

Reconnaissance automatique de domaines d'altération et de résistivité par des méthodes statistiques

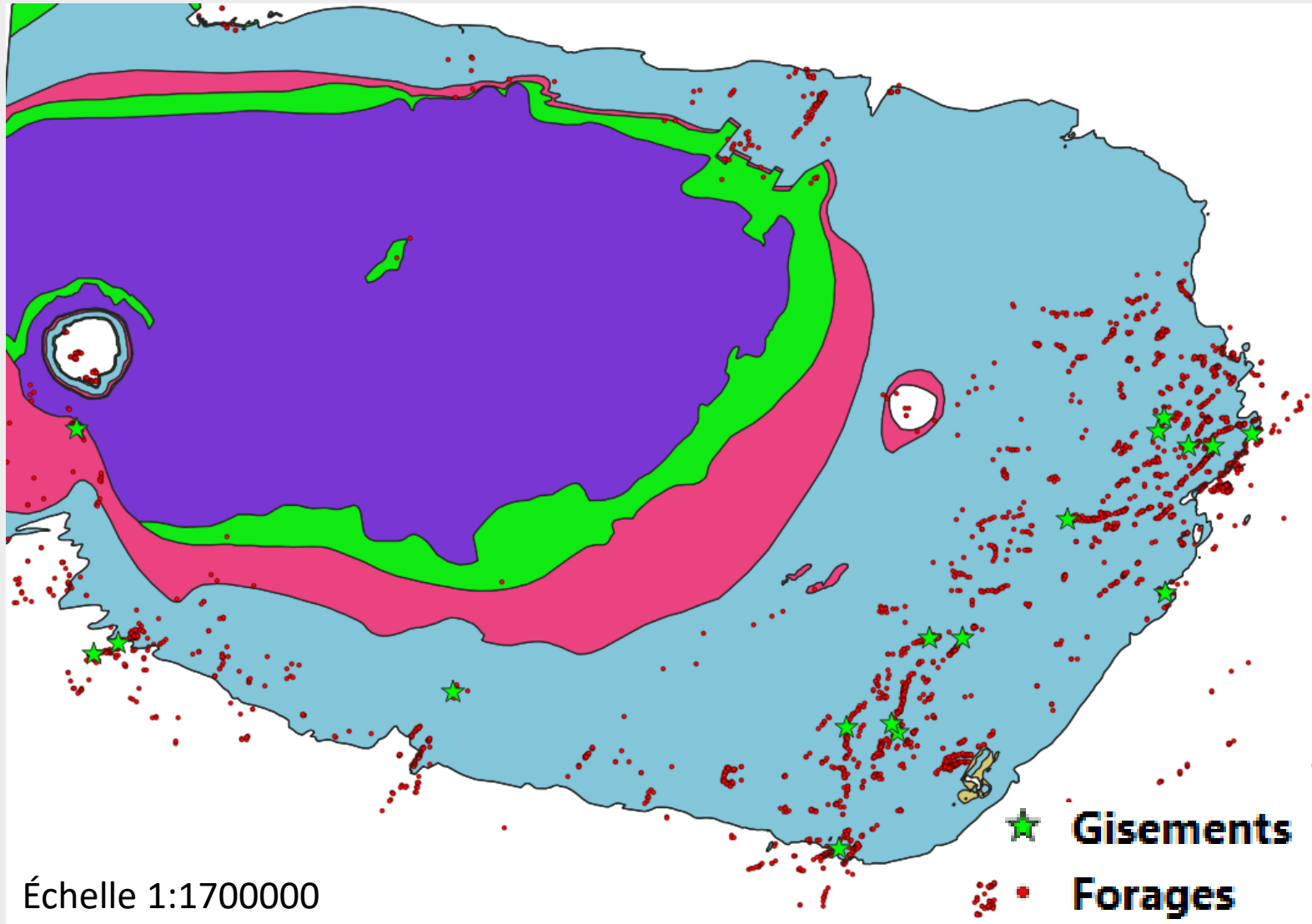
20/01/2022

Mehdi Serdoun - Géoressources

- I) Problématique générale de l'exploration des gisements d'uranium de discordance
 - II) Données disponibles
 - III) Pré-traitement des données / variables
 - IV) Modèles de prédictions de variables
 - V) Conclusions

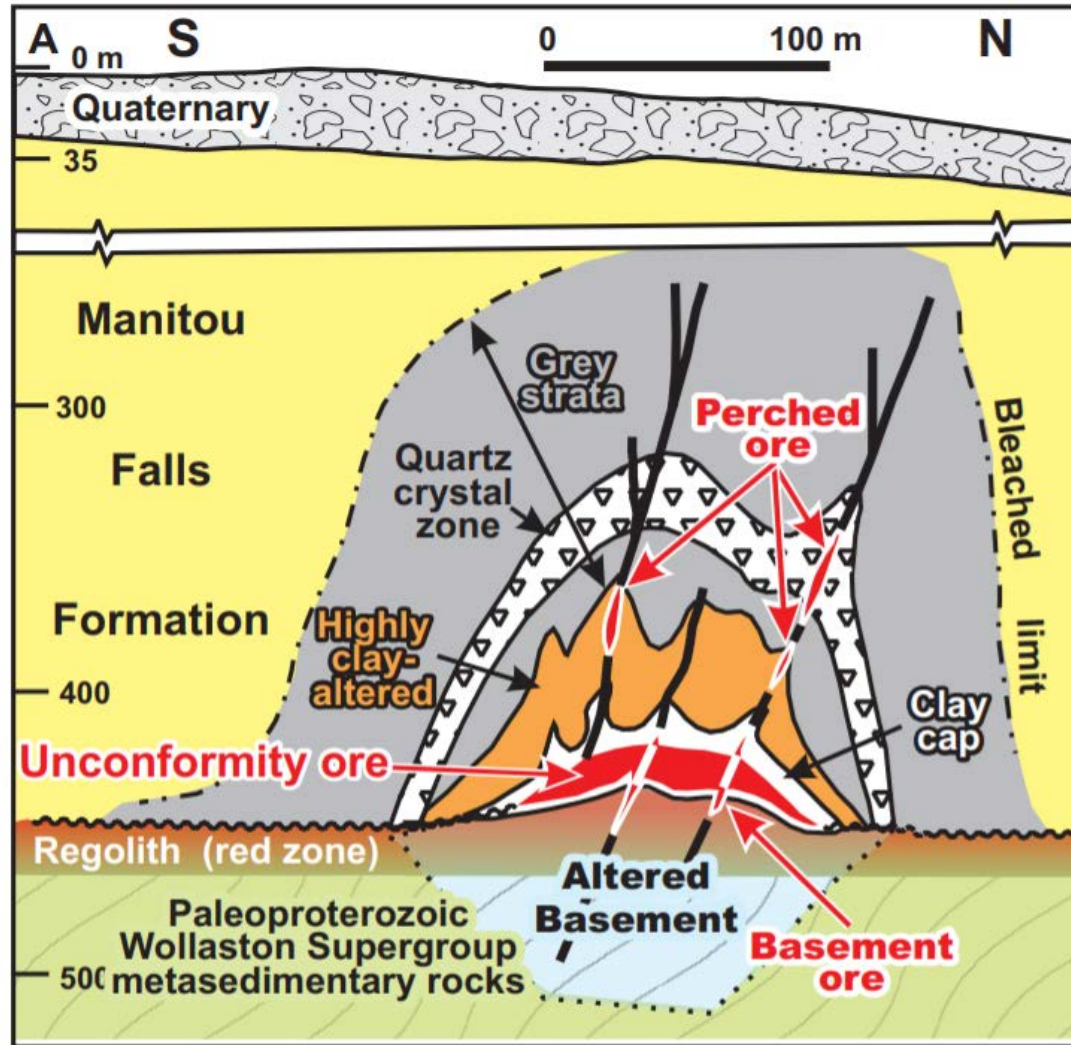


Problématique générale de l'exploration des gisements d'uranium de discordance de l'exploration



- Nombreuses données acquises dans le cadre de l'exploration des gisements de discordance depuis les années 1960
- Routines d'exploration ont peu changé depuis 2000-2005 = vastes jeux de données disponibles
- Cas d'étude = Cigar Lake

Problématique générale de l'exploration des gisements d'uranium de discordance



- Gisements d'uranium limités en taille (quelques dizaines de mètres de hauteur / largeur)
- Cependant, halos d'altération beaucoup plus visibles spatialement, caractérisés par des assemblages d'argiles/chlorites, différents modèles existent
- Halos d'altération comme proxys dans la stratégie d'exploration
- Compréhension de ces halos = IR, géochimie, logs d'altération (hématite, chlorite, graphite,...) dans le socle et le bassin

Types de données

Données issues de
campagnes
d'exploration
annuelles

Systematiques (majorité des
puits)

Non systematiques

*Quelle possibilité d'utiliser
les données instrumentales
et systematiques pour la
valorisation des autres jeux
de données ?*

Instrumentales

IR, géochimie, sonde gamma,...

Non instrumentales

(logs, échelles d'altération, tailles
de grain)

Instrumentales

Résistivité (30% des puits),
pétrophysique
(échantillons)

Non instrumentales

(fracturation,...)

Types de données

Données instrumentales Spectral, Géochimie

*Spectral : 2150 longueurs d'onde (350-2500nm)
Géochimie : digestion totale (majeurs et trace) +
partielle (trace)*

Données de logs

*Un échantillon (géochimie, spectral) = valeurs associées
à la même profondeur, même forage, dans les logs
Hématite, graphite, chlorite, argile, lithologie,
paléorégolithe*

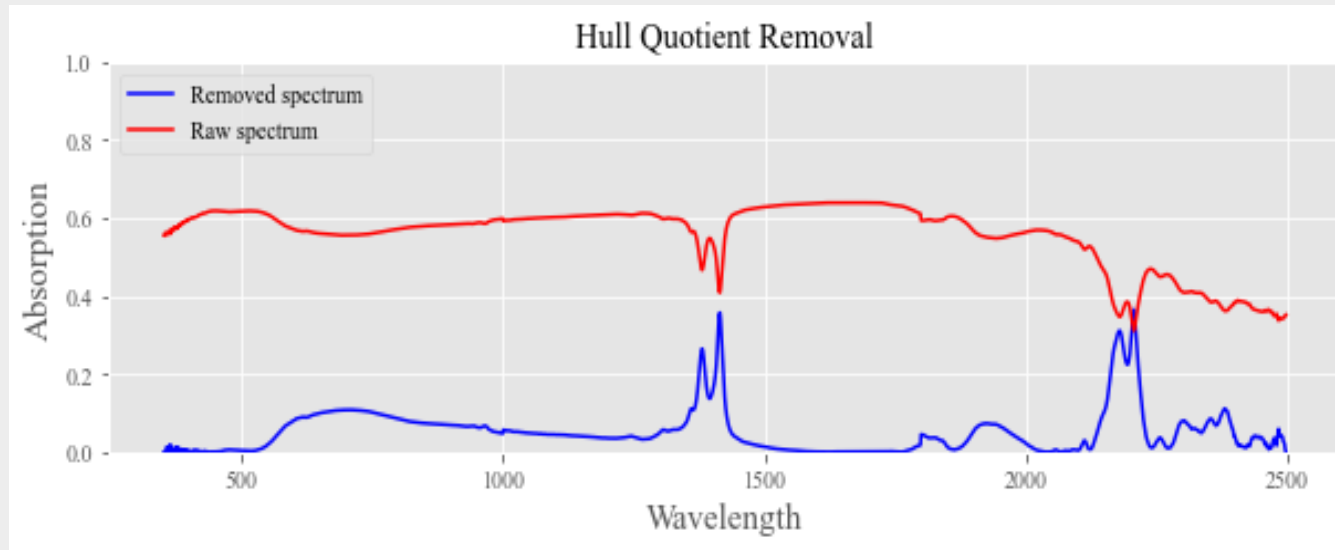
**Observation des corrélations
données instrumentales /
données de logs**

**Sélection des variables + choix
d'un modèle de prédiction**

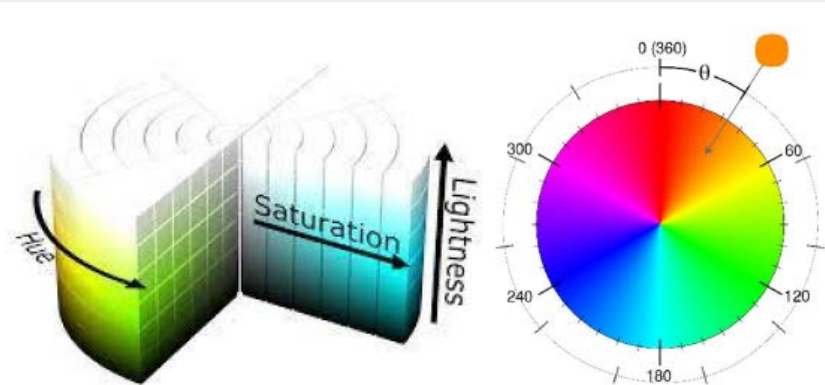
**Réapplication à des échantillons
sans données de logs**
*Apprentissage machine =
apprendre à un modèle à obtenir
des labels*

*Exemple : apprendre à un
modèle à reconnaître les
échantillons graphitisés à
partir d'une combinaison
spectral/géochimie*

Pré-traitement



- Application d'une transformée log-centrée réduite aux données de géochimie
- Données postérieures à 2005 seulement
- 173 forages sur Waterbury-Cigar
- Retrait du continuum
- Extraction des propriétés optiques (HSL + RGB)

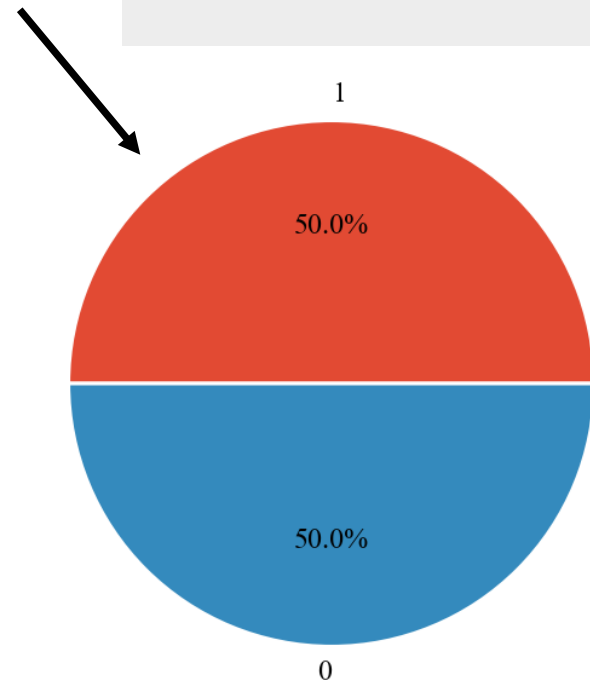
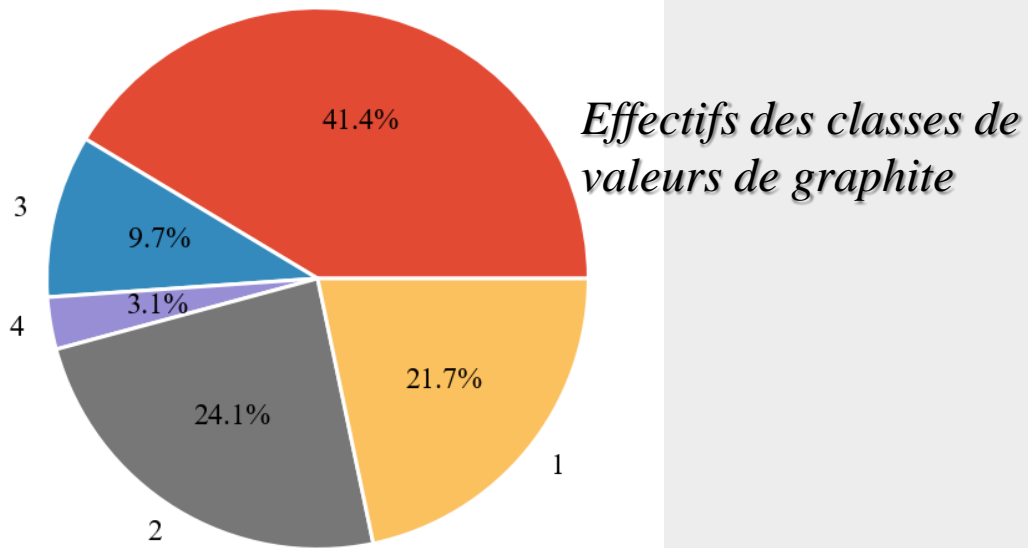


$$\text{Relative Green} = \frac{\text{Green (RGB)}}{\text{Blue (RGB)} + \text{Red (RGB)}}$$

$$\text{Relative Blue} = \frac{\text{Blue (RGB)}}{\text{Red (RGB)} + \text{Green (RGB)}}$$

$$\text{Relative Red} = \frac{\text{Red (RGB)}}{\text{Green (RGB)} + \text{Blue (RGB)}}$$

