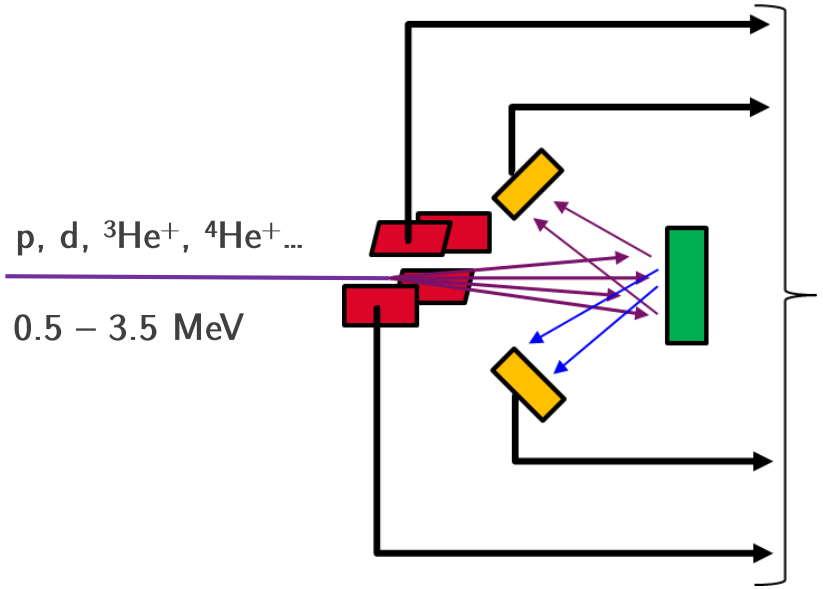
ARTIFICIAL INTELLIGENCE FOR ION BEAM ANALYSIS SPECTROSCOPIES

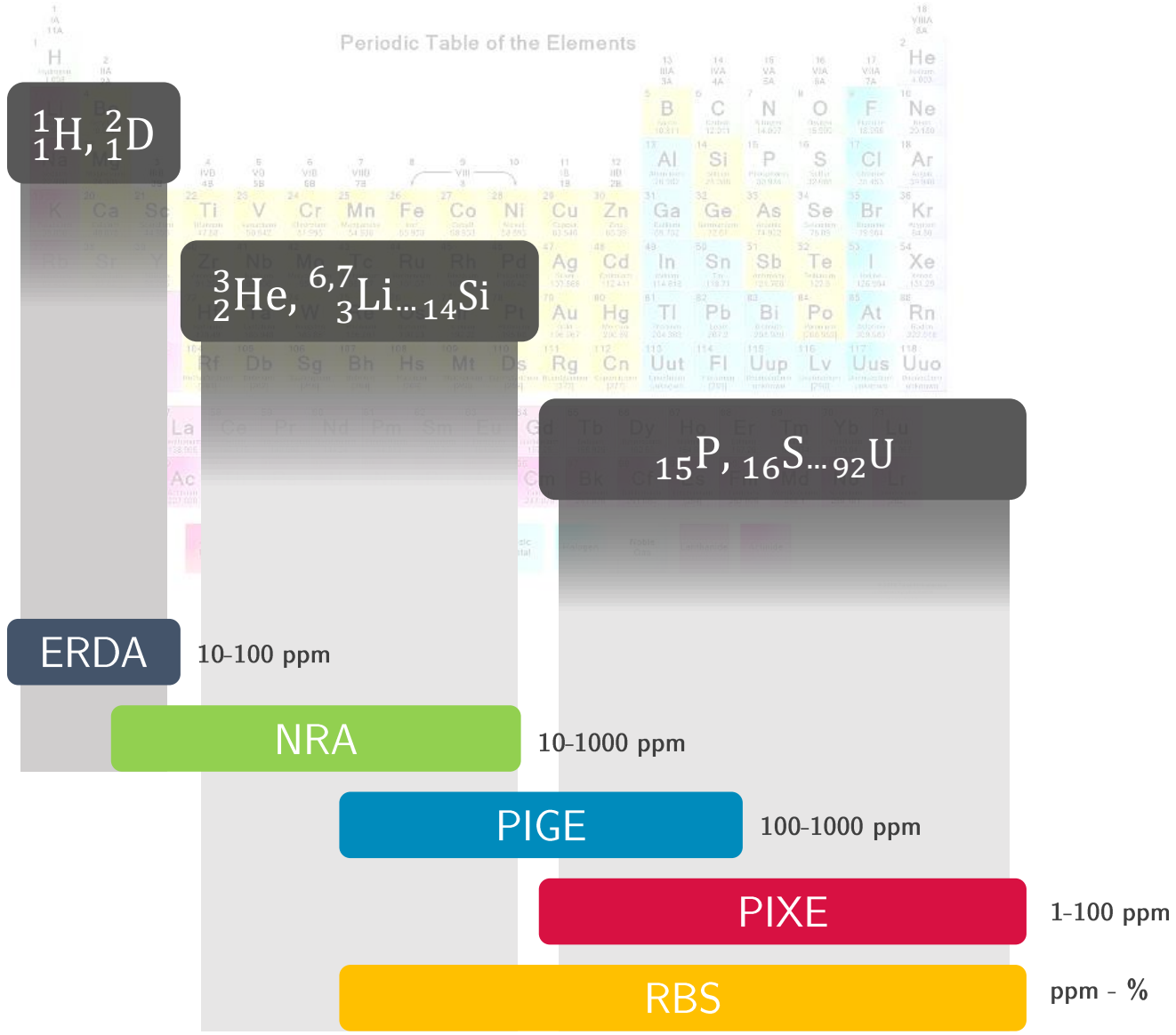
Hicham Khodja

May 31st 2022

DE LA RECHERCHE À L'INDUSTRIE

Motivations: IBA at a glance

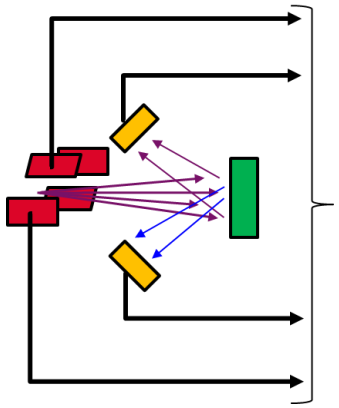


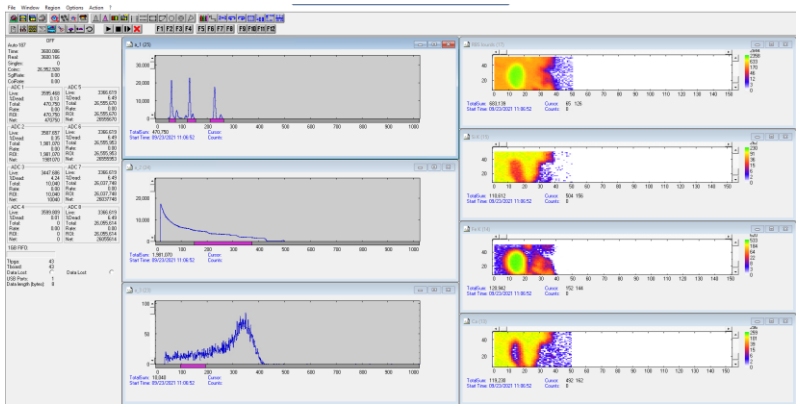


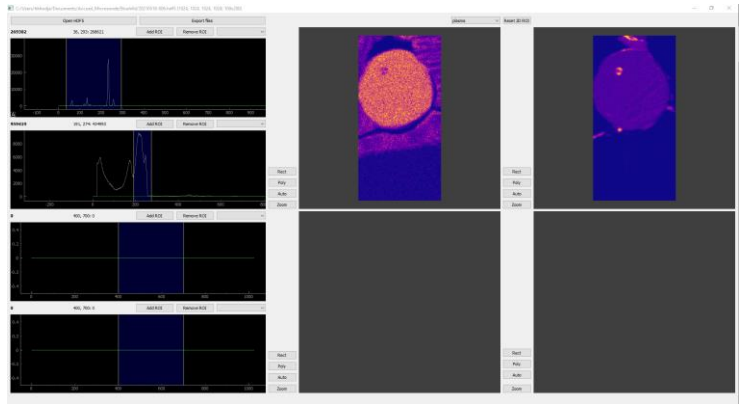
ERDA	Elastic Recoil Detection Analysis
NRA	Nuclear Reaction Analysis
PIGE	Particle Induced Gamma-ray Emission
PIXE	Particle Induced X-ray Emission
RBS	Rutherford BackScattering

