priority on:
  - 1) identify operation anomaly for High intensity runs
  - 2) Clusterisation of the anomalies for potential accelerator system involvement
    - For this, identify possible ML model on time series through:
      - First exploration of models (low level but enough to select one candidate that will be explored more thoroughly) for monovariate data:
        - » Performance analysis
        - » Including test on pre-existing data library
      - Application on multivariate data if adequate
  - 3) root cause analysis
  - 4) Prevention (operation/maintenance) + Prediction
- Contribution to the French network on ML and European network
- Contribution to the construction of the accelerator data library:
  - Dedicated operation time will be allocated if necessary (3 days/3years)
  - Arronax can provide data and will work on adapting them to the required standard (within its data limitation eg frequency/synchronization,…)
    - storage
  - These operation time & data are opened to the Accelerator community through collaboration
    - Open also for benchmarking on data and anomaly detection model
- Contribution to the usage of GENCI calculation computing

GENCI: Grand Equipement National de Calcul Intensif

- Extra

- pyDO: Python library for detecting anomalies in multivariate data.
  – 50 detection algorithms
  – Not all applicable to time series
  – https://pyod.readthedocs.io/en/latest/#

Conclusion

- **Preliminary results = Beginning of the exploration of the algorithms**
  - Experimenting on the whole process from data to root cause
- **Result:**
  - **Several algorithms are being employed for the analysis** → towards a selection of best suited algorithms?
    - Shows some interesting results and can help the identification/definition of the outliers
- **Precautions:**
  - The events associated to anomalies have to be connected to real parameters (ex machines): This is not direct or takes time → and here requires in addition **robustness method** (indices, SHAP, evaluation metrics) to eliminate the aboundance of outcomes or unstabilities (sampling strategies).
  - Probably need to dig further in the algorithms (+ more data, multivariates, limits): We are at the beginning (!!!)
- **Future:**
  - Time series:
    - On short term: reconsider/rework some of the present algorithms
    - Temporal Flux to be analysed with more understood algorithms and new algo (?)
    - Define what could be dedicated operational runs and specific to anomaly detection
  - Global studies (statistics):
    - Impact of the results of the rejection of anomalies on the final yields of the analysis (mCi/uAh)
    - Taking into account more information and « impression » from the members of the team

Main metrics

- **Accuracy**: number of correct predictions divided by the total number of predictions, multiplied by 100

$$ACC = \frac{tp + tn}{tp + fp + tn + fn}$$

- A **Recall** is essentially the ratio of true positives to all the positives in ground truth

- **F1 score:** combines precision and recall into one metric by calculating the harmonic mean between those two.
-

$$F_{beta} = (1+\beta^2)\frac{precision * recall}{\beta^2 * precision + recall}$$

- **ROC AUC:** area under the curve:
  - chart that visualizes the trade-off between the true positive rate and the false positive rate. Basically, for every threshold, we calculate TPR and FPR and plot them on one chart

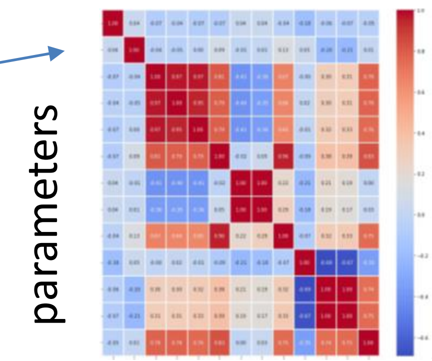
- **PR AUC:** Precision-recall curve – area under the curve: combines precision and recall