Définition du jeu d'apprentissage



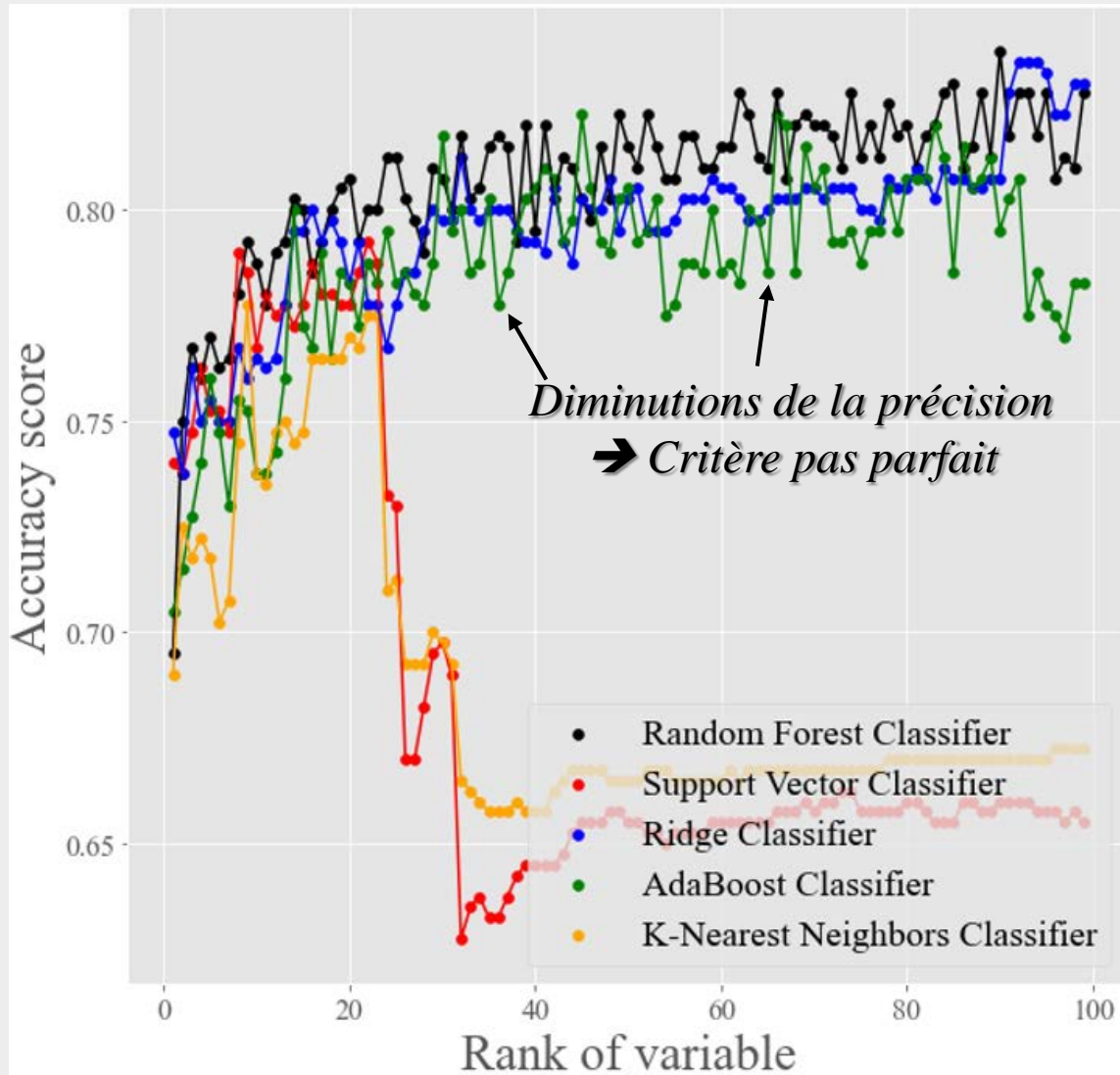
- Graphite = échelle de 0 (absence) à 5 (très forte)
- Labels lithologiques compliqués, besoin de simplifier pour obtenir un traitement.
- Zone faiblement graphitisée = 0 à 2 // zone fortement graphitisée = 3 à 5 → minimisation du biais humain dans la notation
- Rééchantillonnage. 577 échantillons pour chaque classe (modèle aléatoire = 50% de succès)

Sélection des variables

Méthode	Méthodes par filtre (<i>filter methods</i>) = <i>filtre avant le modèle</i>	Méthodes enveloppantes (<i>wrapper methods</i>) = <i>sélection depuis le modèle</i>
Principe	Sélection des variables selon un critère <i>a priori</i> (corrélation, information mutuelle,...)	Sélection itérative des variables selon leur capacité à améliorer le modèle
Inconvénients	Difficile de savoir si le critère est judicieux dans la sélection	Lenteur lorsque vaste jeux de données
Avantage	Plus rapide sur des vastes données	Meilleures performances globales

- Quelles variables utiliser pour prédire une valeur de graphite ?
- 100 teneurs géochimiques + 2150 longueurs d'onde IR + gamma + visible
- Première étape = réduction de dimensionalité (ACP sur spectral)
- 203 variables
- Méthodes enveloppantes globalement plus efficaces.

Graphitisation

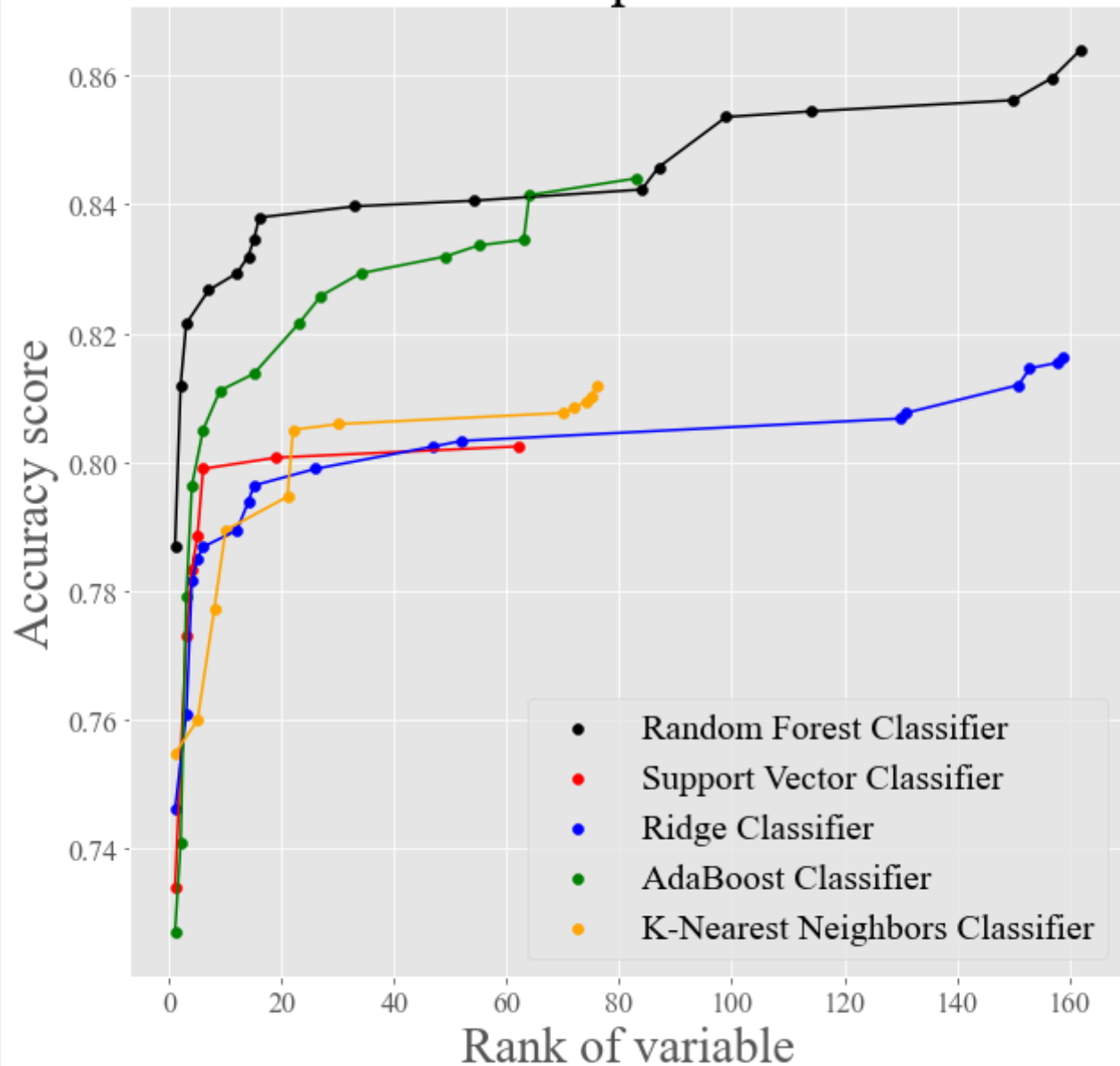


- Sélection par filtre pas idéal
- Méthodes enveloppantes : on choisit les critères en fonction du modèle

Prédiction avec sélection par filtre (critère = information mutuelle)

Graphitisation

Graphite



Etape initiale : Calcul de n prédictions pour chaque variable

Classement des prédictions par précision décroissante

Premier modèle = Prédiction avec variable de rang 1

$i = 1$

Calcul d'une prédiction avec la variable de rang $i + 1$ + variable de rang $i + 1$

→ Si précision augmente, variable de rang $i+1$ ajoutée au modèle

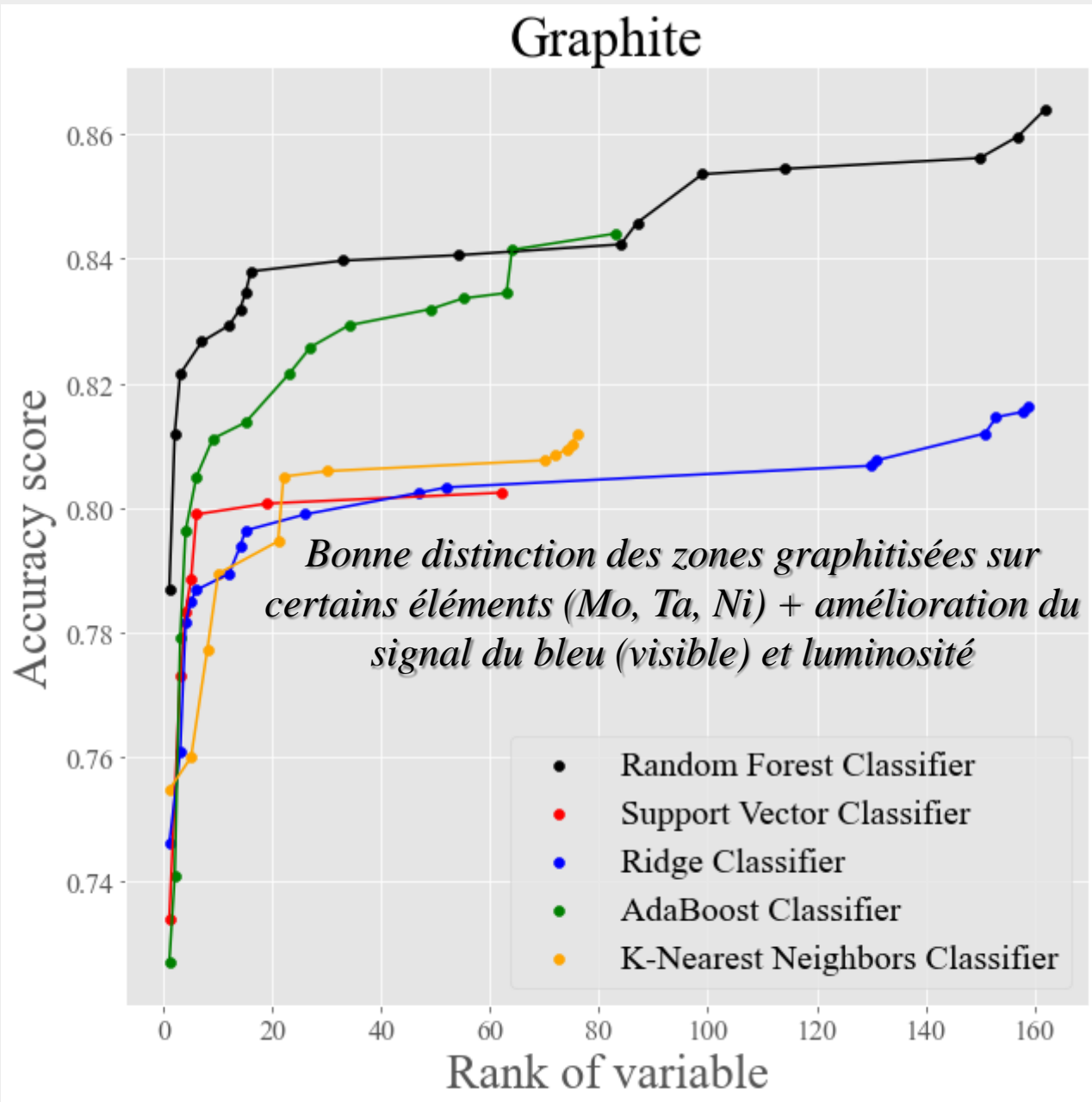
→ Si précision diminue, variable de rang $i+1$ retirée du modèle

$i = i+1$

...

→ Classification ascendante ; alternative = descendante (commencer avec toutes les variables et les retirer pour observer l'évolution du modèle)

Graphitisation



Etape initiale : Calcul de n prédictions pour chaque variable

Classement des prédictions par précision décroissante

Premier modèle = Prédiction avec variable de rang 1

$i = 1$

Calcul d'une prédiction avec la variable de rang $i + 1$ + variable de rang $i + 1$

→ Si précision augmente, variable de rang $i+1$ ajoutée au modèle

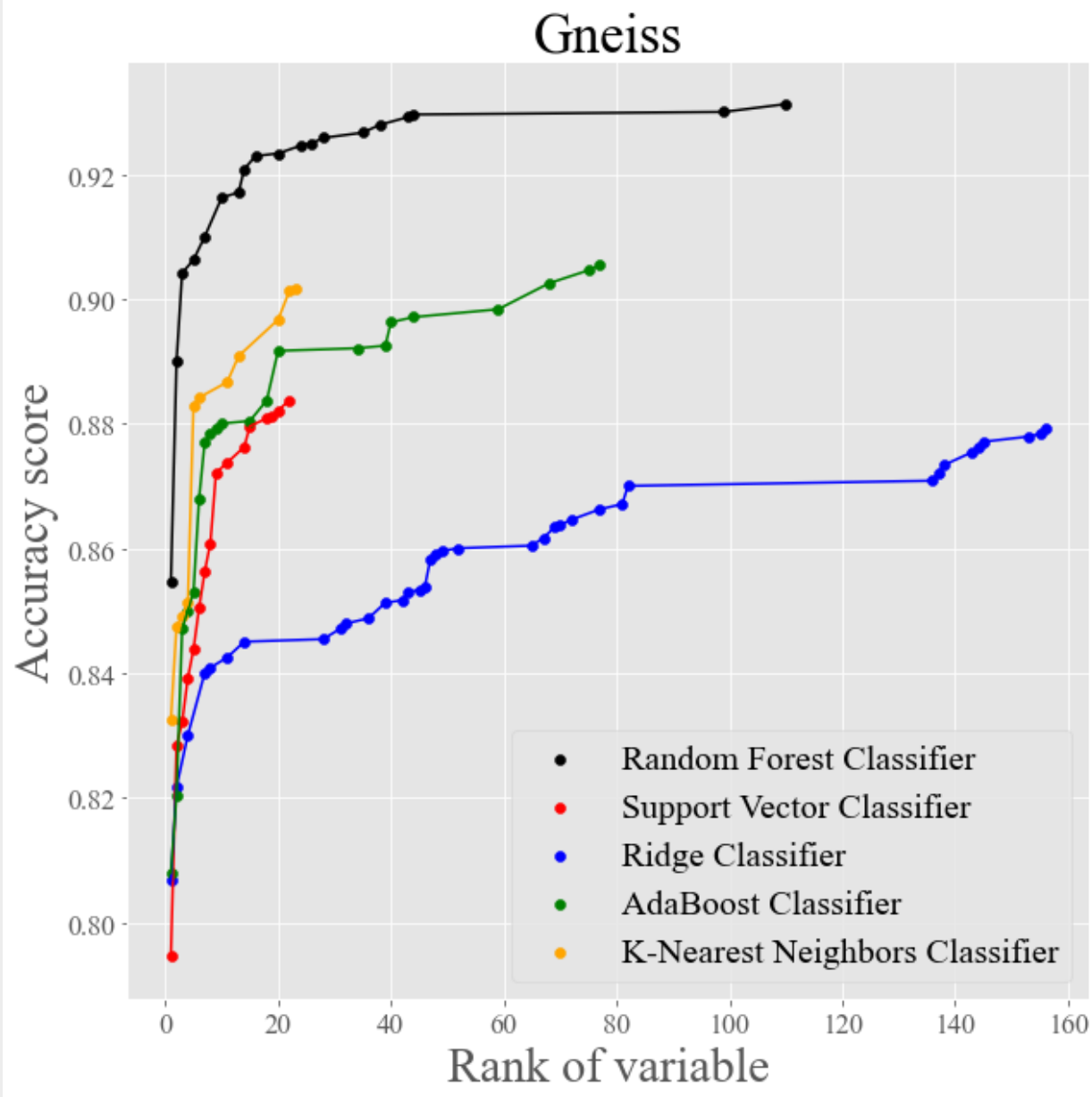
→ Si précision diminue, variable de rang $i+1$ retirée du modèle

$i = i+1$

...