Motivations: from data to compositions

- Data workfow

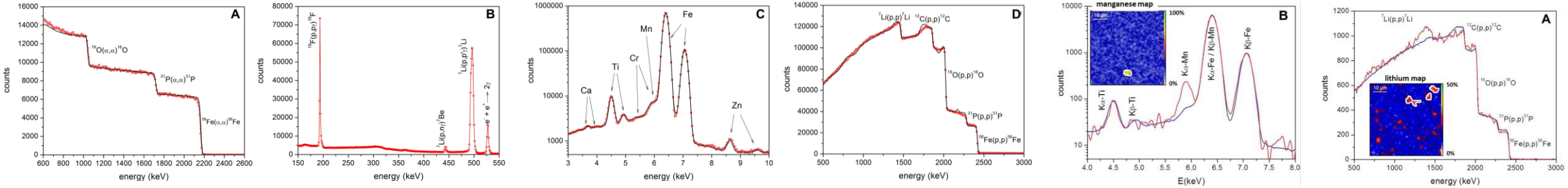






- Data processing

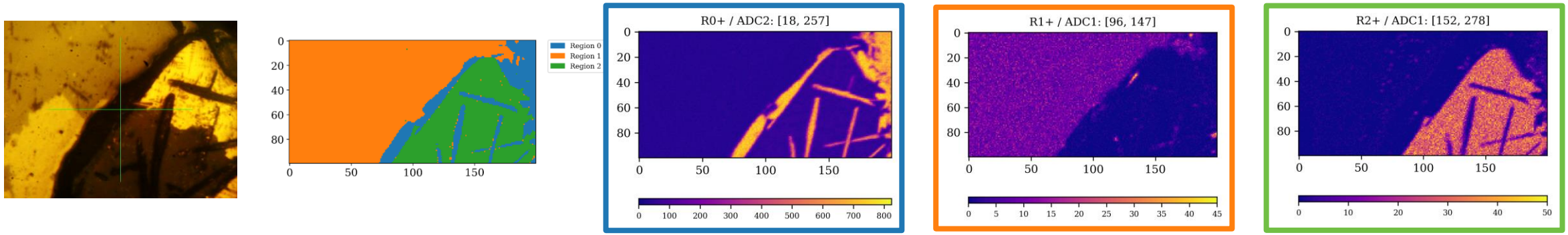
https://doi.org/10.1016/j.nimb.2012.08.014

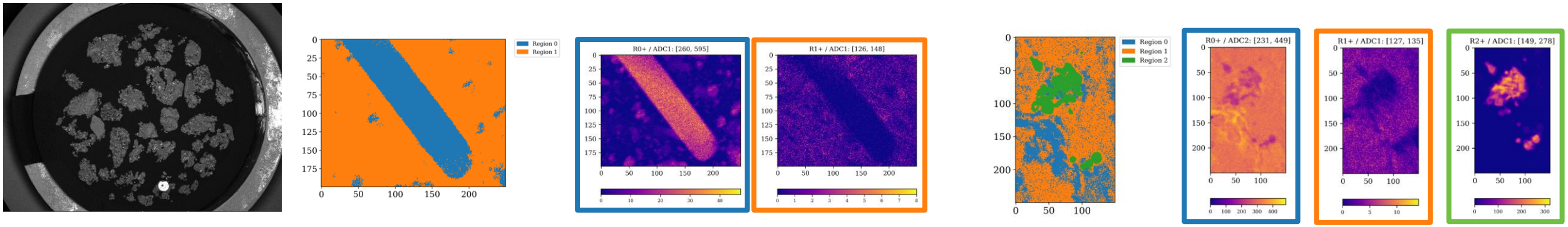


- Multiple steps
- Multiple software programs
- **Reaching full IBA convergence requires brave engagement…**

Automated hyperspectral image analysis

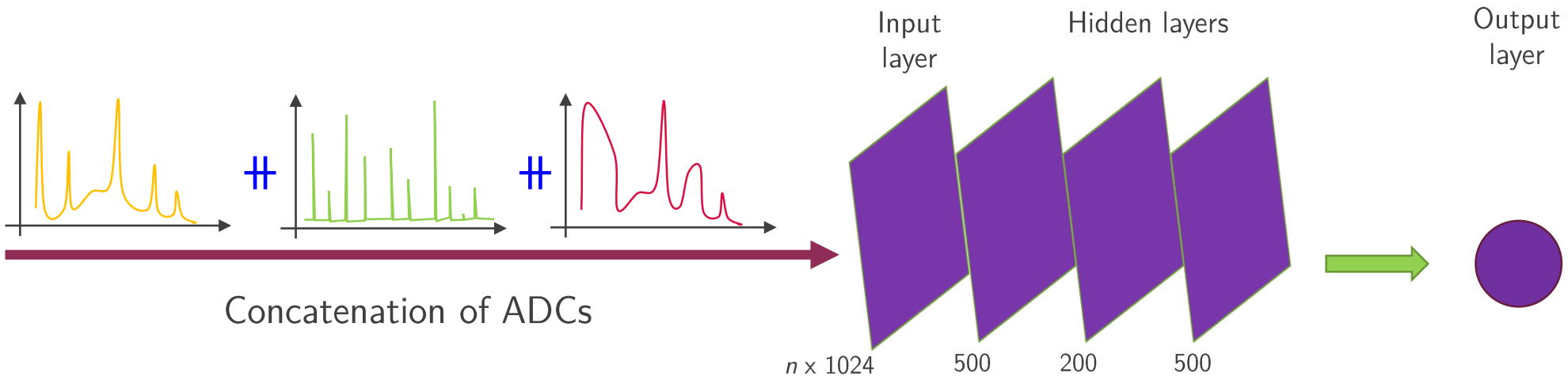
- PCA + k-means clustering; chemical imaging by highest contrast ROI automated selection