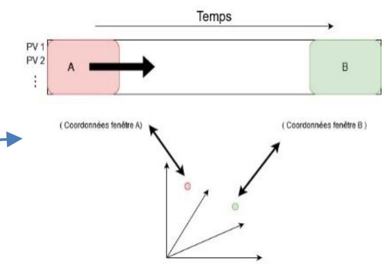
First approach: single run

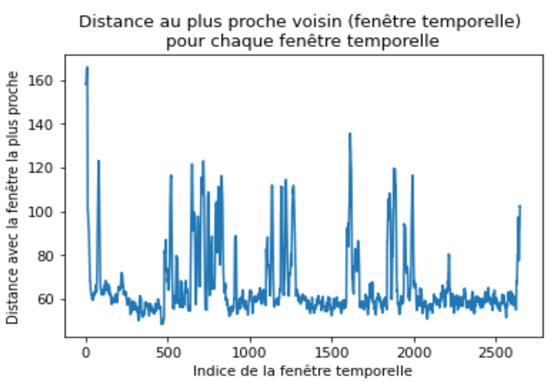
Steps that were used for our study:

- Data mapping with correlation that can be eliminated for later studies
- Verification using Principal Component Analysis (PCA=dimension reduction)
- Take into account evolution of data over time through time windows:
  - Ex: 1h rolling windows
- Analysis of the window through distances to its “sisters” taking into account all data:
- Not all windows look alike, and a rather large amount of data are still bursting out

Machine & diag parameters

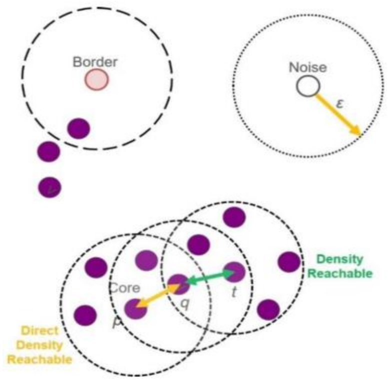






Clusterisation: DBSCAN

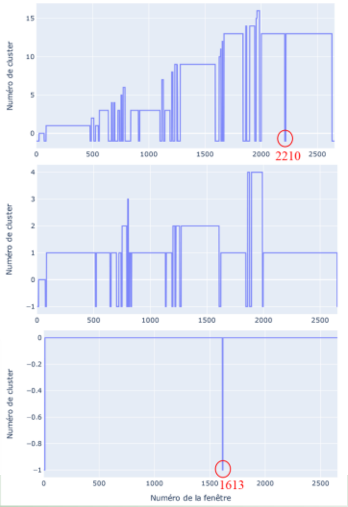
- We wanted to gather the data to find out if some windows would come out
- For this: clusterisation through DBSCAN:
  - DBSCAN (Density-Based Spatial Clustering of Applications with Noise):
    - Clustering algorithm based on the density of the data.
    - No prerequisite on cluster.
    - Cluster using hyper-parameters (minimum nb of points in the cluster and euclidian distance)