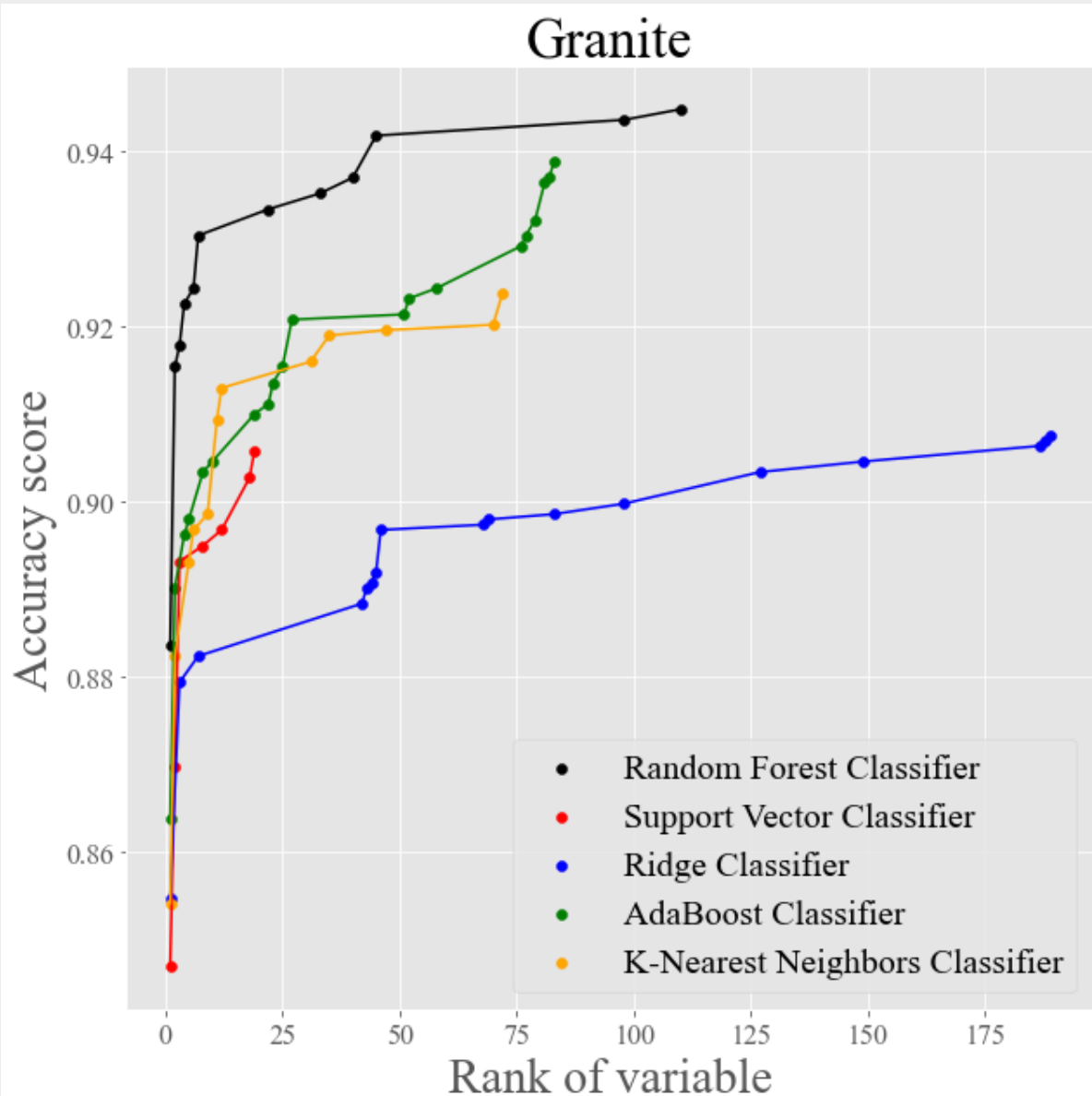
→ Classification ascendante ; alternative = descendante (commencer avec toutes les variables et les retirer pour observer l'évolution du modèle)

Lithologies



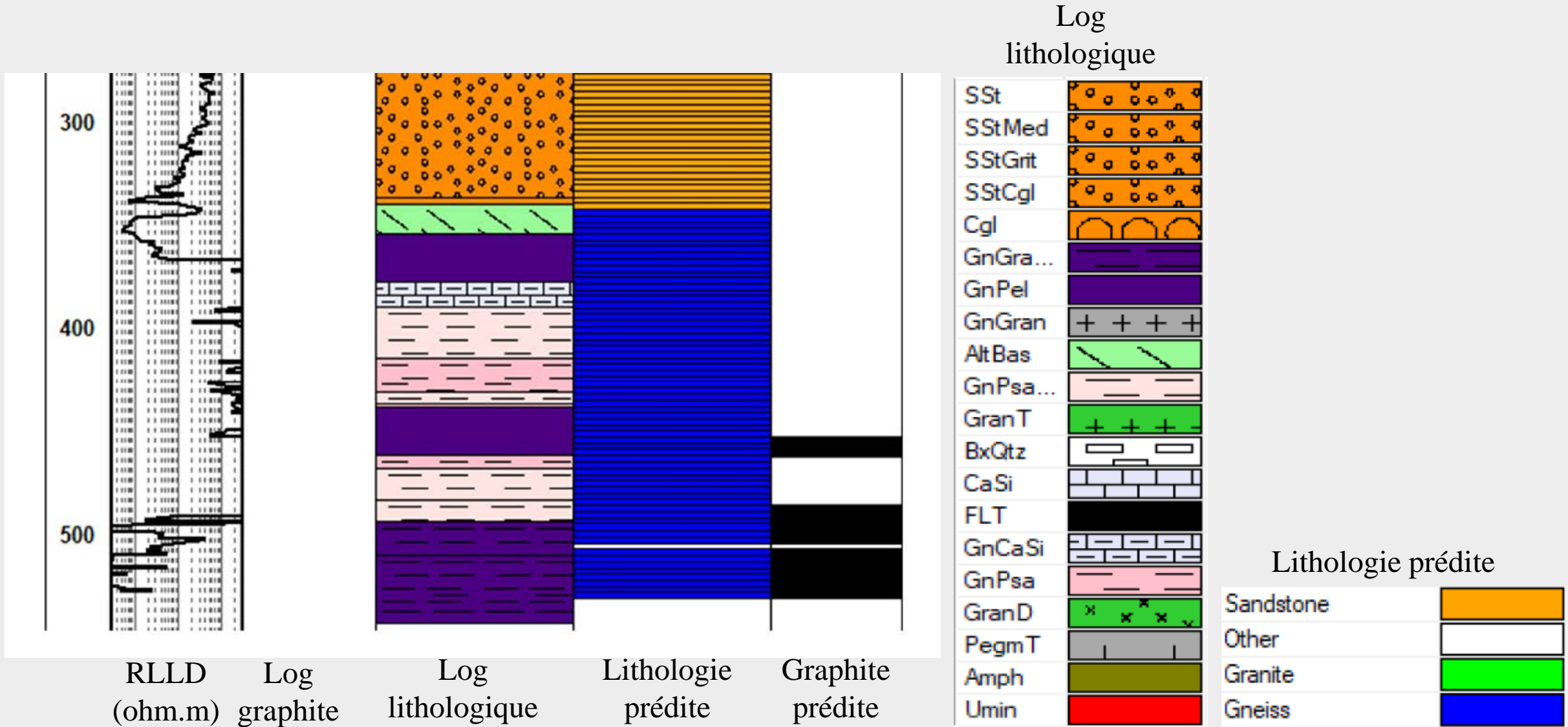
- Conversion des lithologies en *variables muettes* (0 si absence de la lithologie, 1 si présence)
- Simplification des lithologies :
$$\text{Gneiss} = \text{GnGraGra} + \text{GnPel} + \text{GnGra} + \text{GnPsaPel} + \text{GnBi} + \text{GnAlt} + \text{GnCaSi} + \text{GnPsa} + \text{GnCo}$$
- Possibilité de relocaliser les résultats des prédictions sur les logs
- Détection des lithologies gneissiques = 93,2%

Lithologies

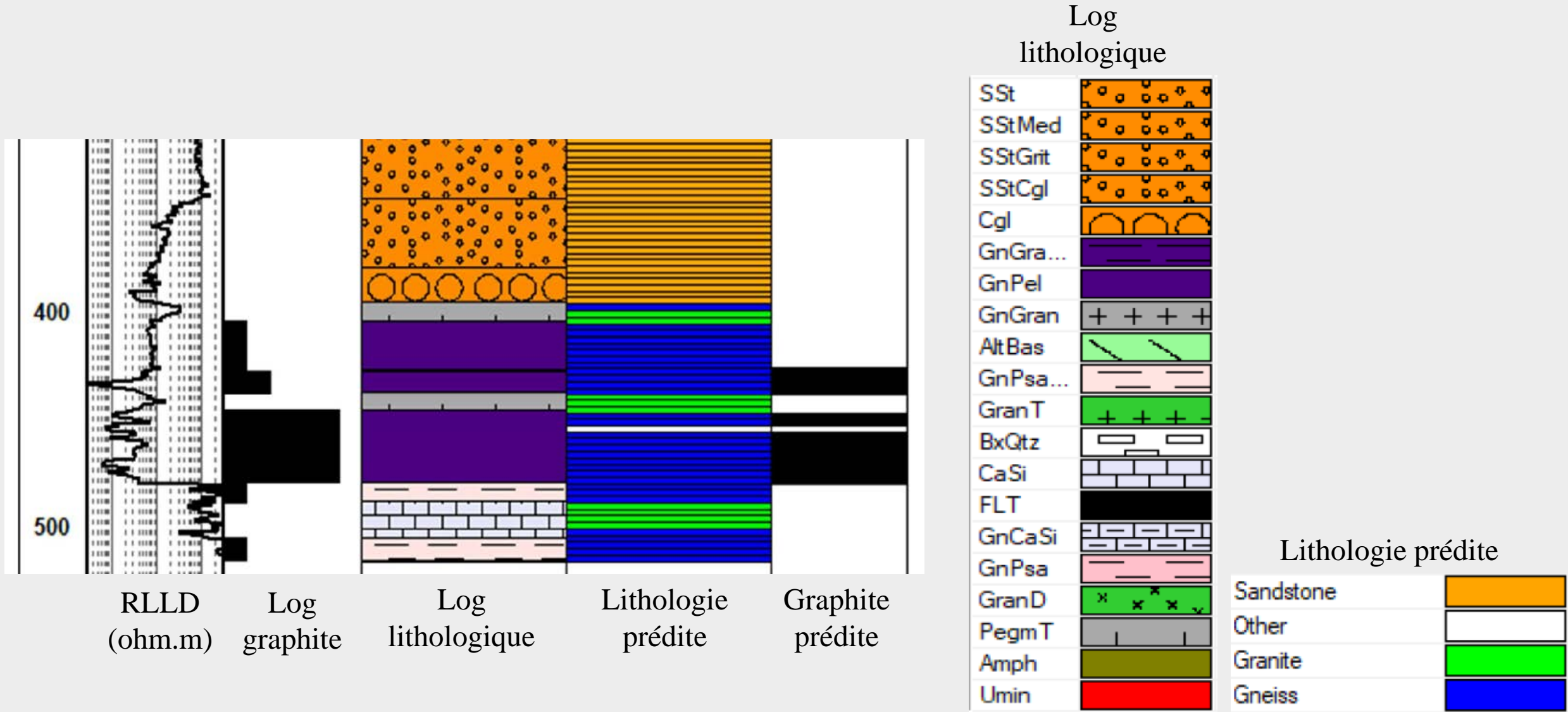


- Granite = regroupement de tous les labels à lithologie granitique
- Bonne distinction gneiss/granite sur certains éléments chimiques (Ta, K₂O, Hg, Cr, Rb)

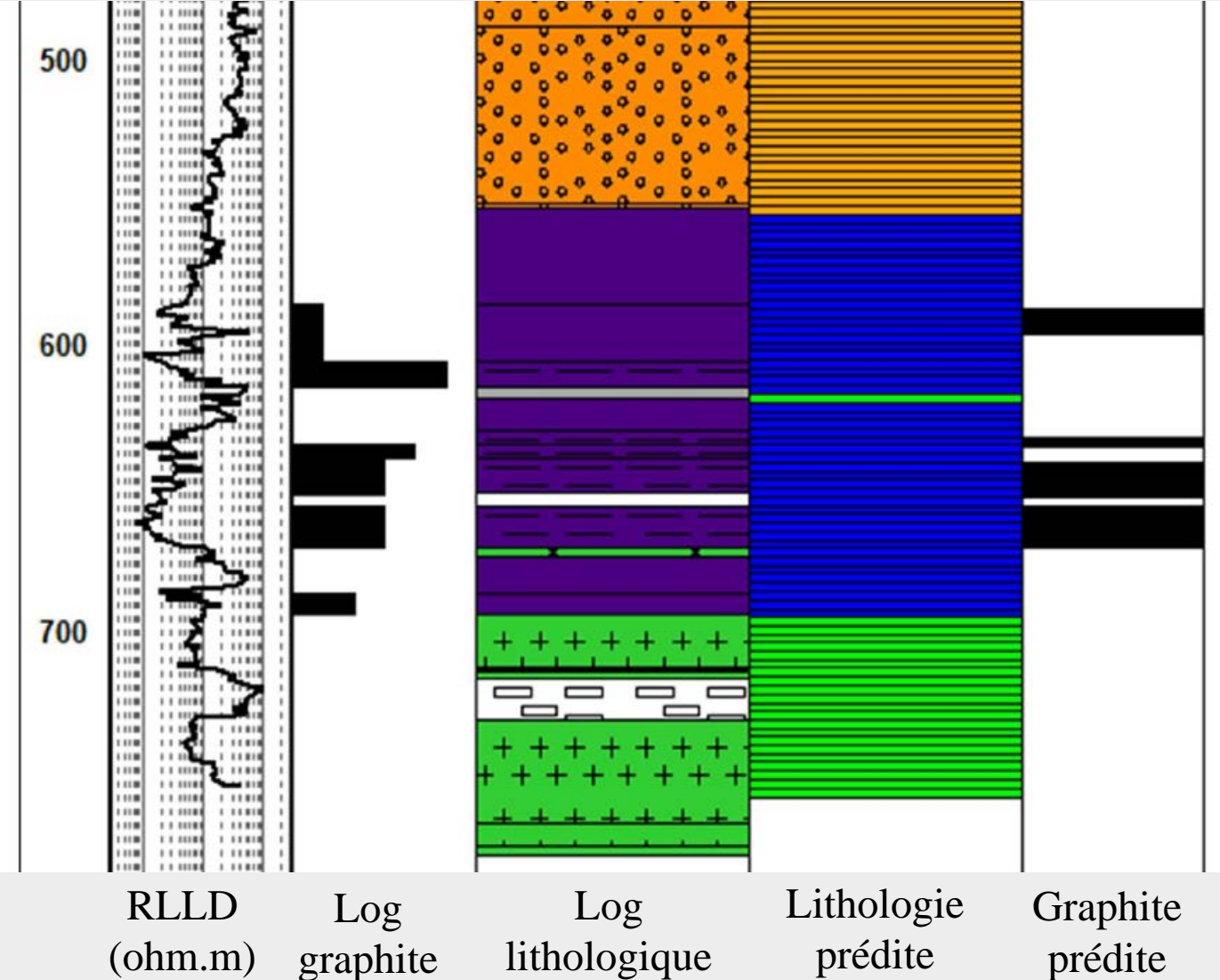
Lithologies



Lithologies



Lithologies



Log lithologique

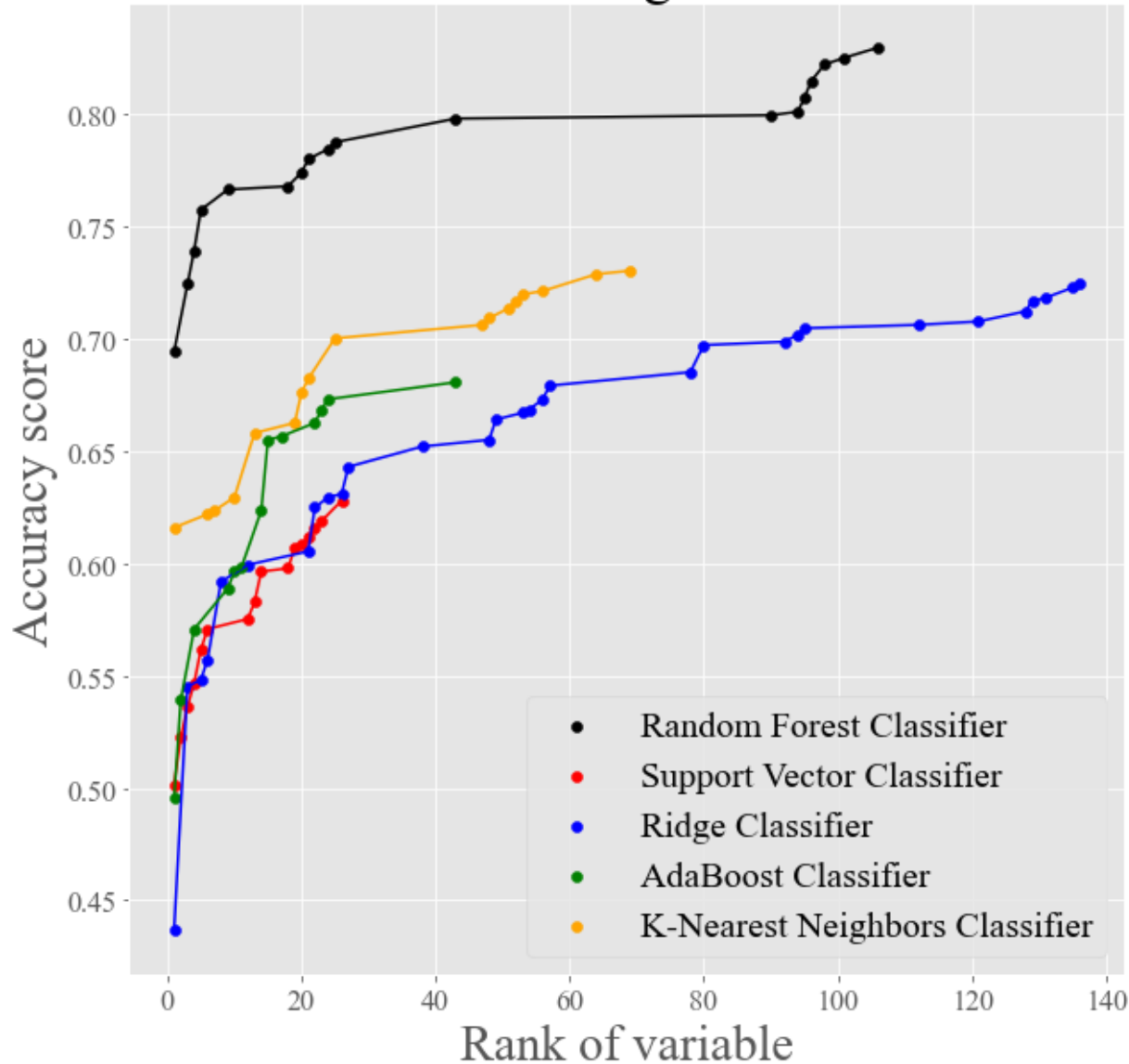
SSt	
SStMed	
SStGrit	
SStCgl	
Cgl	
GnGra...	
GnPel	
GnGran	
AltBas	
GnPsa...	
GranT	
BxQtz	
CaSi	
FLT	
GnCaSi	
GnPsa	
GranD	
PegmT	
Amph	
Umin	