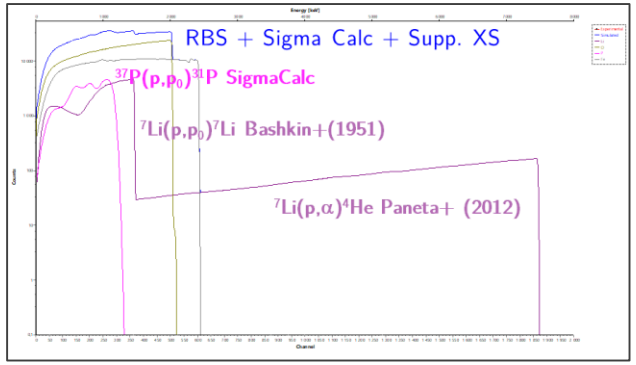


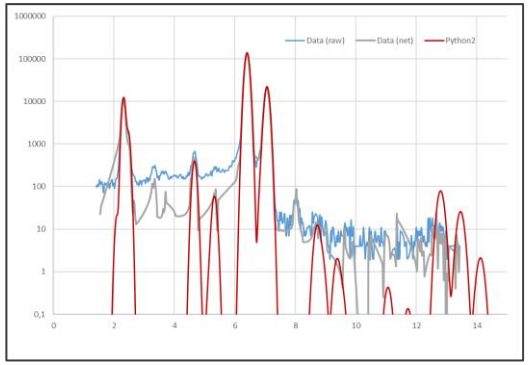
IA for composition prediction

- MLP regressor



- MLP training performed by IBA simulation software programs hijacking

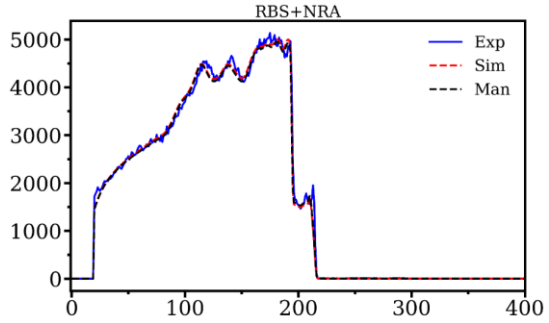


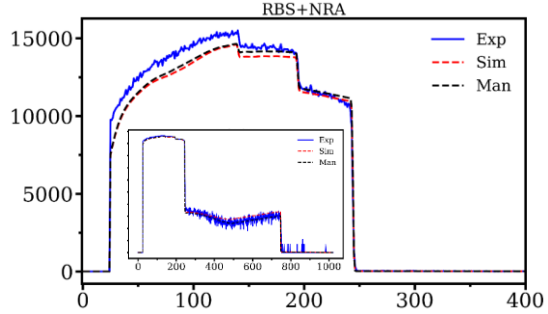


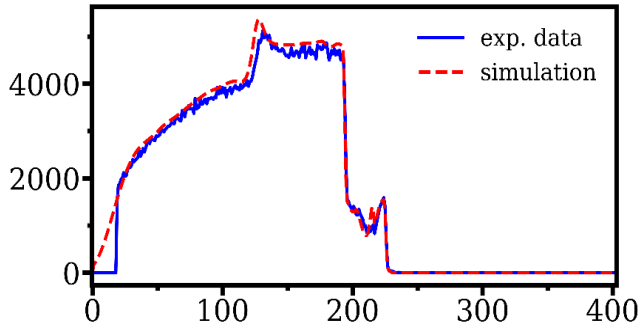
Preliminary results

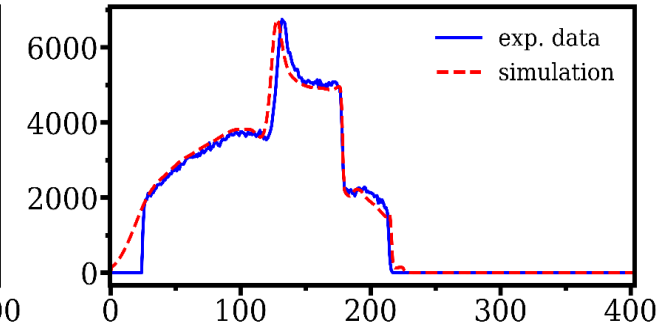
Training: Mg, Al, O | Li, Ta, O | Ca, Si, O, C, Ti

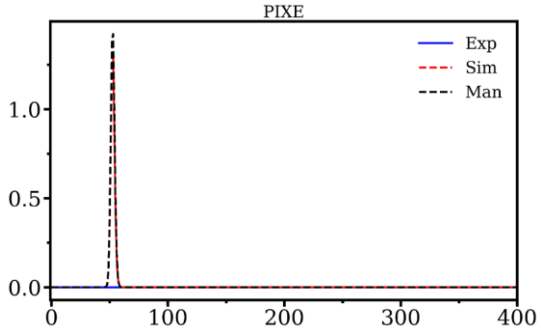
$MgAl_2O_4$ / $Mg_{1.02}Al_{1.83}O_{4.05}$ $Li_2Ta_2O_6$ / $Li_{2.74}Ta_{1.85}O_{5.4}$ $CaSiO_3$ / $Ca_{0.8}Si_{1.1}O_3C_{0.03}Ti_{0.05}$ SiC / $SiC_{0.86}Ca_{0.01}O_{0.03}Ti_{0.04}$