$MinPts = 3$
$\varepsilon = 75$

$MinPts = 3$
$\varepsilon = 100$

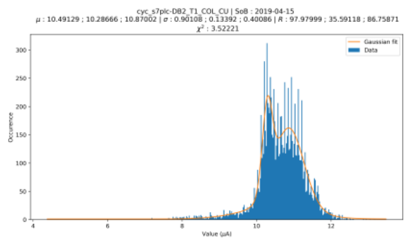
$MinPts = 3$
$\varepsilon = 125$



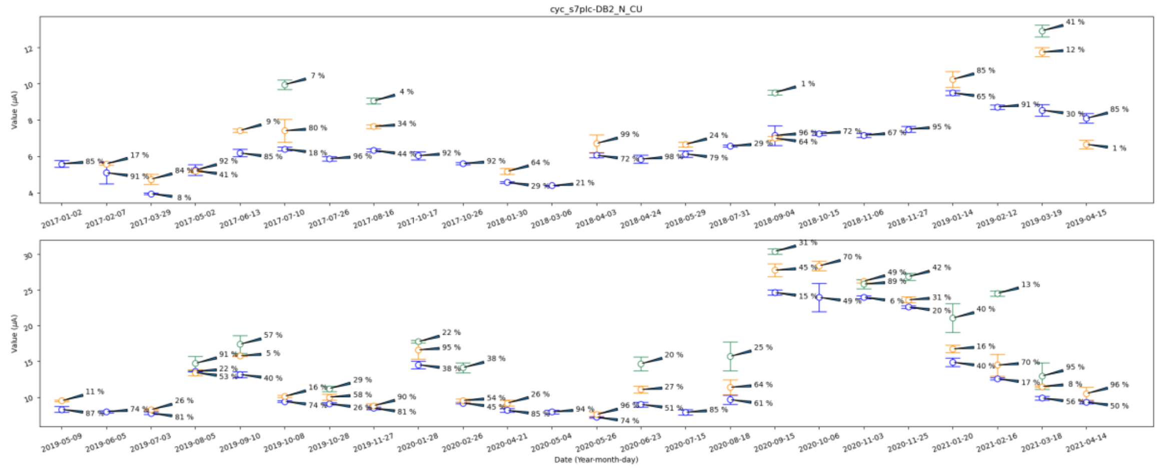
One events over 3000 sliding windows comes out → what is it? Anomaly?

Second extreme approach: multiple runs

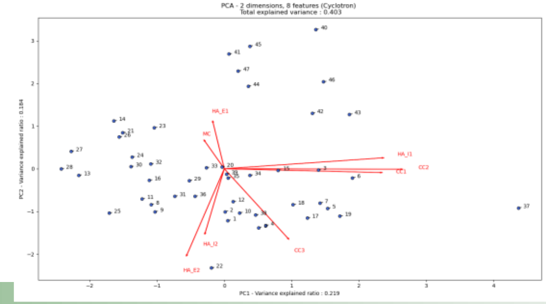
Classic statistical study of the machine parameters: several distribution are necessary to ajust to a gaussian model



Evolution of the results and coverage power of each gaussian distribution for each run

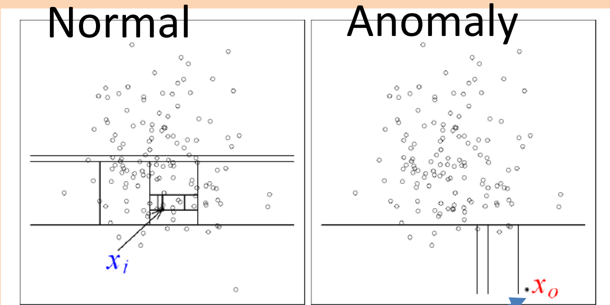


And then analysis in principal component (PCA) as earlier



Anomalies detection through Isolation Forest (IF)

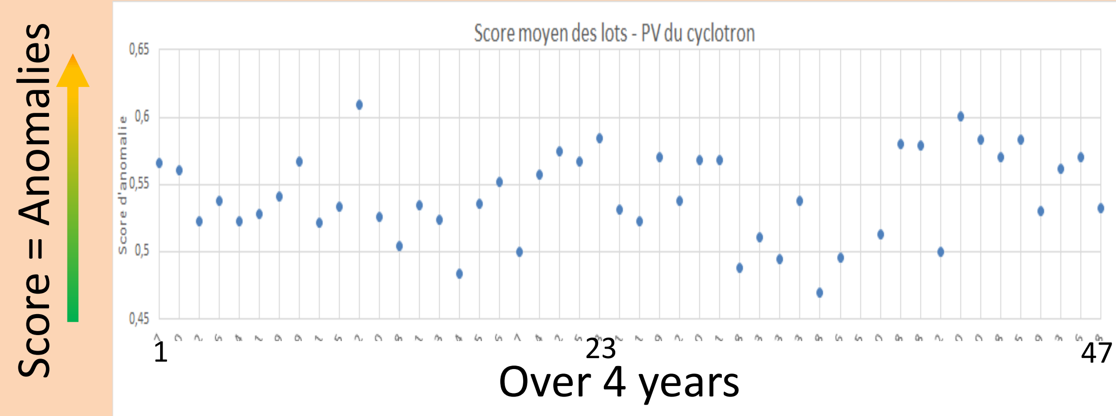
Isolation of data through random decision in branch = isolation forest