Lithologie prédite

Sandstone	
Other	
Granite	
Gneiss	

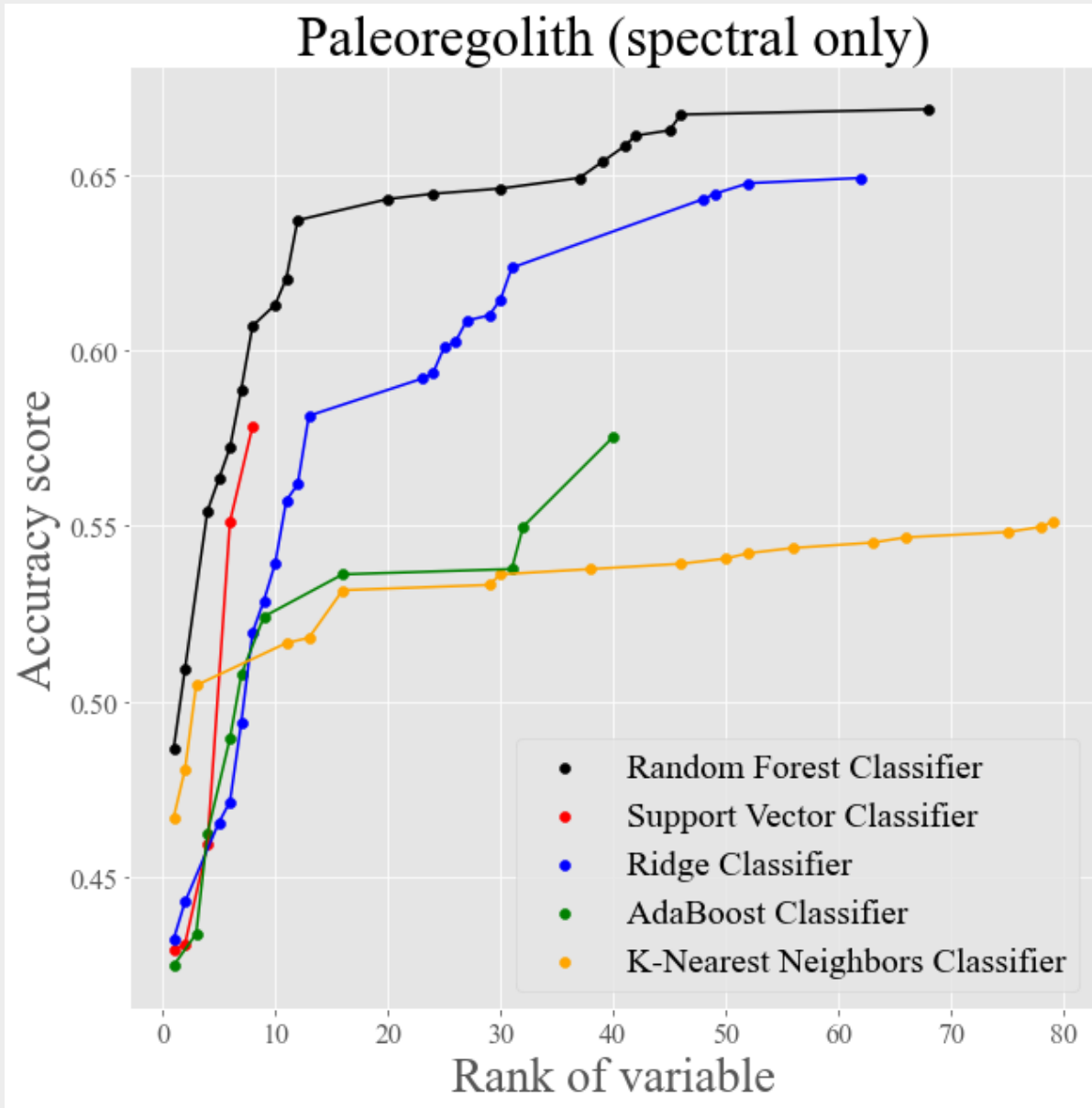
Paléorégolithe

Paleoregolith



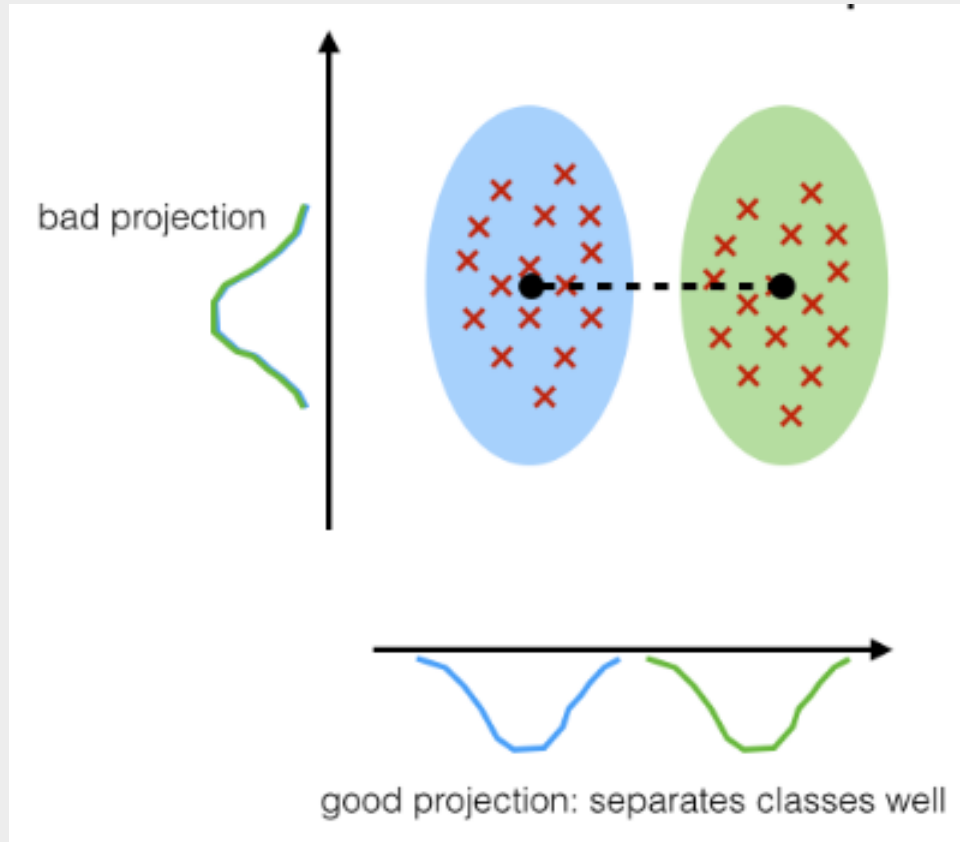
- Paléorégolithe noté de manière constante dans les logs
- Spectral adapté à sa description ?
- Besoin de rééchantillonner, 99% d'échantillons de bassin = non altérés
- Classes notées : Green, Red, Red-Green, Green-Red, None
- Simplification : 4 classes / Red, Green, None, Mixed
- 82,98% de précision

Paléorégolithe (spectral seul)



- Cas du spectral seul
- 66% de précision

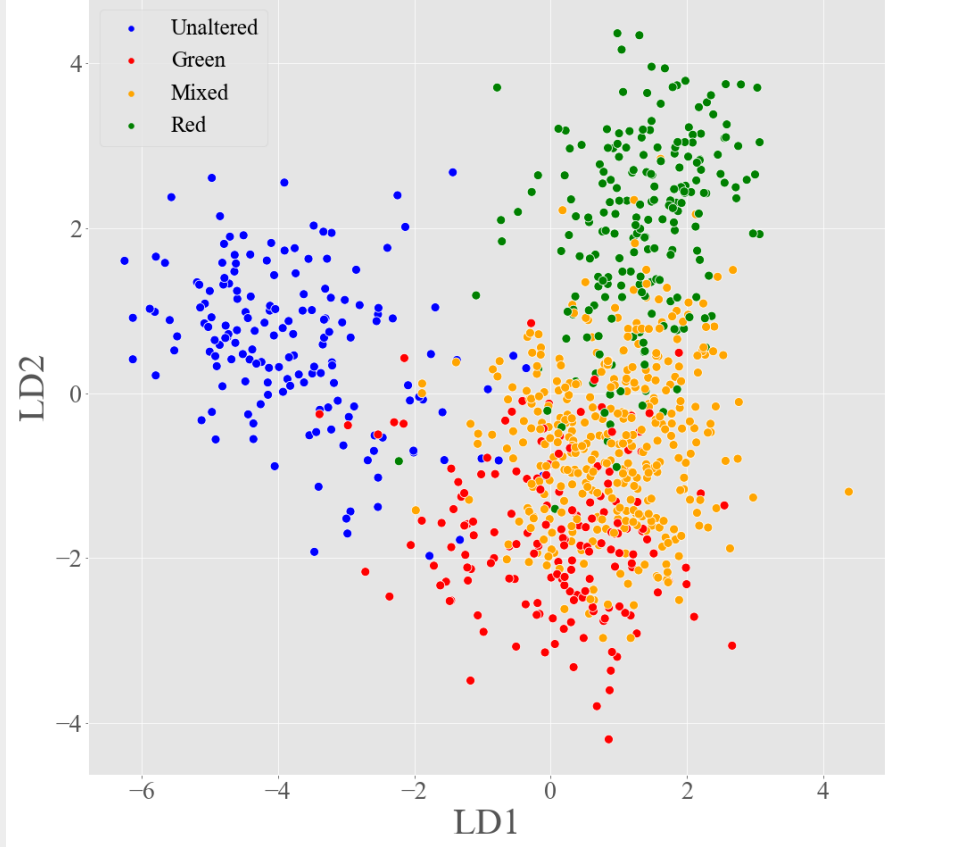
Paléorégolithe



- Analyse discriminante = technique supervisée de discrimination (=/= classification)
- Reprojection des données d'entrée dans un espace à n discriminants décorrés

Paléorégolithe

Lineary Discriminant Analysis, Wavelet Transform, scale 3, wavelet = db10

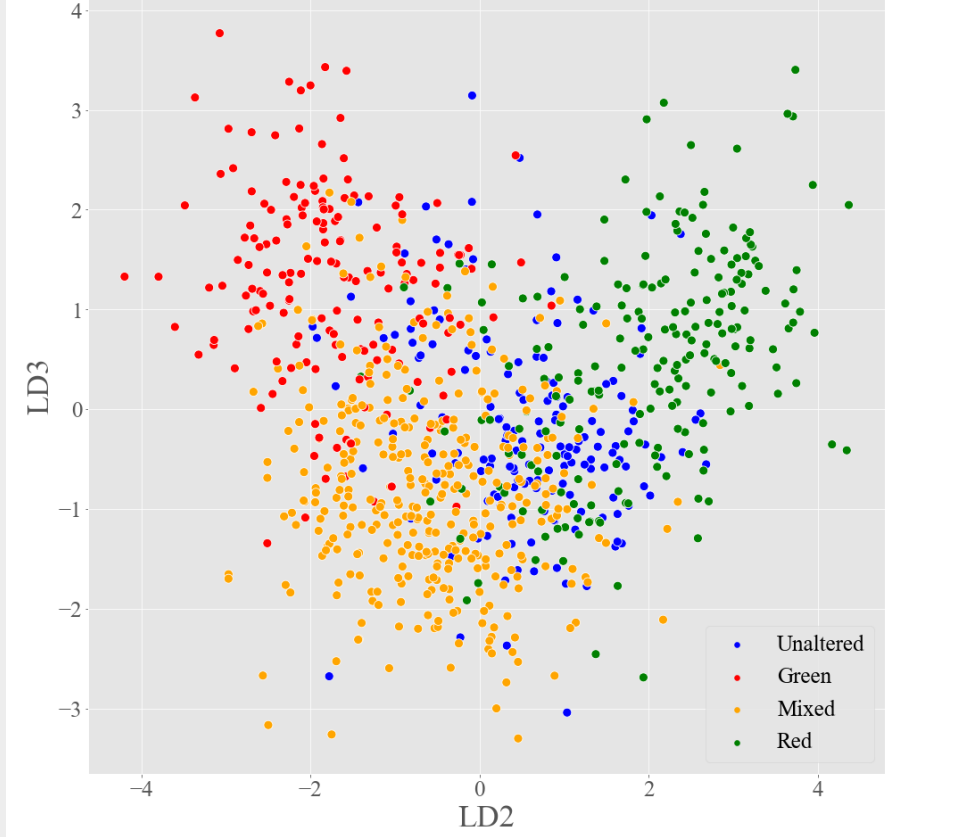


- Test de différentes stratégies de pré-traitement
 - ➔ Réduction de dimensions par ondelettes/ACP
 - ➔ Retrait / non retrait du continuum
- Meilleur résultat = 71,2%

	All data	PCA								Wavelet							
		6	8	10	12	14	16	18	20	1	2	3	4	5	6	7	8
Raw																	
LDA	0,519	0,575	0,622	0,624	0,629	0,639	0,646	0,648	0,648	0,633	0,672	0,688	0,674	0,664	0,625	0,570	0,574
QDA	0,258	0,598	0,619	0,615	0,616	0,614	0,602	0,606	0,606	0,272	0,366	0,617	0,670	0,625	0,603	0,597	0,570
Continuum removed																	
LDA	0,472	0,556	0,573	0,586	0,589	0,615	0,617	0,632	0,633	0,610	0,663	0,712	0,696	0,652	0,609	0,549	0,508
QDA	0,264	0,518	0,516	0,528	0,543	0,556	0,559	0,583	0,597	0,283	0,367	0,606	0,671	0,616	0,553	0,524	0,463

Paléorégolithe

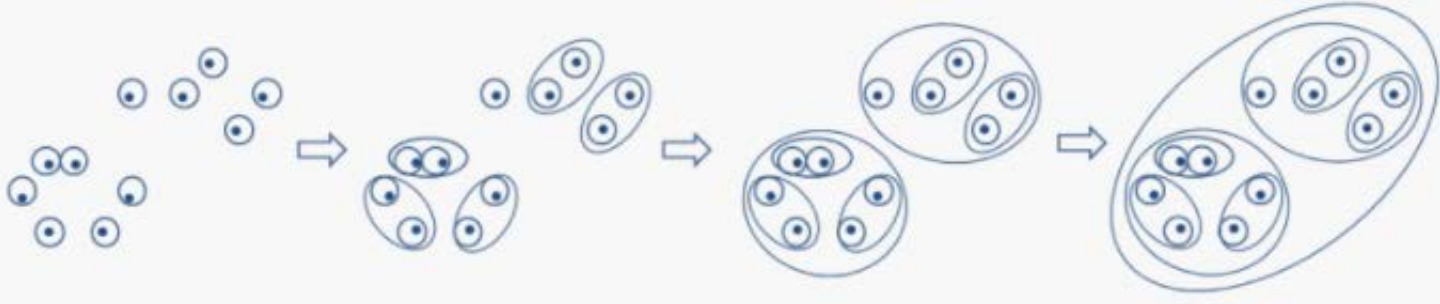
Lineary Discriminant Analysis, Wavelet Transform, scale 3, wavelet = db10



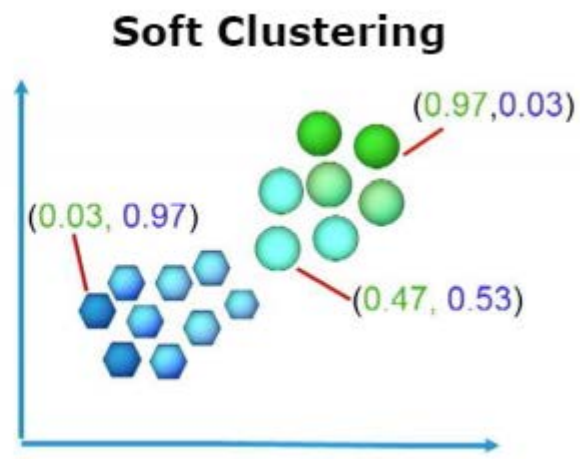
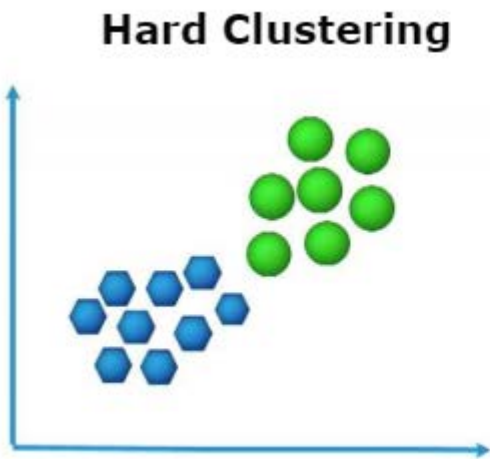
- Test de différentes stratégies de pré-traitement
- ➔ Réduction de dimensions par ondelettes/ACP
- ➔ Retrait / non retrait du continuum
- Meilleur résultat = 71,2% de précision

	All data	PCA								Wavelet							
		6	8	10	12	14	16	18	20	1	2	3	4	5	6	7	8
Raw																	
LDA	0,519	0,575	0,622	0,624	0,629	0,639	0,646	0,648	0,648	0,633	0,672	0,688	0,674	0,664	0,625	0,570	0,574
QDA	0,258	0,598	0,619	0,615	0,616	0,614	0,602	0,606	0,606	0,272	0,366	0,617	0,670	0,625	0,603	0,597	0,570
Continuum removed																	
LDA	0,472	0,556	0,573	0,586	0,589	0,615	0,617	0,632	0,633	0,610	0,663	0,712	0,696	0,652	0,609	0,549	0,508
QDA	0,264	0,518	0,516	0,528	0,543	0,556	0,559	0,583	0,597	0,283	0,367	0,606	0,671	0,616	0,553	0,524	0,463