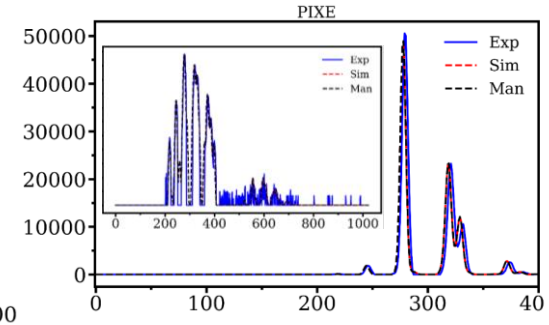


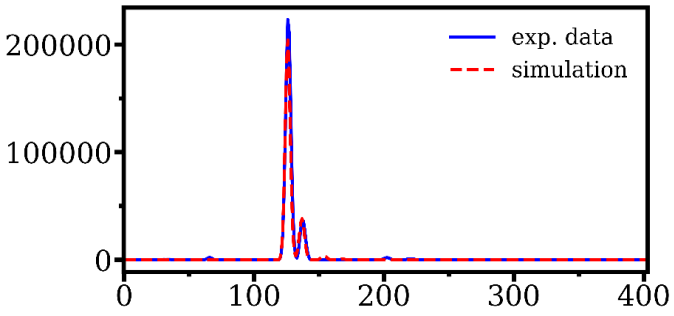


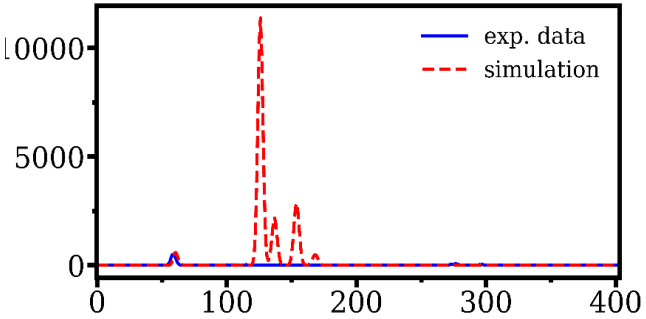










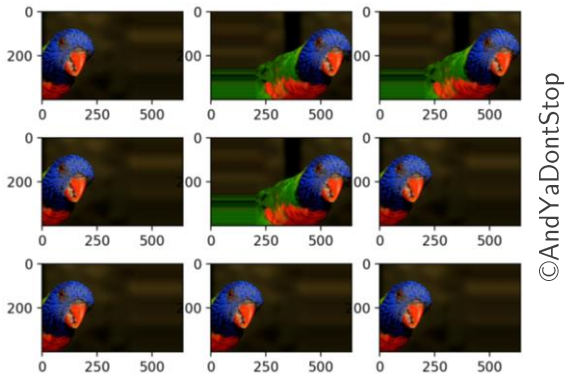


- Well-choosen labels, ~ 10 kS for training: good performances
- If extra labels, loss of precision

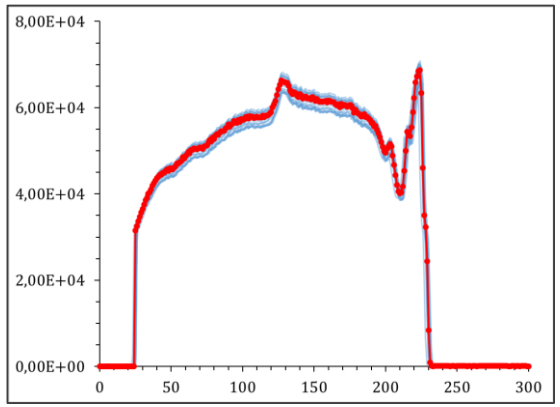
→ Specific training set for each group of labels (Simul. time ~ 1-10h, Train. time ~ 1-10 min)

Optimizations

- **Artificial data augmentation**

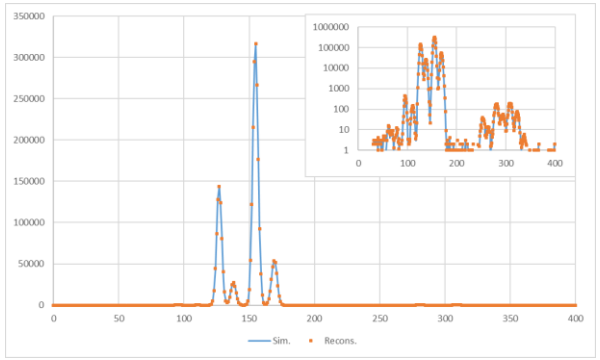


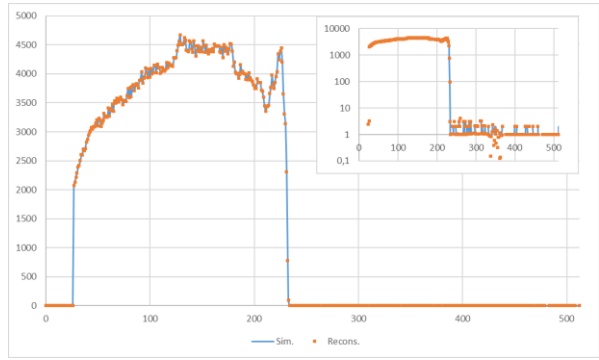
  - Introducing small (and realistic) alterations to learning set spectra

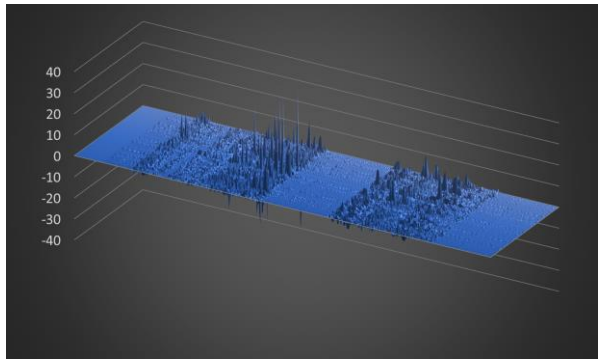


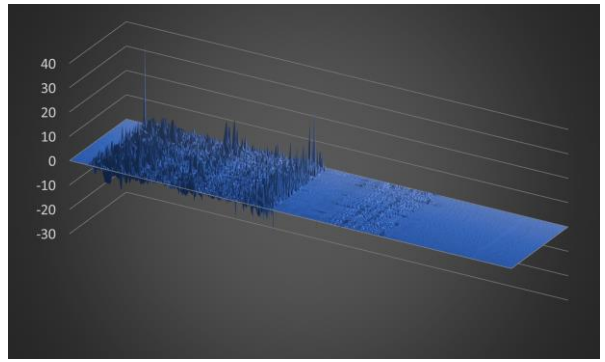
  - Minimal computing cost
  - 5-50 x faster

- **Dimensionality reduction**
  - Reduce the size of the input layer and thus the numbers of NN weights
  - Loop on $n\_components$ of PCA over the training set to ensure correct spectra reconstructions (1-$<R^2>$=1e-6)
  - Typical size reduction: 4096 → 150