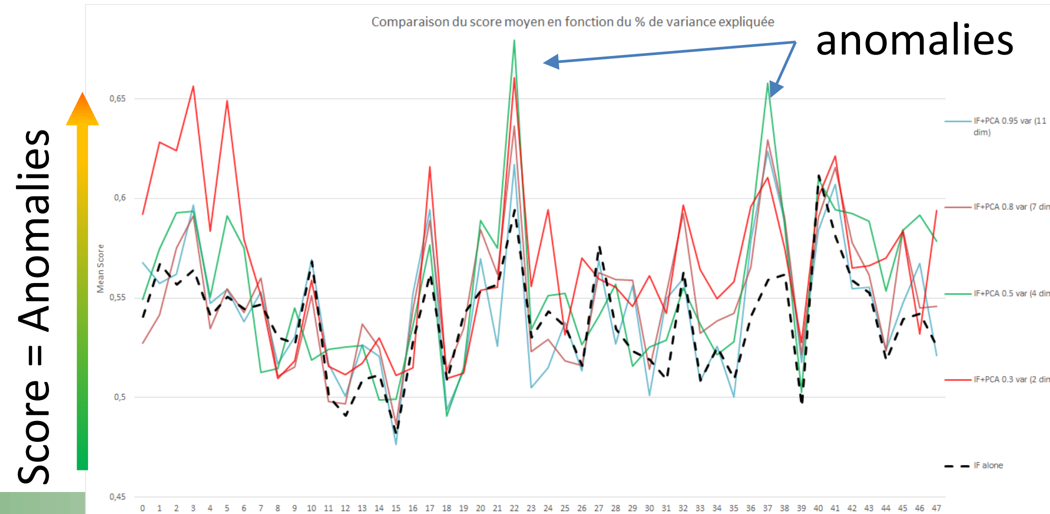
Isolation of the data

A Score calculated from the cuts give normality vs anomaly

Training with 30% of the runs



Combinaison of PCA and isolation forest allows to determine wich runs comes out: non-orthogonal in the isolation



ARRONAX

afaq ISO 9001 Qualité AFNOR CERTIFICATION

Questions on IF

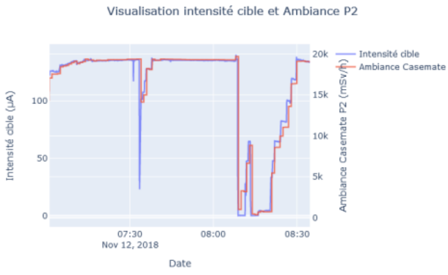
- Isolation forest is a random method so new questions:
  - Is it a **robust method**? Are the runs, tagged as anomaly, all the time present? whatever is the initial set for the learning model ?→ SHAP method
  - **Why a run comes out as an anomaly**? → get back to the selecting tree in the isolation forest in order to identify the leading parameters

On the study - Robustness

1) First approach with the time windows:

- Exploration of the robustness through selection indices. This helped to study the variation of the identified cluster as a fonction of the input hyper-parameters (distance, cluster):
  - Davies-Bouldin indices
  - Calinkski-Harabasz indices → the most decisive as it gives a stability and level of results less independent of the hyper-parameters

DBSCAN event 1613



2) Second approach on the multiple runs:

- SHapley Additive exPlanations (SHAP): SHAP Iteration on the runs.
- Each decision tree in the forest is studied. The stability of the anomalies found is here taken as a stable indication of an unexpected event
- Through this method, several runs came out as anomalies. Some were associated with high neutral current in the machine, some are still under investigation

i.e so far we discovered already known anomalies – but it is not the end of the story

We used these knowledge to push forward concerning our study of algorithms and evaluation of the results

F.Poirier et al, “First anomalies exploration from data mining and machine learning at the ARRONAX cyclotron C70XP”, 10.18429/JACoW-IPAC2023-TUPM036