Spectroscopie IR



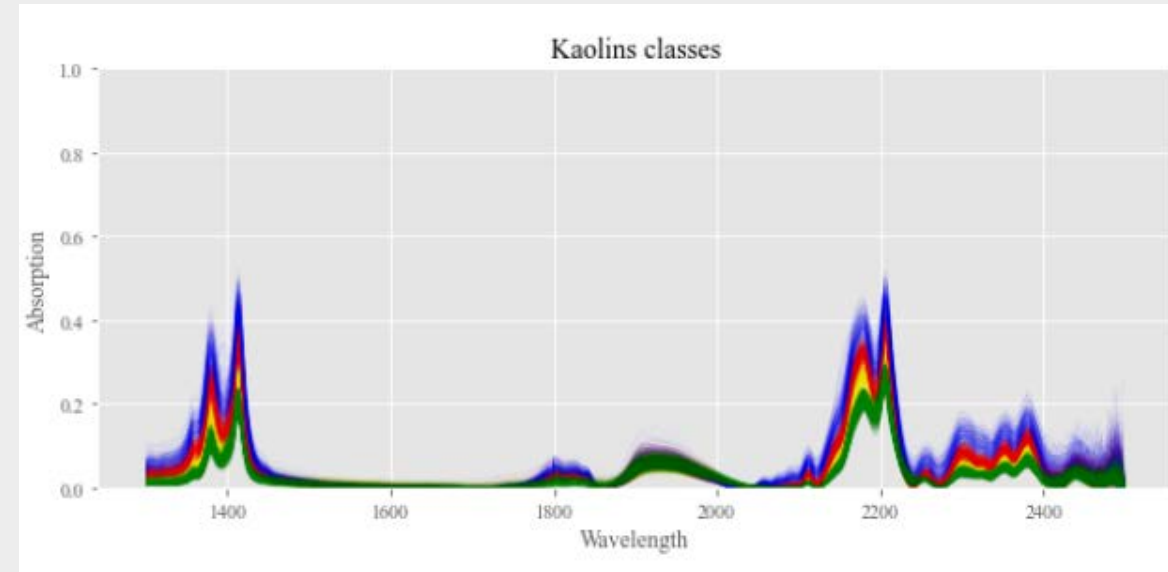
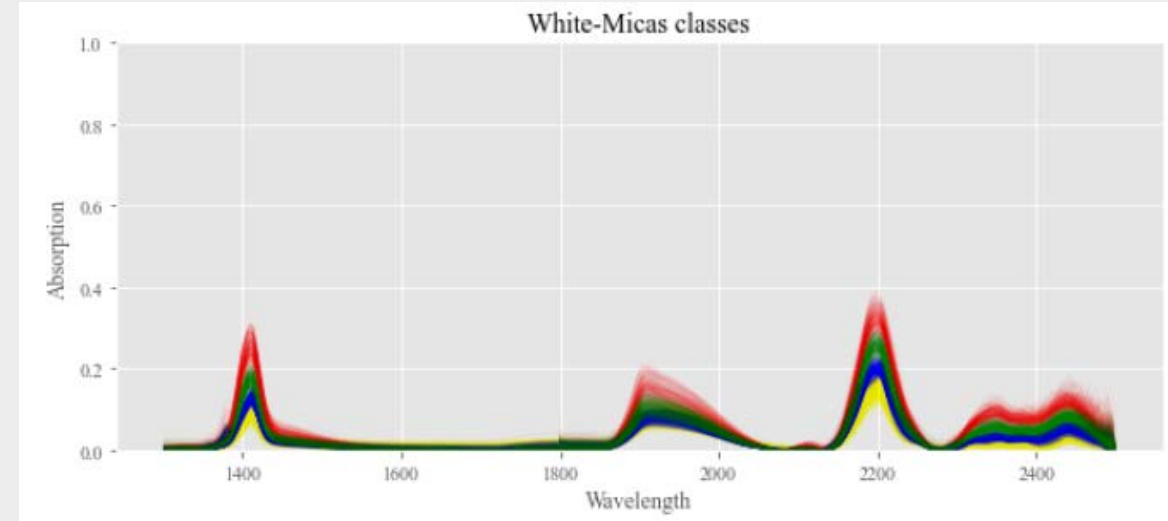
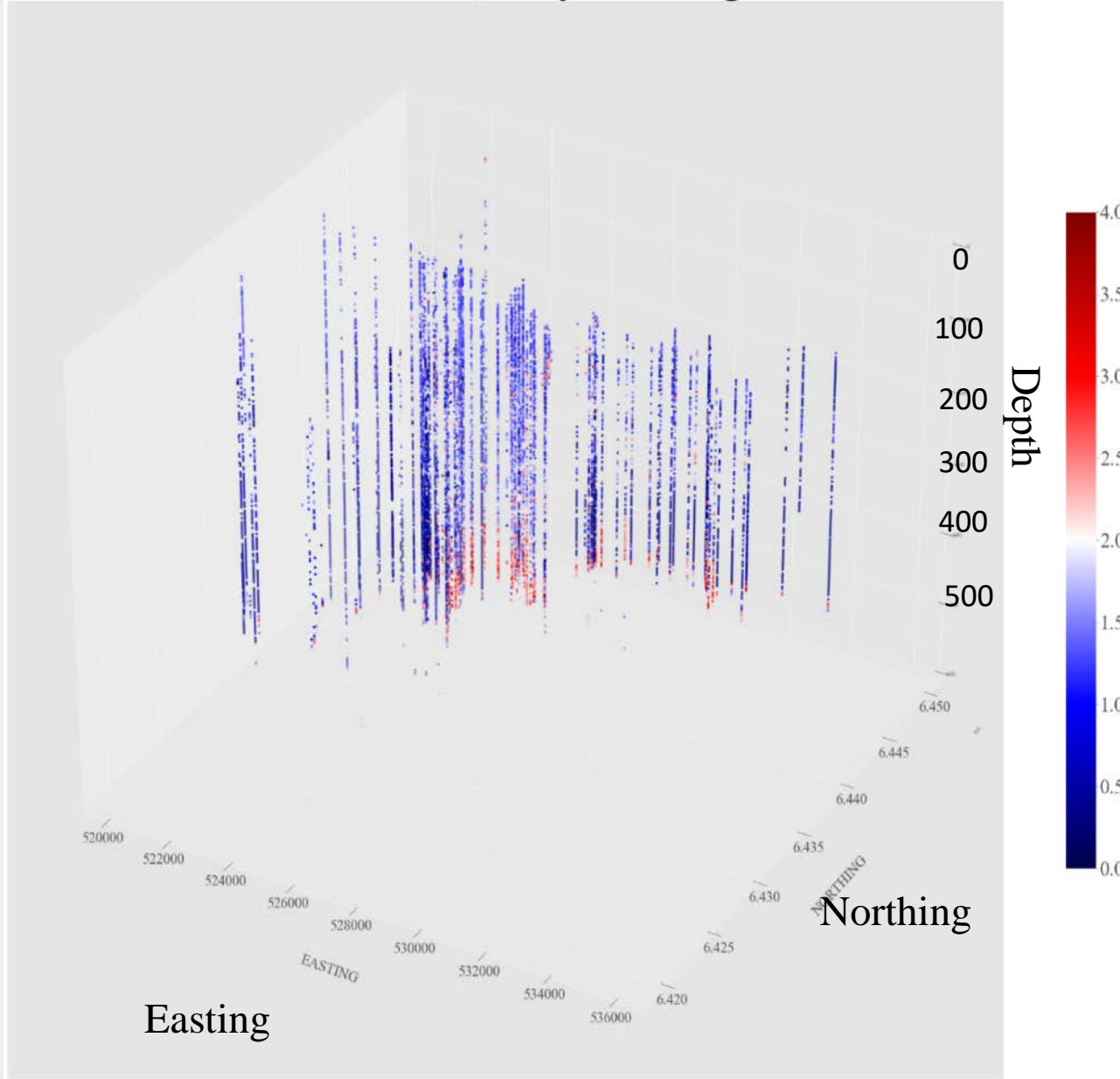
Classification par ascendance hiérarchique



- Lorsqu'absence de labels = classification non supervisée (*clustering*).
- Objectif : découvrir les regroupements naturels entre individus d'un jeu de données
- Méthodes les plus connues : hiérarchique, k-moyennes
- Méthode des c-moyennes (logique floue, *fuzzy clustering*)
- Regroupement flou = plus adapté à la description de mixtures binaires/ternaires des zones d'altération

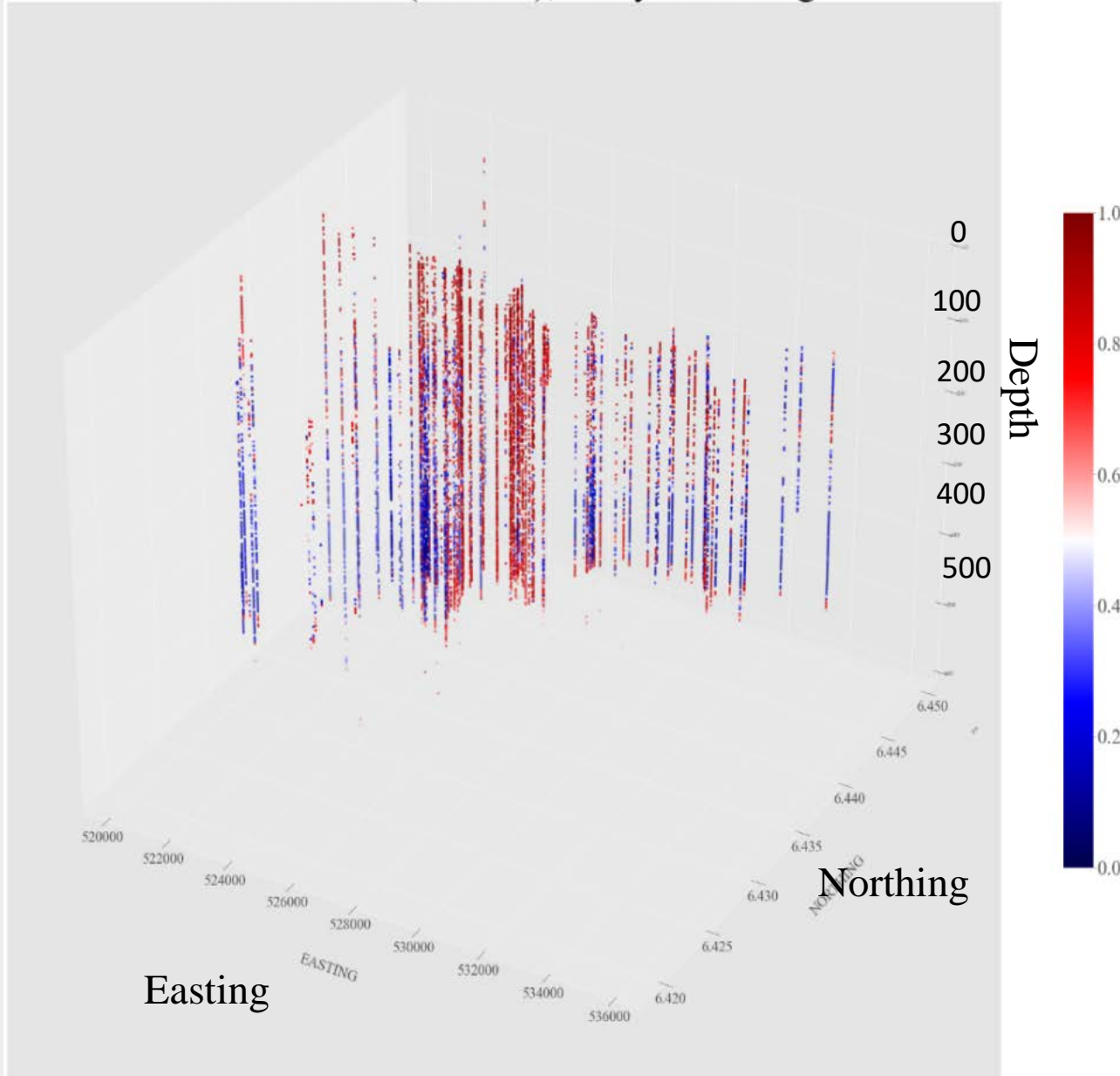
Spectroscopic IR

White-Micas, fuzzy clustering

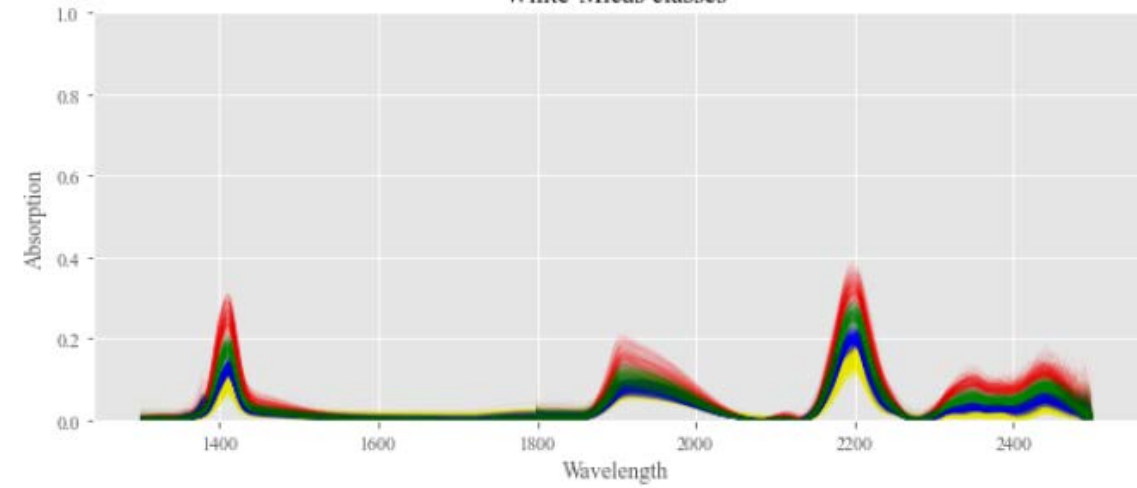


Spectroscopic IR

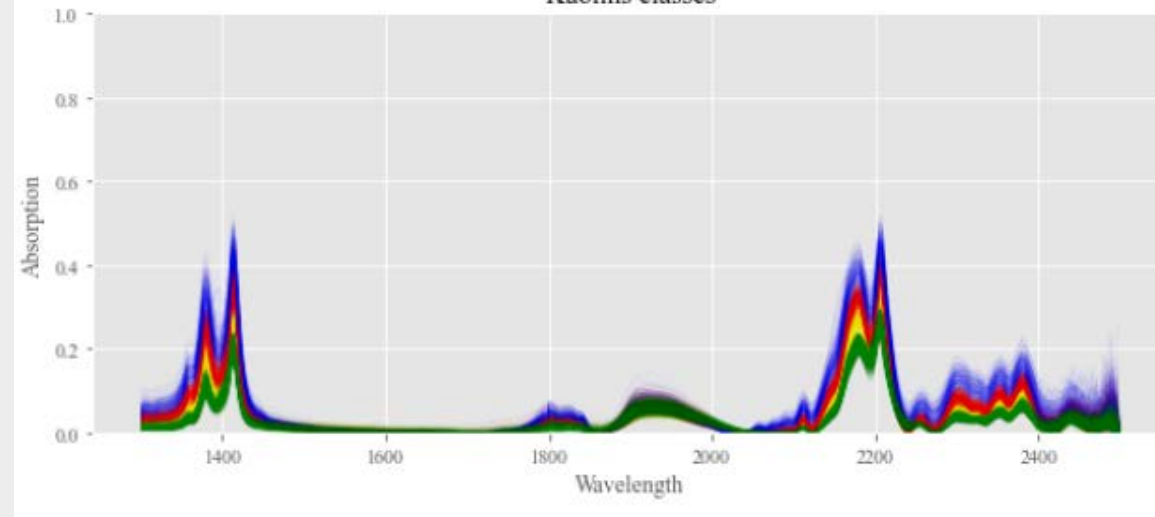
White-Micas (relative), fuzzy clustering



White-Micas classes

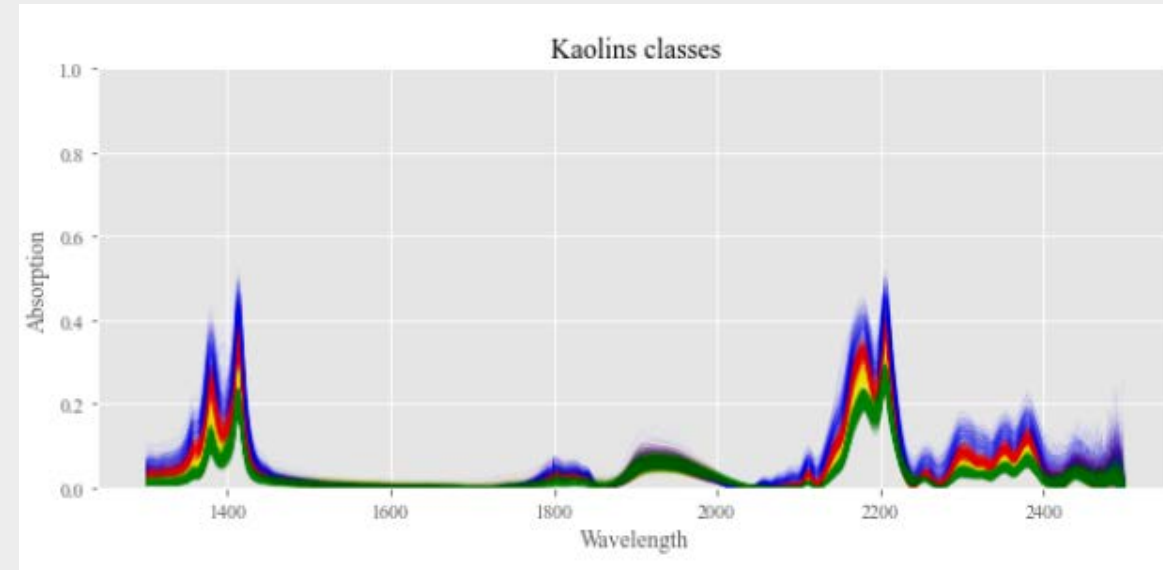
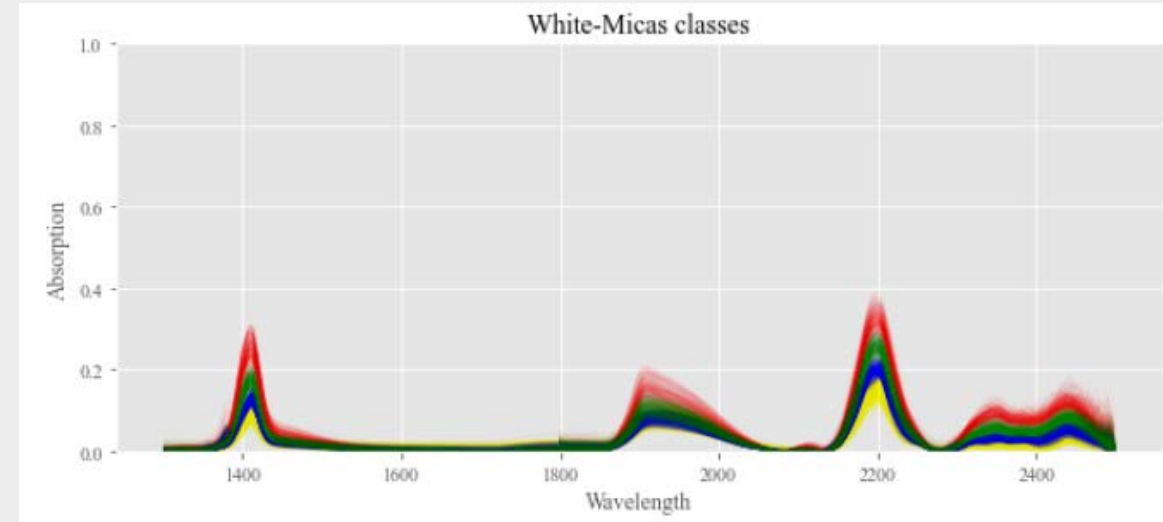
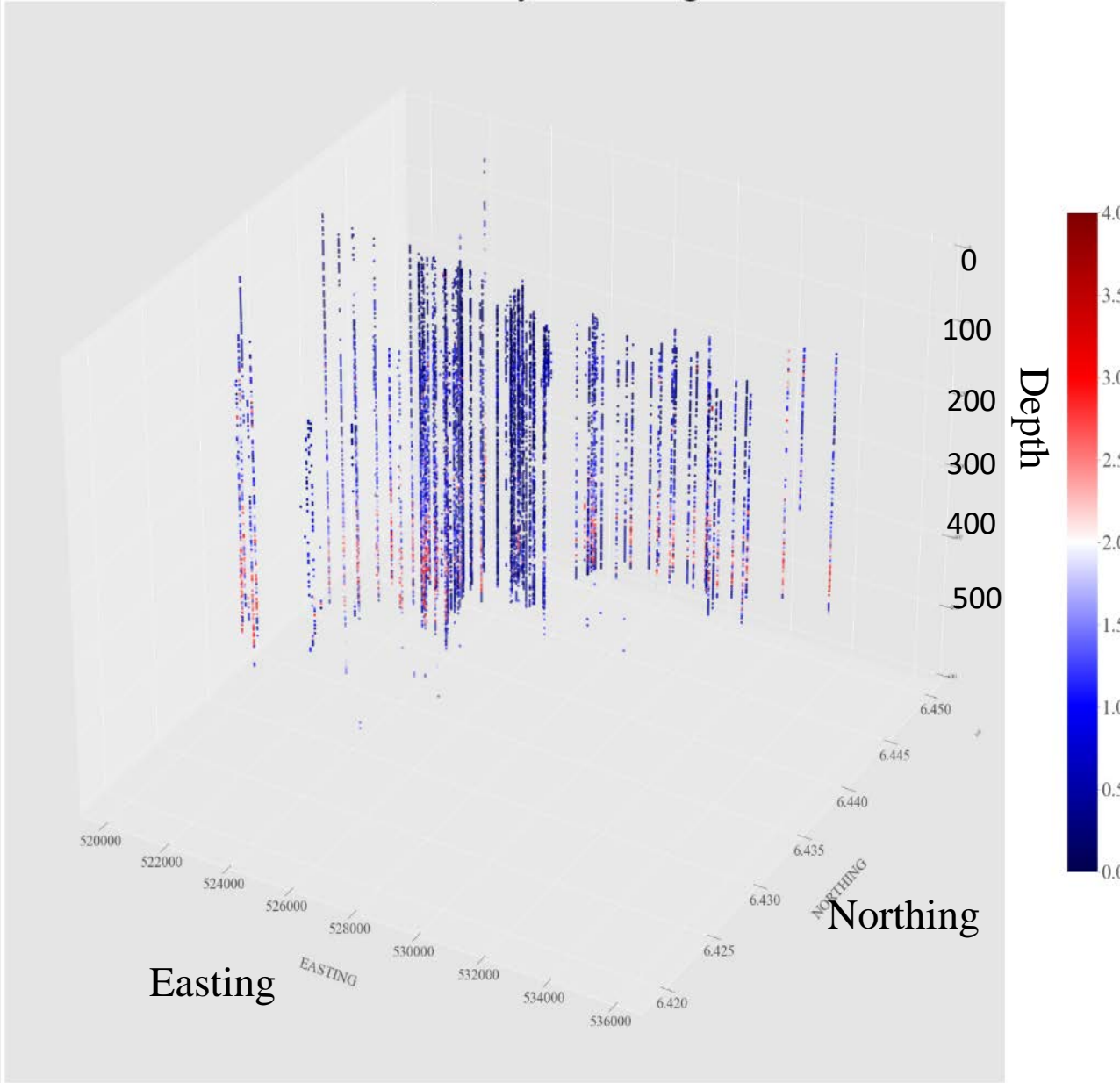