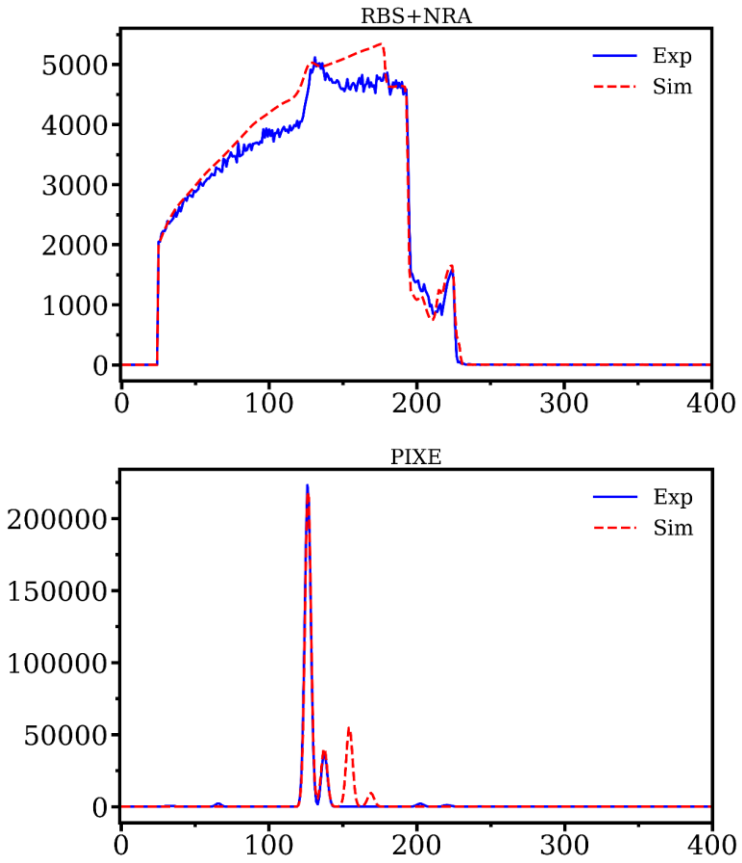




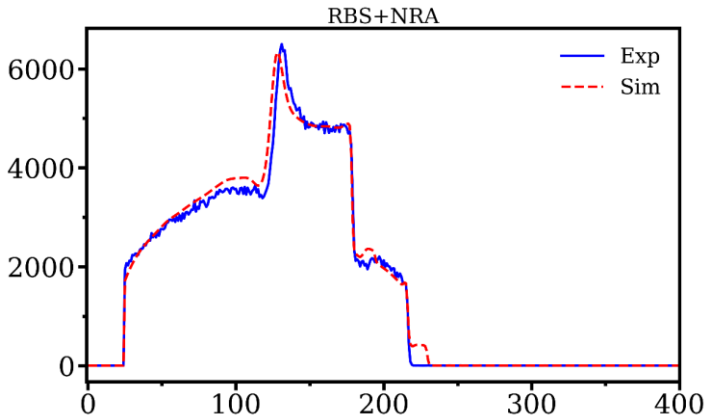
Optimizations

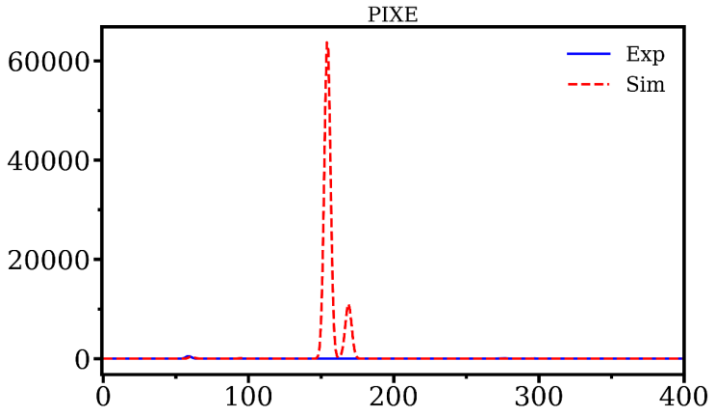
Data augmentation: x10 / PCA $n\_comp = 66$ NN: (2048, 500, 200, 500, 5) → (66, 66, 66, 5)

$CaSiO_3$ / $Ca_{0.69}Si_{0.34}O_3$ $C_{0.68}Ti_{0.24}$

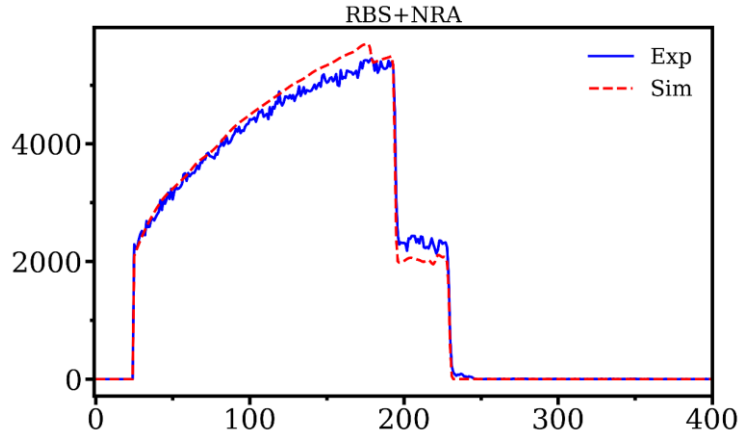


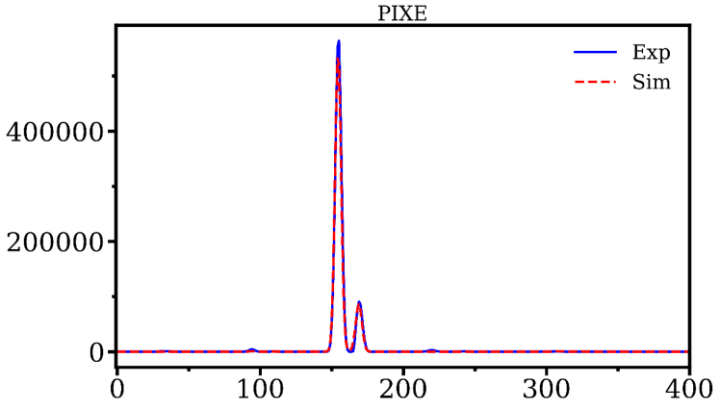
SiC / $SiC_{1.07}$ $Ca_0O_{0.09}Ti_{0.11}$





TiO_2 / $Ti_{0.77}O_2$ $Si_0C_{0.21}Ca_0$

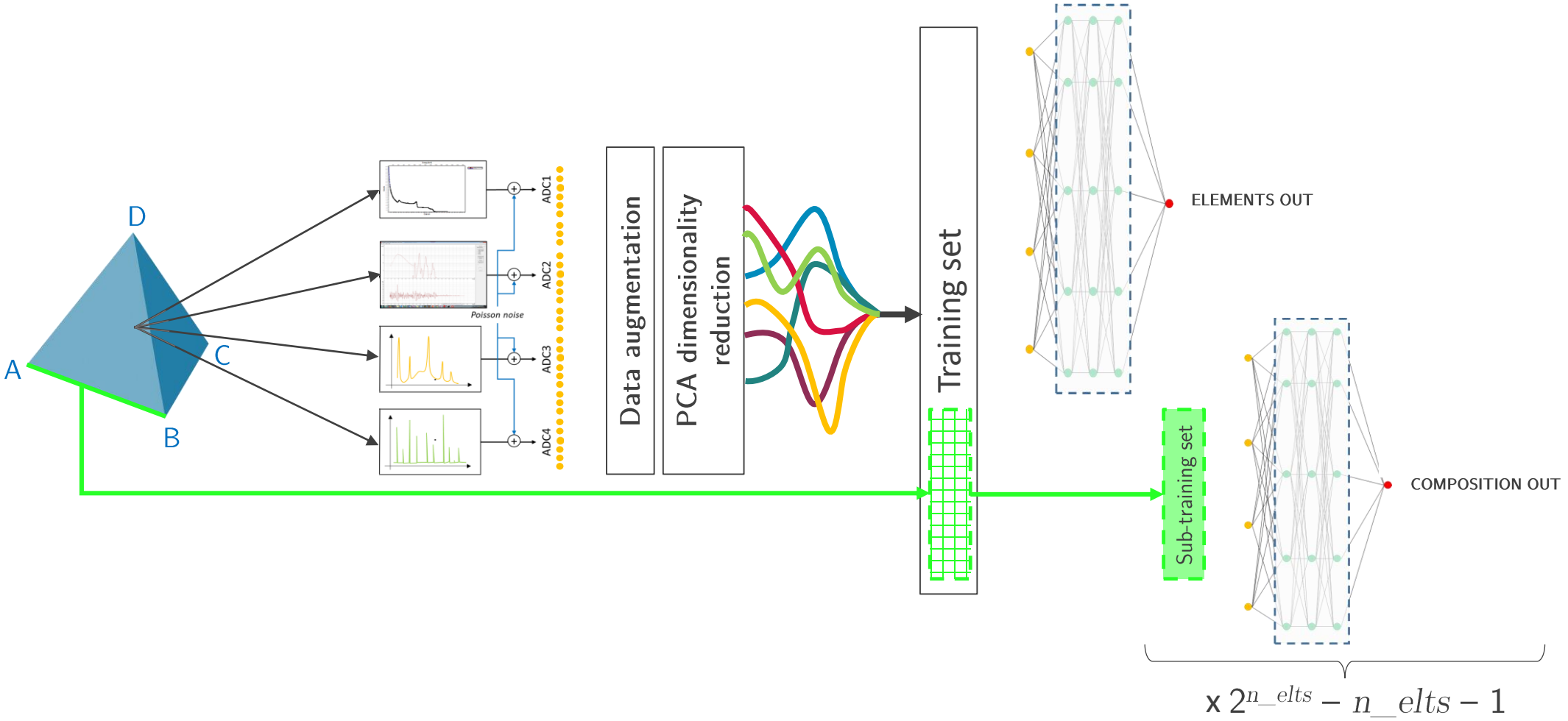




- Much faster results
- Predictions remain acceptable, but MLP still finds extra label if present in the training set

Associating classification and regression

- Built a training set with all possible labels
- Use a classifier to identify the effective labels
- Train for each multi-class a MLP regressor with selected data



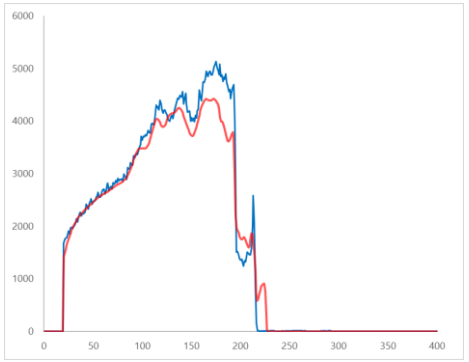
Associating classification and regression