Kaolins classes



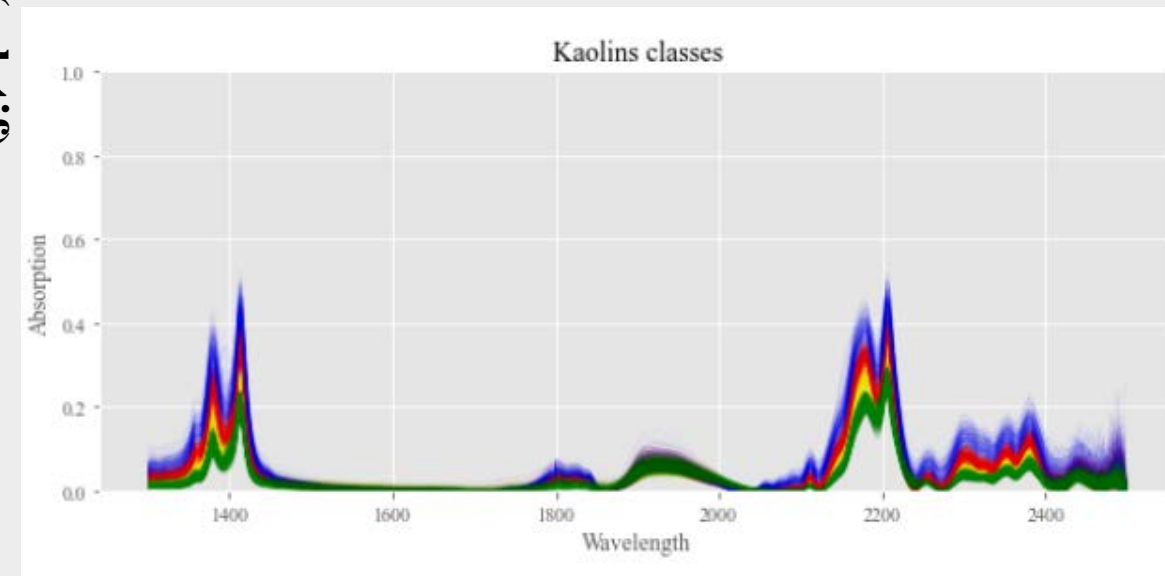
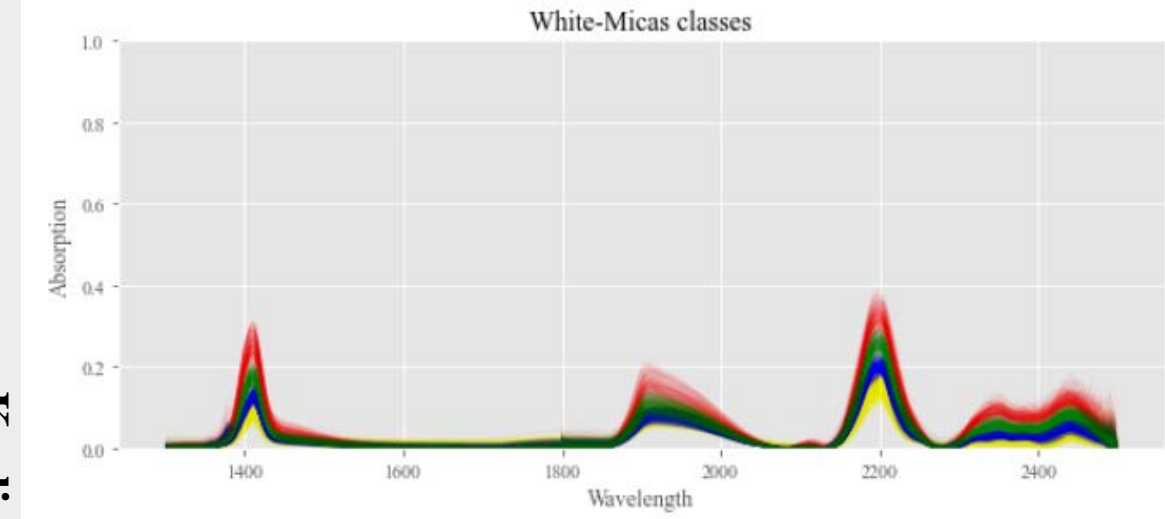
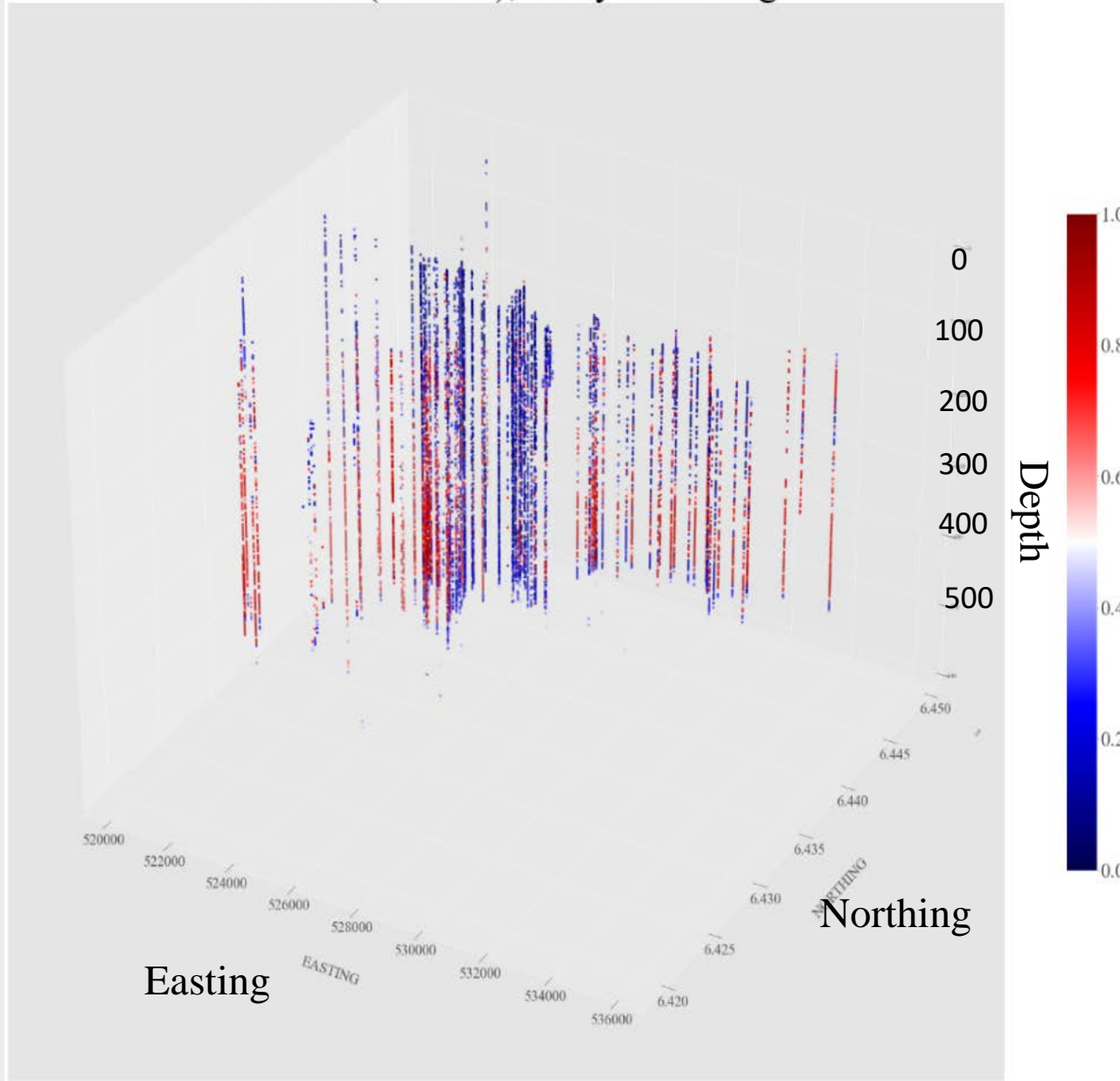
Spectroscopic IR

Kaolins, fuzzy clustering



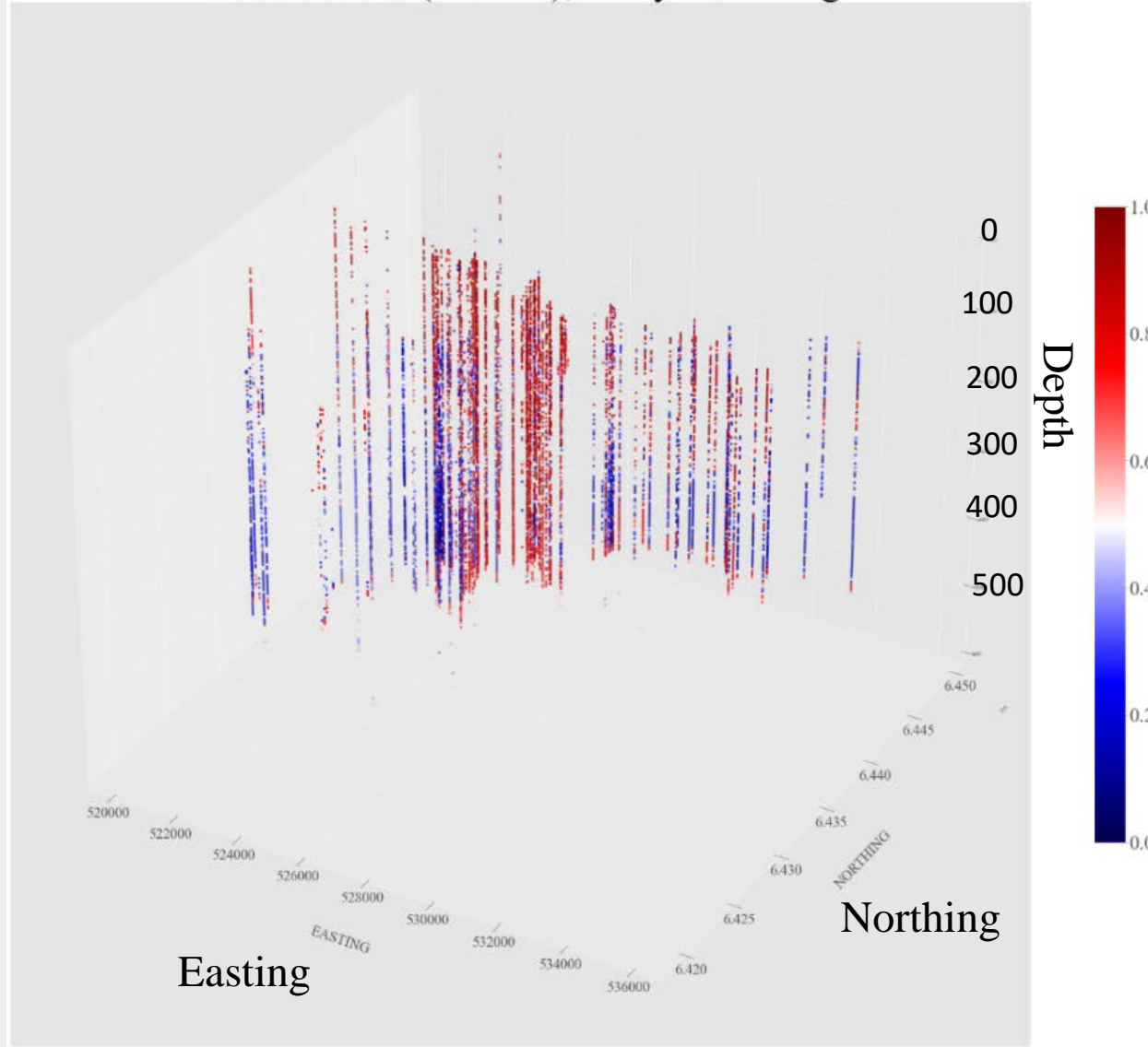
Spectroscopic IR

Kaolins (relative), fuzzy clustering

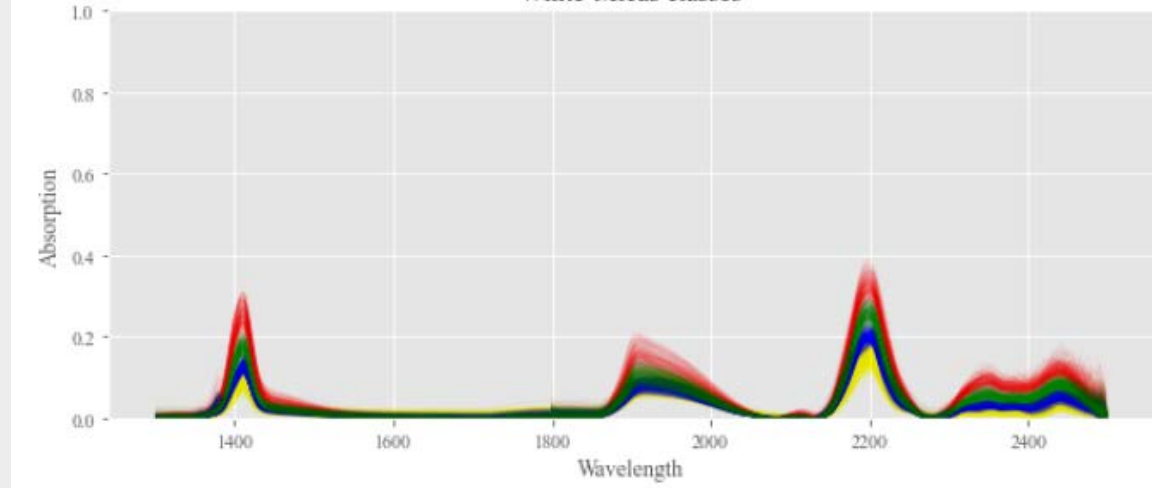


Spectroscopic IR

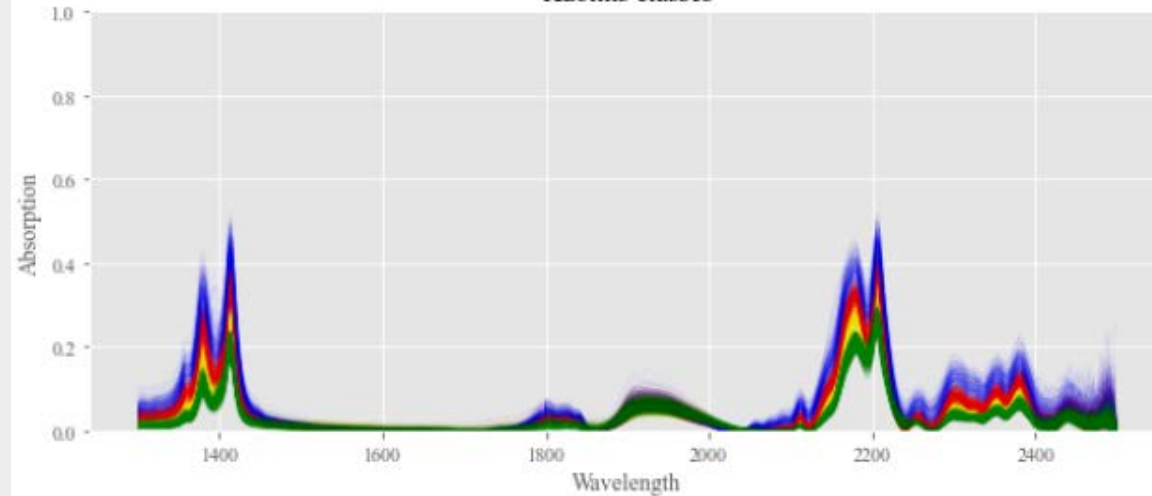
White-Micas (relative), fuzzy clustering