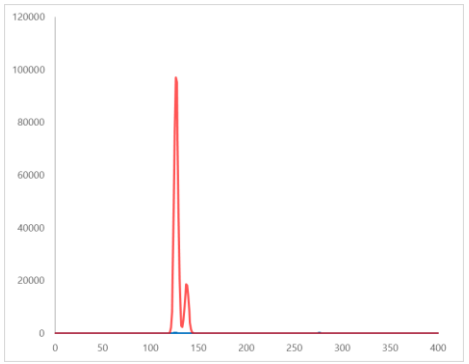
Training: Ca, Si, O, C, Ti, Al

Classifier OFF

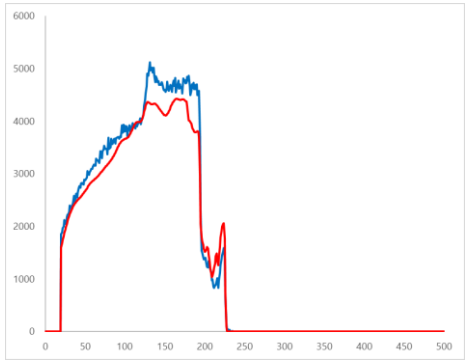
Acquisition time: 15 min

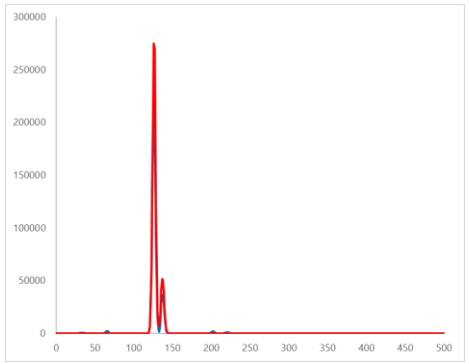
Al_2O_3 / O Si C Ca Ti Al



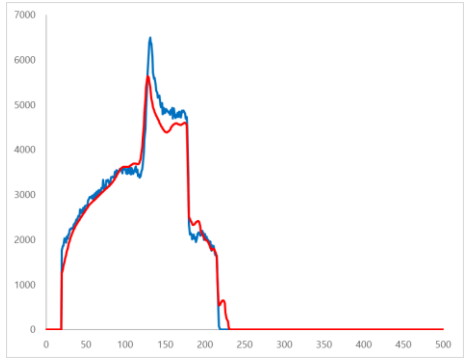


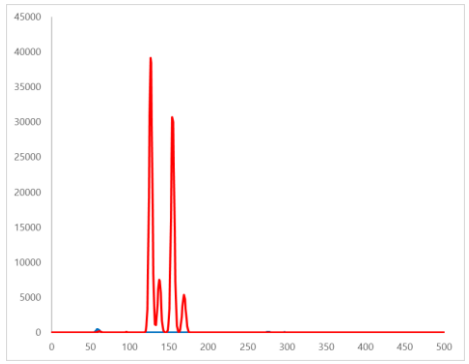
$CaSiO_3$ / O Si C Ca Ti Al



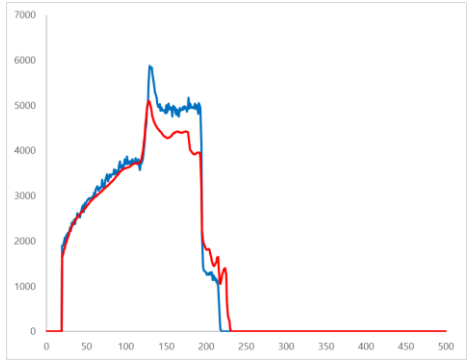


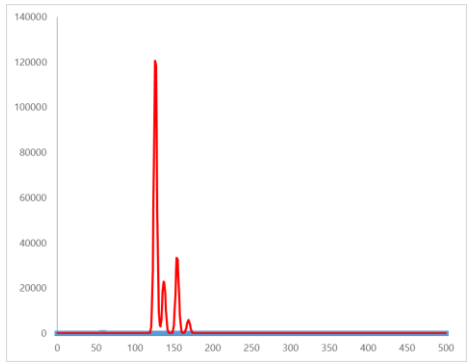
SiC / O Si C Ca Ti Al



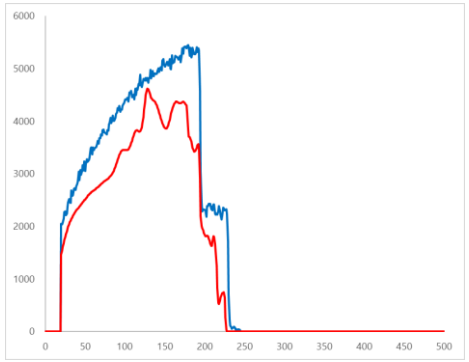


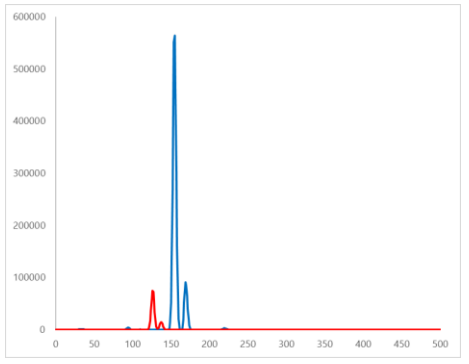
SiO_2 / O Si C Ca Ti Al





TiO_2 / O Si C Ca Ti Al





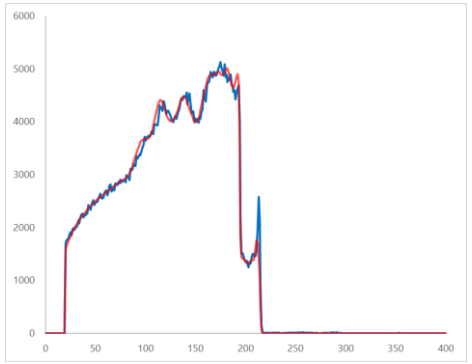
Associating classification and regression

Training: Ca, Si, O, C, Ti, Al

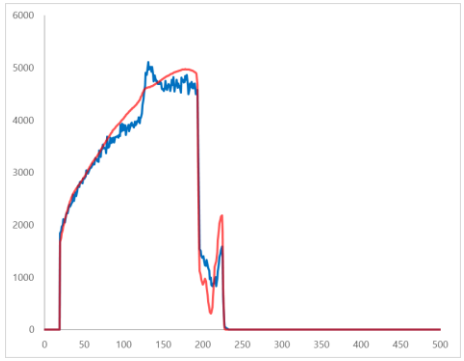
Classifier ON

Acquisition time: 15 min

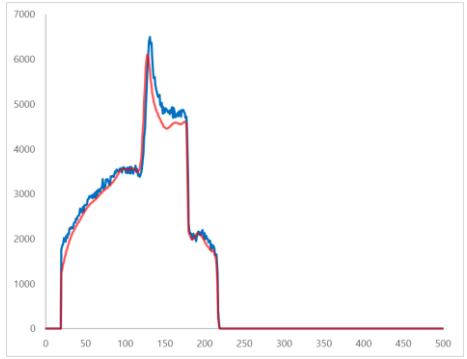
Al_2O_3 / O Si C Ca Ti **Al**



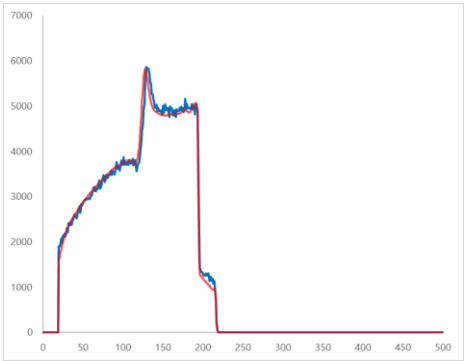
$CaSiO_3$ / **O** **Si** C **Ca** Ti Al



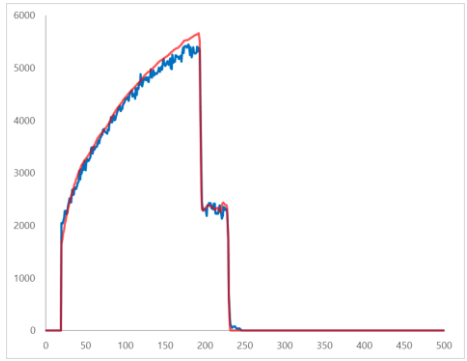
SiC / O **Si** **C** Ca Ti **Al**

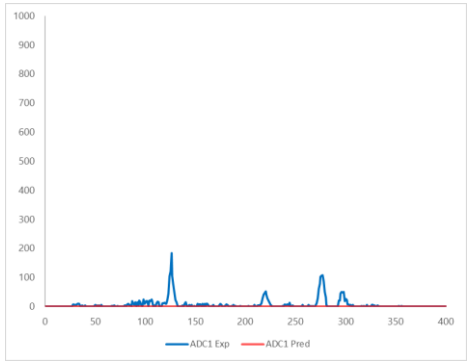


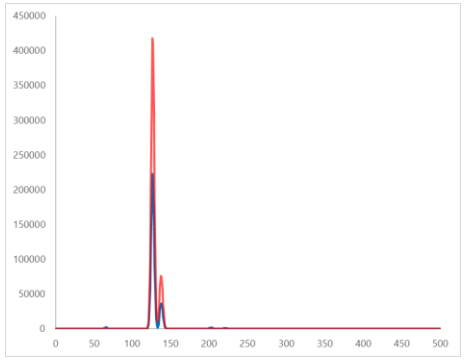
SiO_2 / **O** **Si** C Ca Ti Al

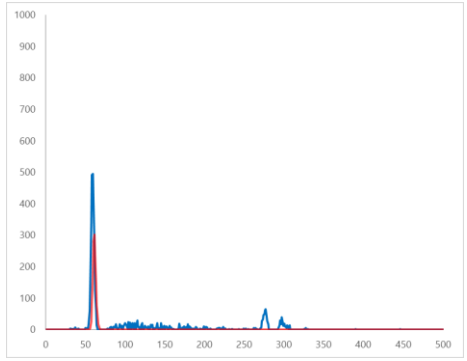


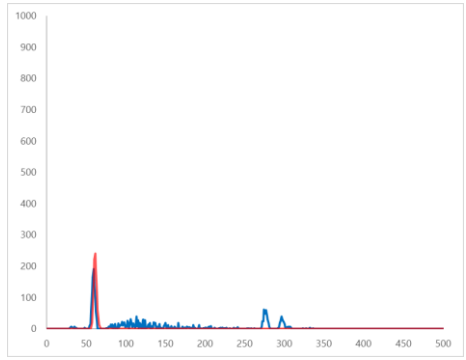
TiO_2 / **O** Si C Ca **Ti** Al

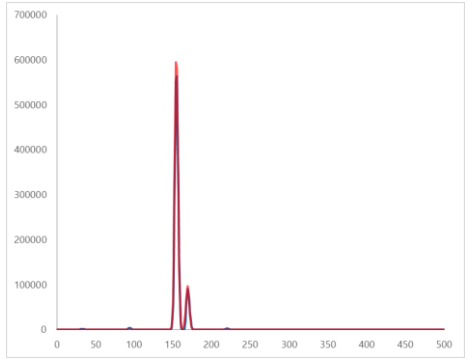












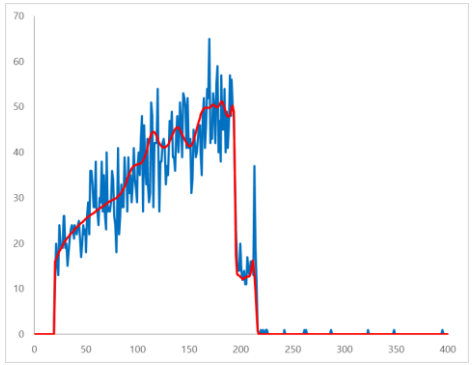
Associating classification and regression

Training: Ca, Si, O, C, Ti, Al

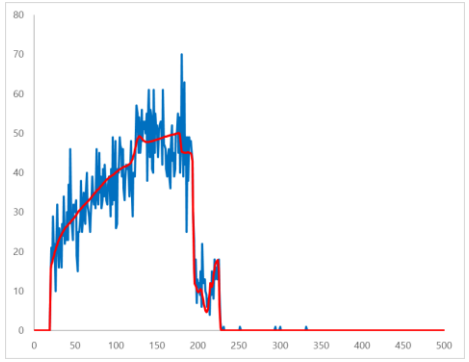
Classifier ON

Acquisition time: 9 s

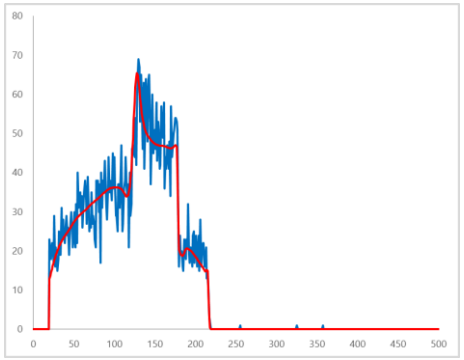
Al_2O_3 / **O** Si C Ca Ti **Al**



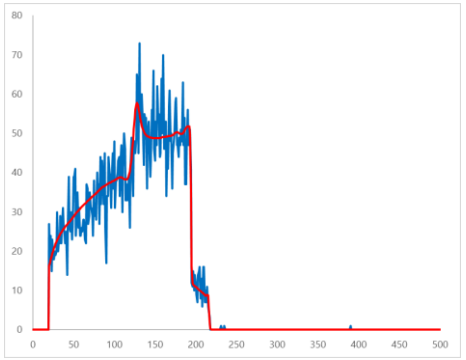
$CaSiO_3$ / **O** **Si** **C** **Ca** Ti Al



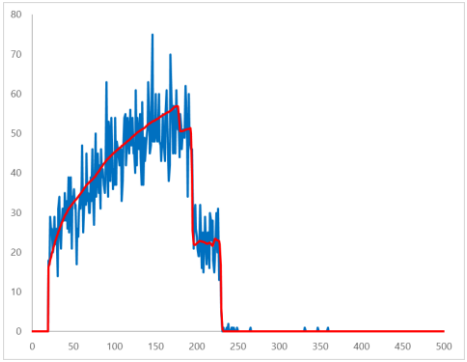
SiC / O **Si** **C** Ca Ti Al

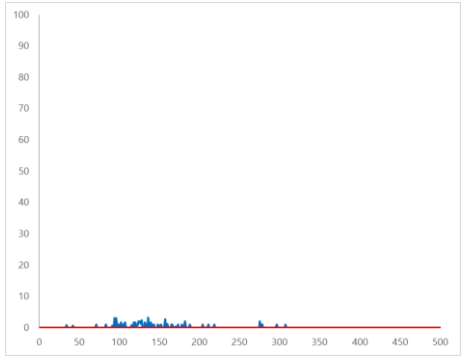


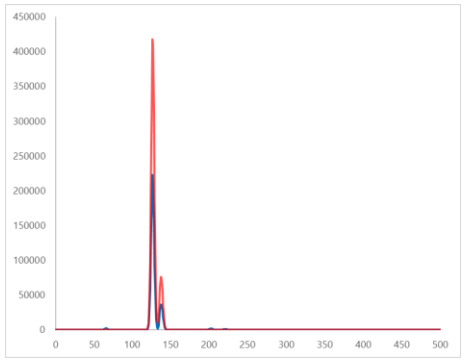
SiO_2 / **O** **Si** C Ca Ti Al

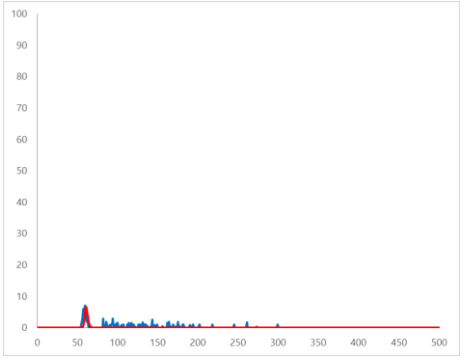


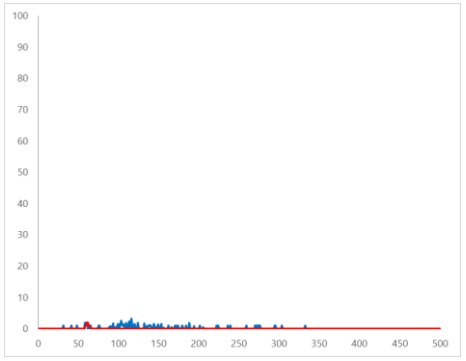
TiO_2 / **O** Si **C** Ca **Ti** Al

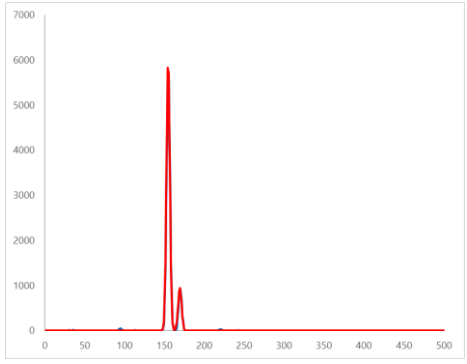












Conclusions

- Hyperspectral image analysis helps at finding chemical phases of heterogonous samples and extract for each phase IBA spectra collected from different detectors
- From coupled IBA techniques, composition predictions can be obtained using a trained MLP regressor
- Artificial data augmentation and PCA dimensionality reduction considerably improve execution times
- When extra labels (elements) are used for training, they are almost always found by the regressor
- Instantaneous elemental discrimination and identification is achieved using a well-trained MLP classifier
- Even with noisy spectra, a MLP classifier associated with MLP regressor produce valuable predictions, paving the way to live data processing
- There is still room to improve regressor performances: weighting of input data, topology...
- Extend the classifier + regressor method to multilayer samples
- Trace element detection and quantification (linear contribution) still to be implemented