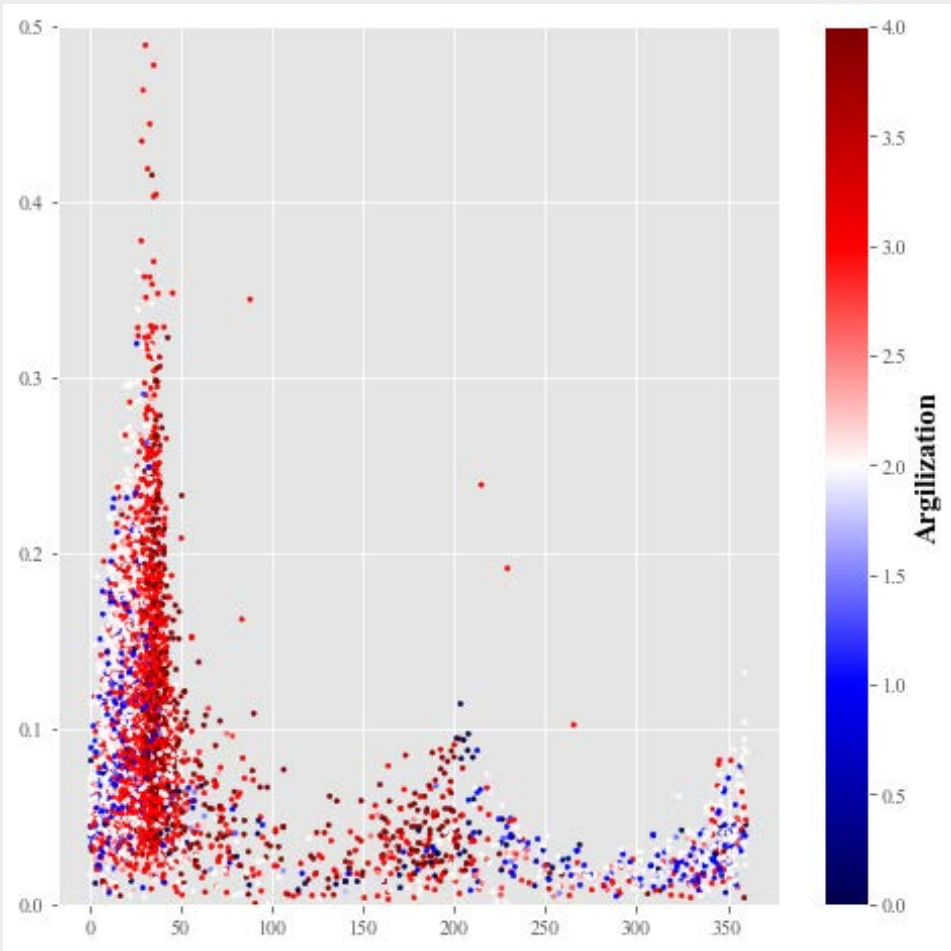
White-Micas classes



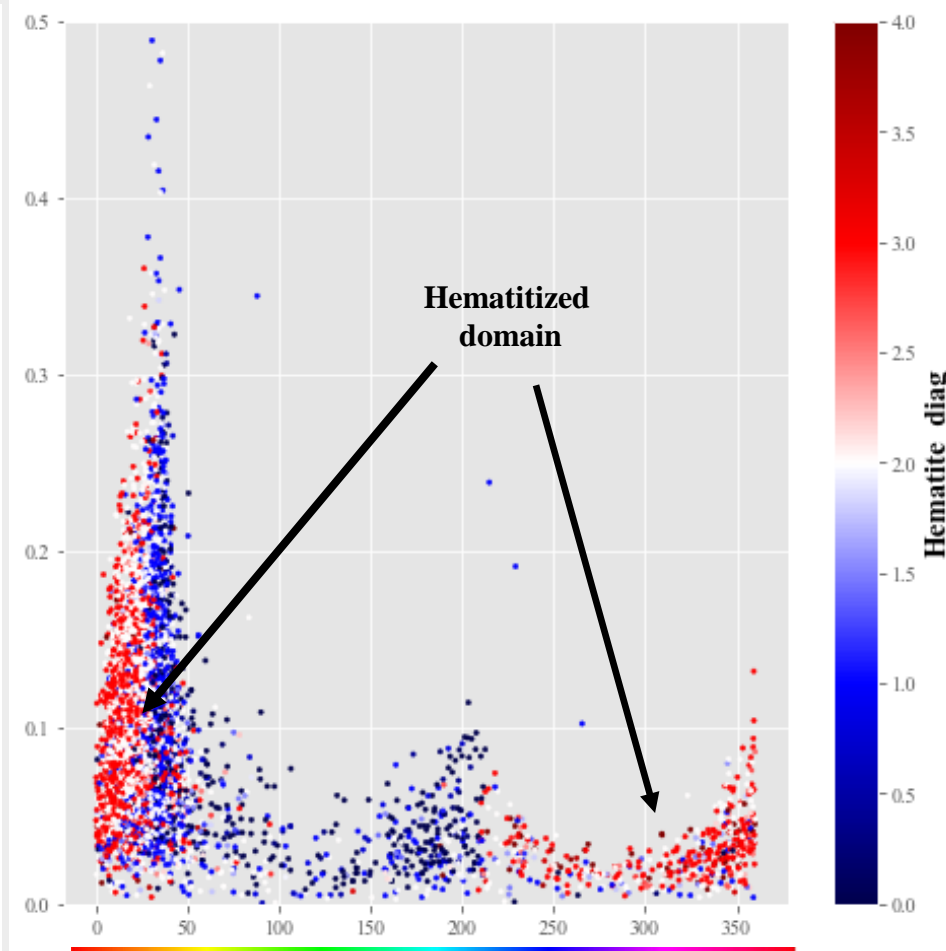
Kaolins classes



Conclusions



Hue



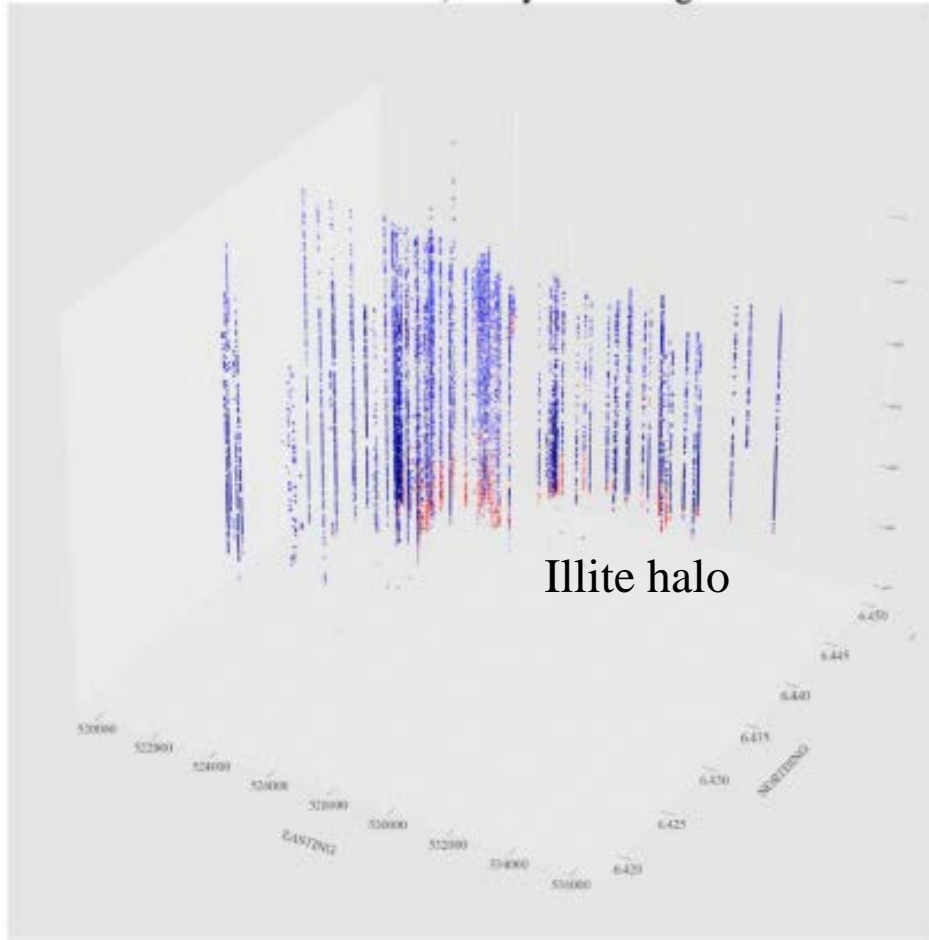
Hue

- Possibilité de déterminer plusieurs variables (paléorégolithe, lithologie, graphite) et des échelles d'altération (kaolins, micas blancs, chlorites) par des méthodes statistiques simples et accessibles
- Certaines données systématiques synchrones au processus d'exploration (spectral), ou le seront peut-être dans le futur (géochimie, avec évolution du XRF portable ?)
- Possibilité d'étendre ces routines à d'autres logs : granulométrie, hématite (propriétés optiques)

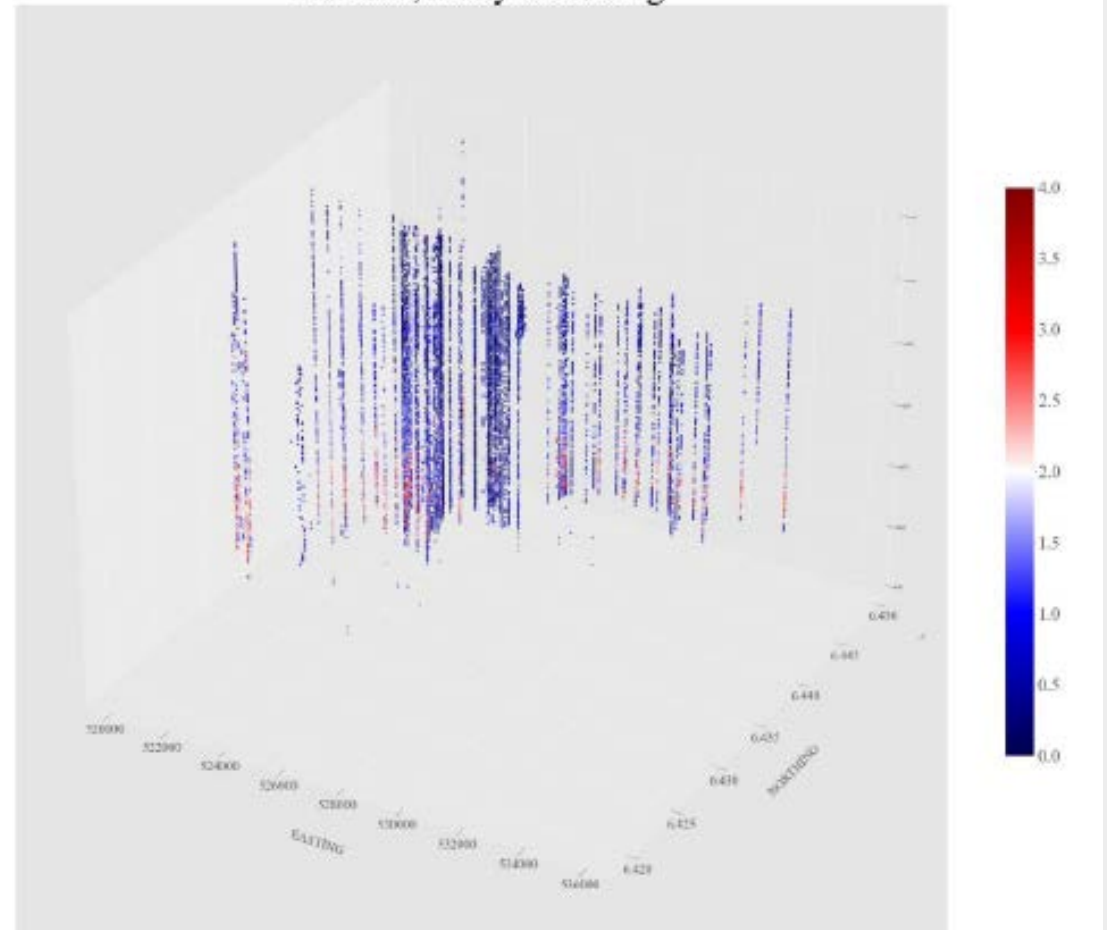
IR-spectra

Fuzzy clustering

White-Micas, fuzzy clustering

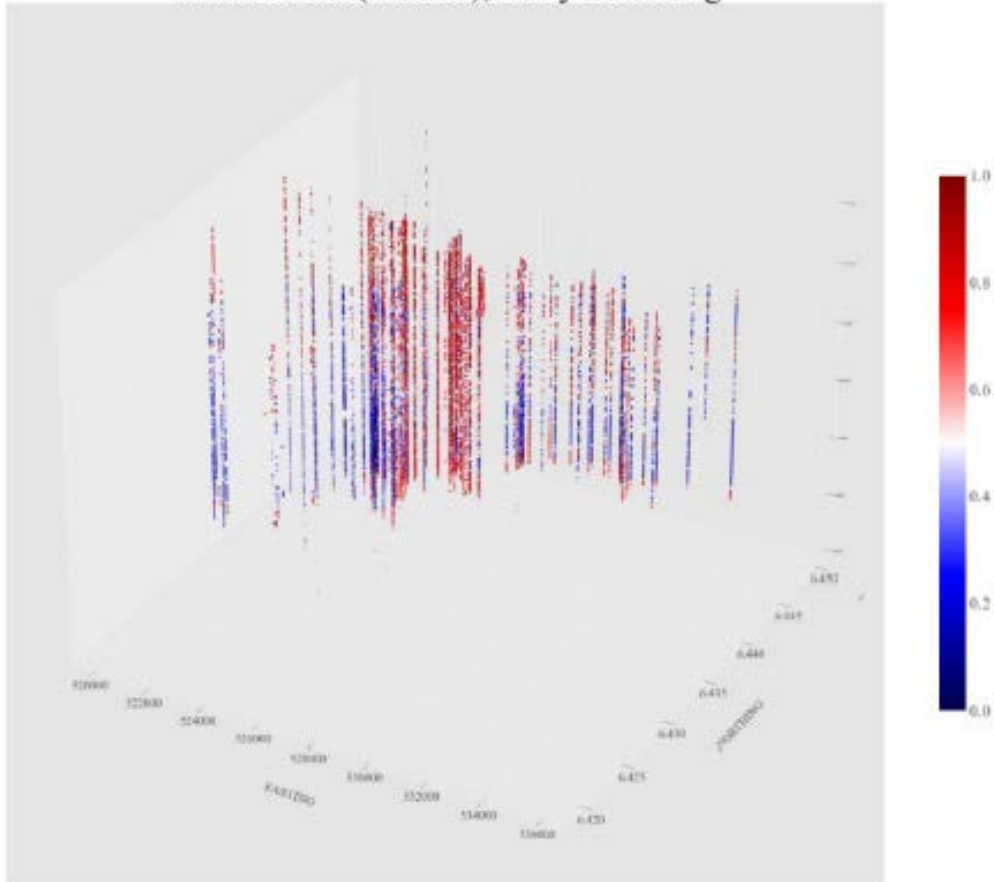


Kaolins, fuzzy clustering

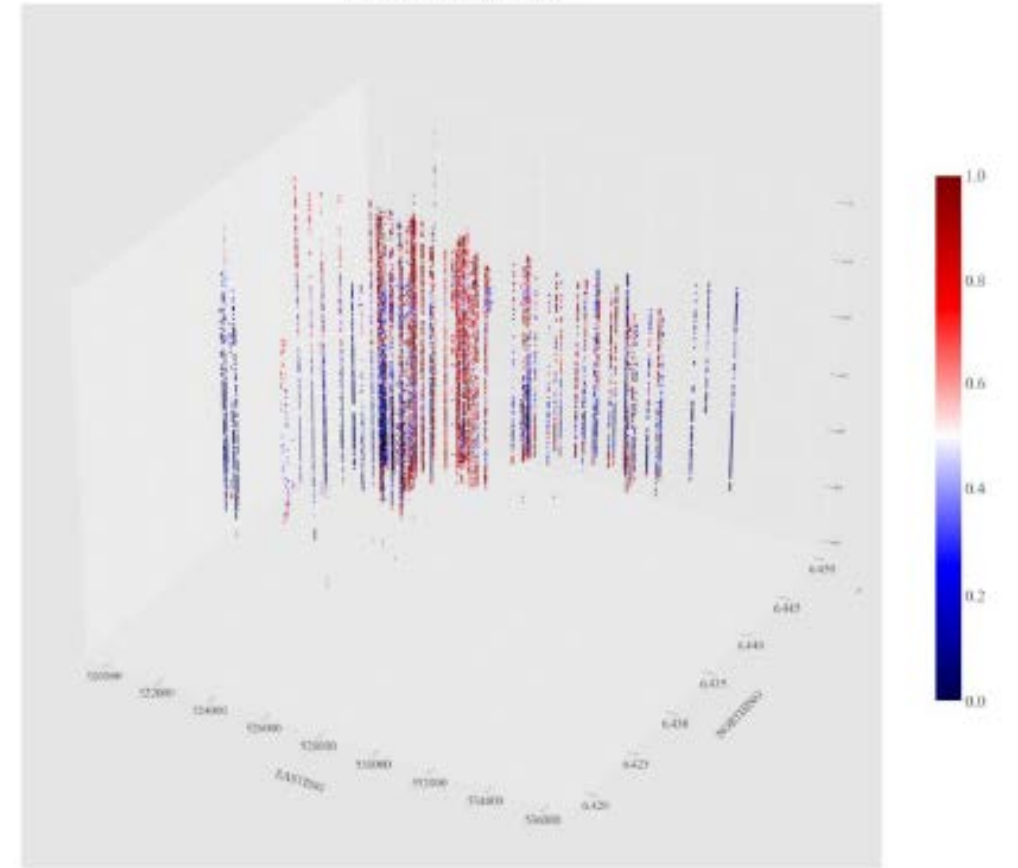


IR-spectra Fuzzy clustering

White-Micas (relative), fuzzy clustering

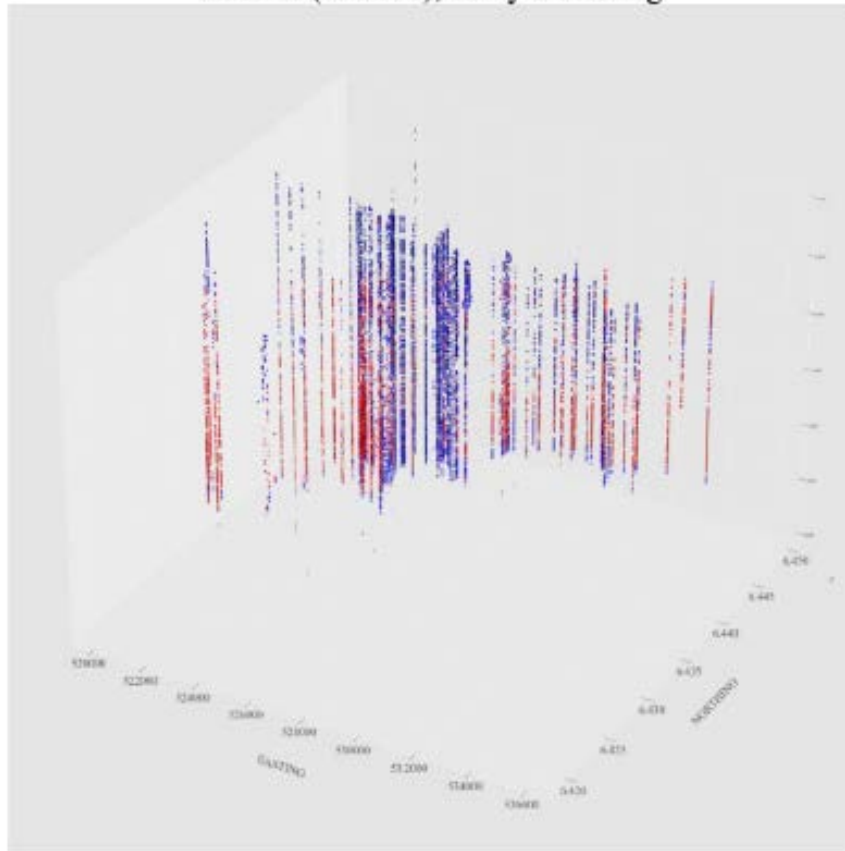


TSGWhiteMicas

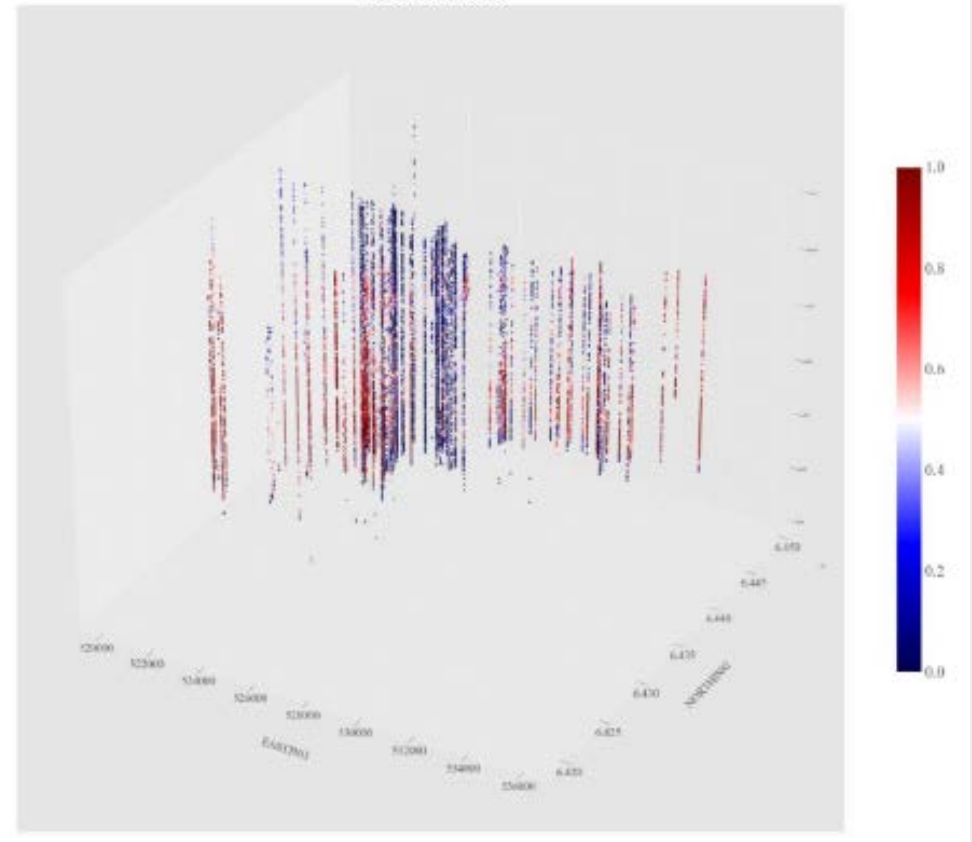


IR-spectra Fuzzy clustering

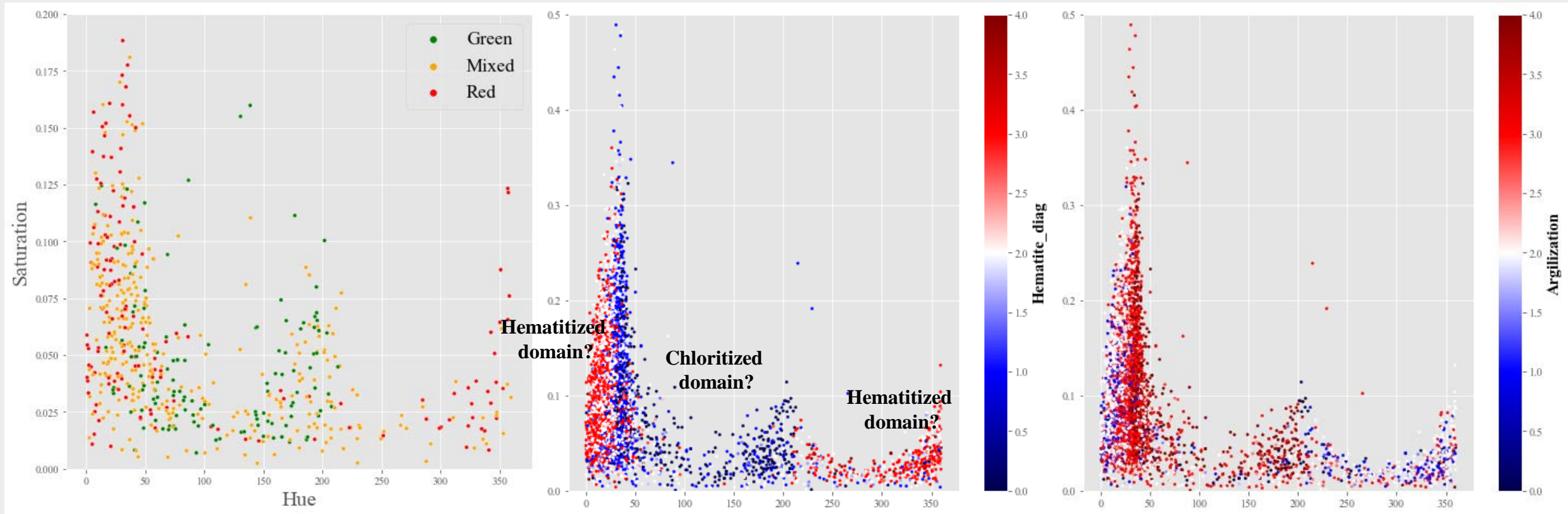
Kaolins (relative), fuzzy clustering



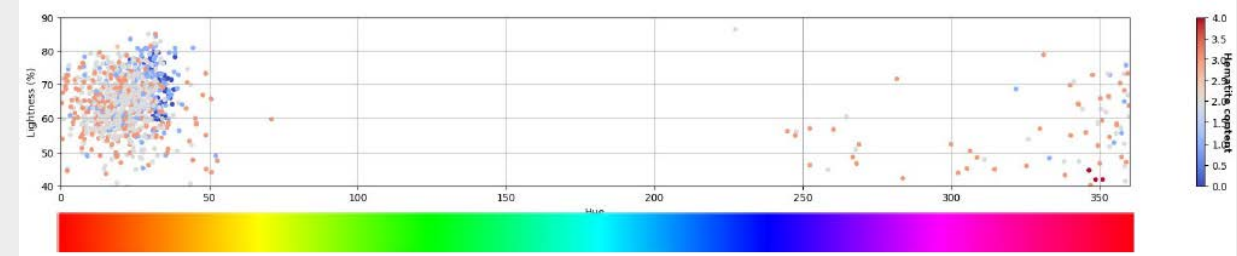
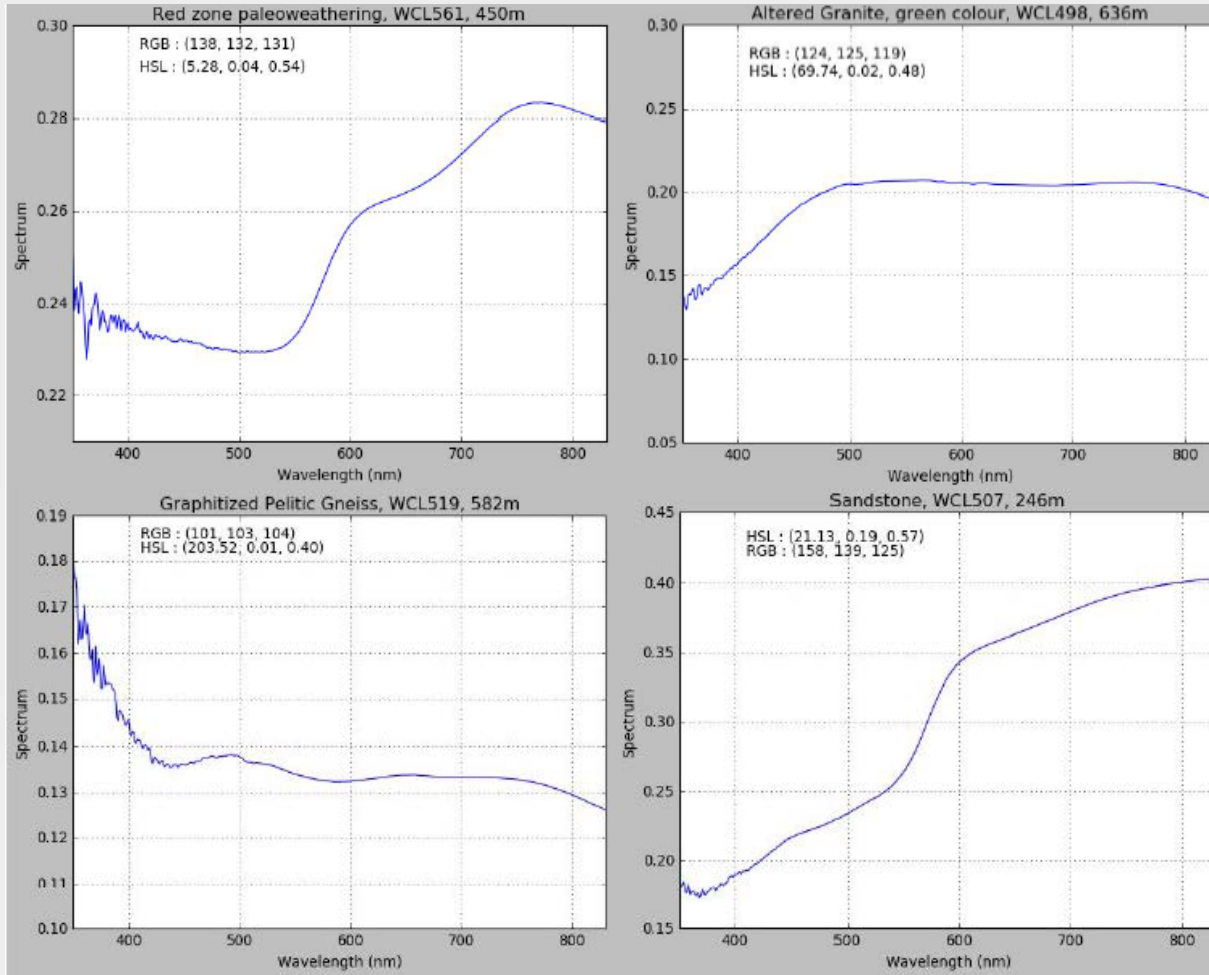
TSGDickite



Optical properties



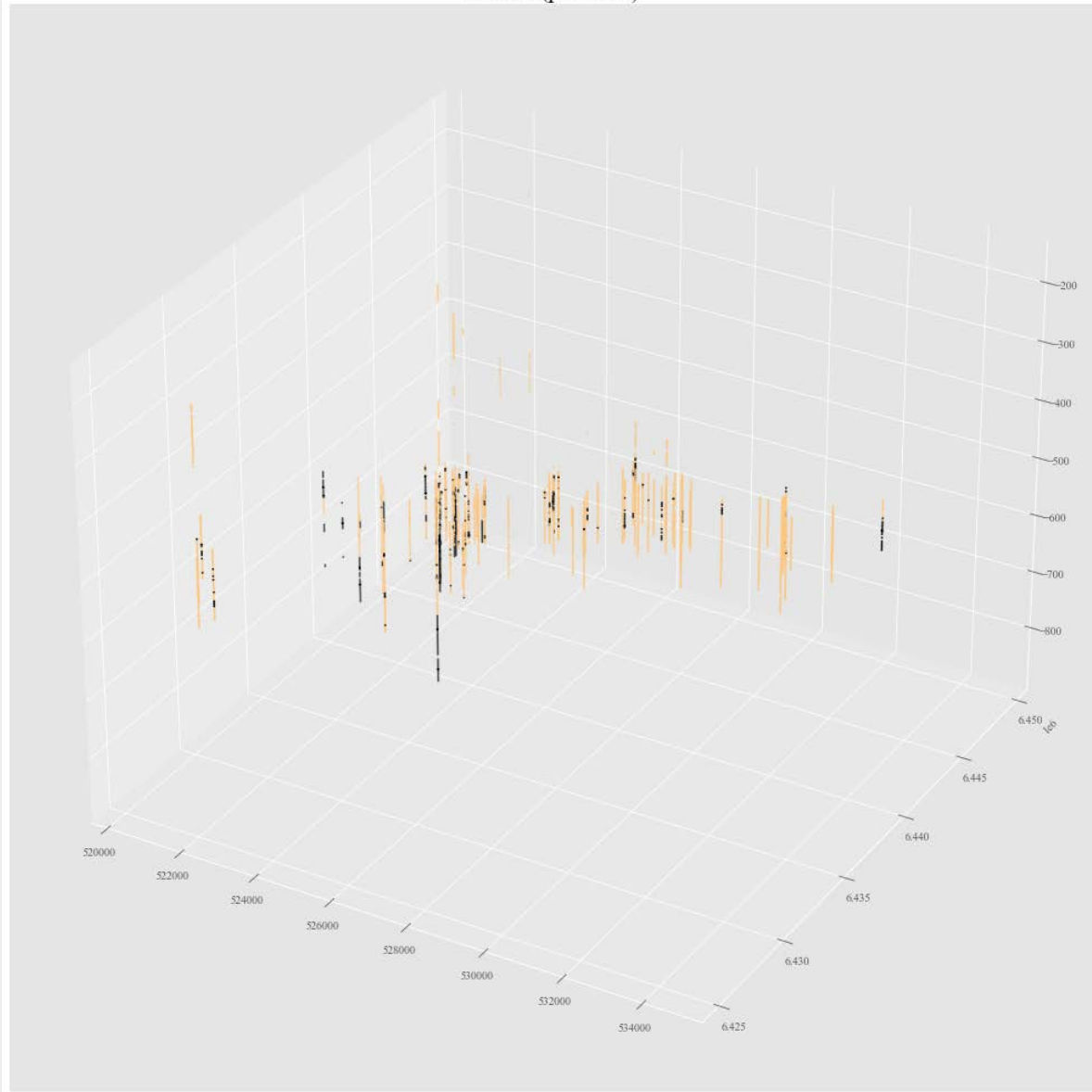
Optical properties



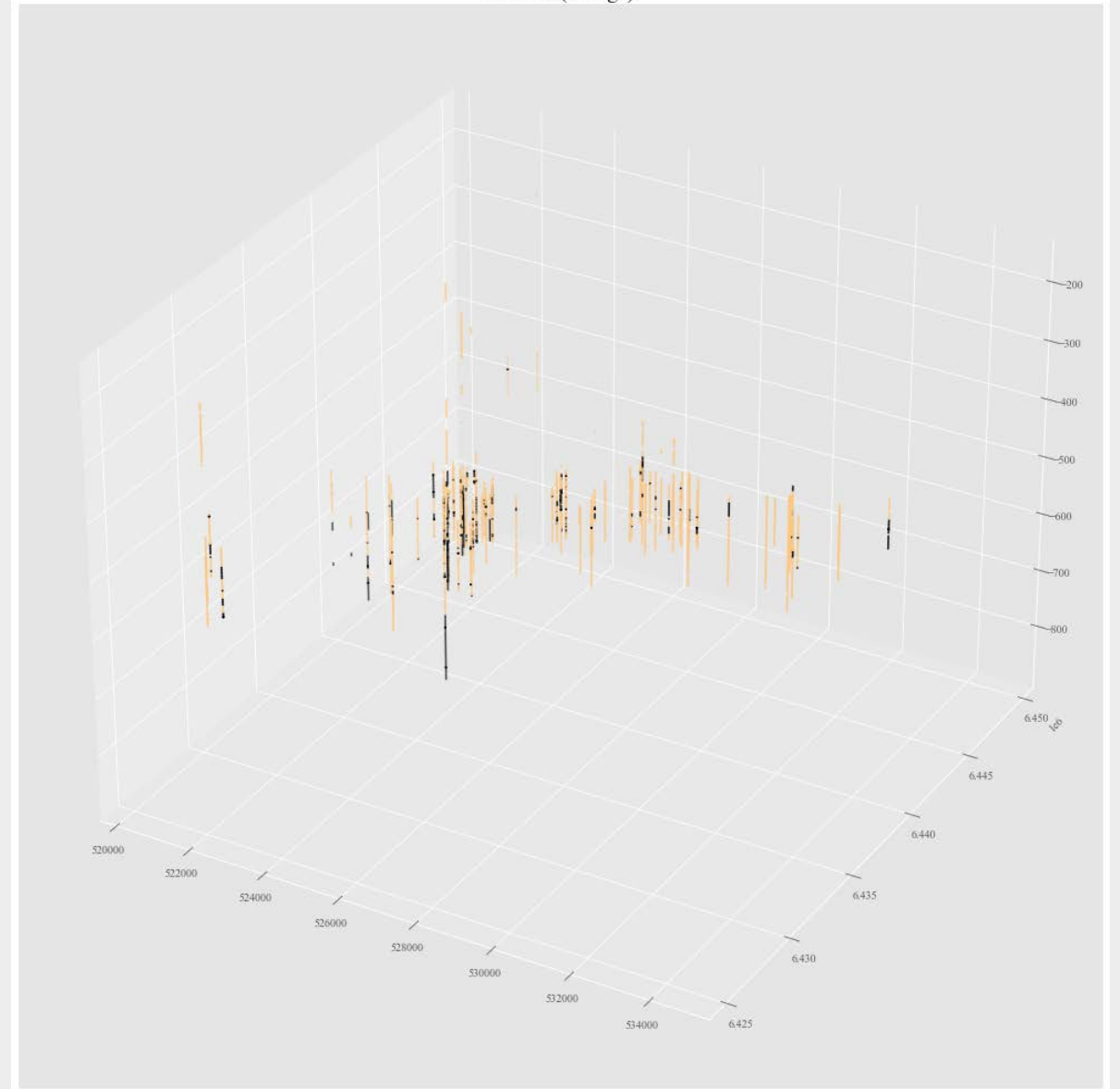
- As a whole, chloritized rocks appear greener, graphitized rocks darker, sandstone rocks lighter, paleoweathering horizon more reddish
- Is it possible to extract 'optical' classes and describe alteration facies with it ?

Lithology logs

Granite (predicted)

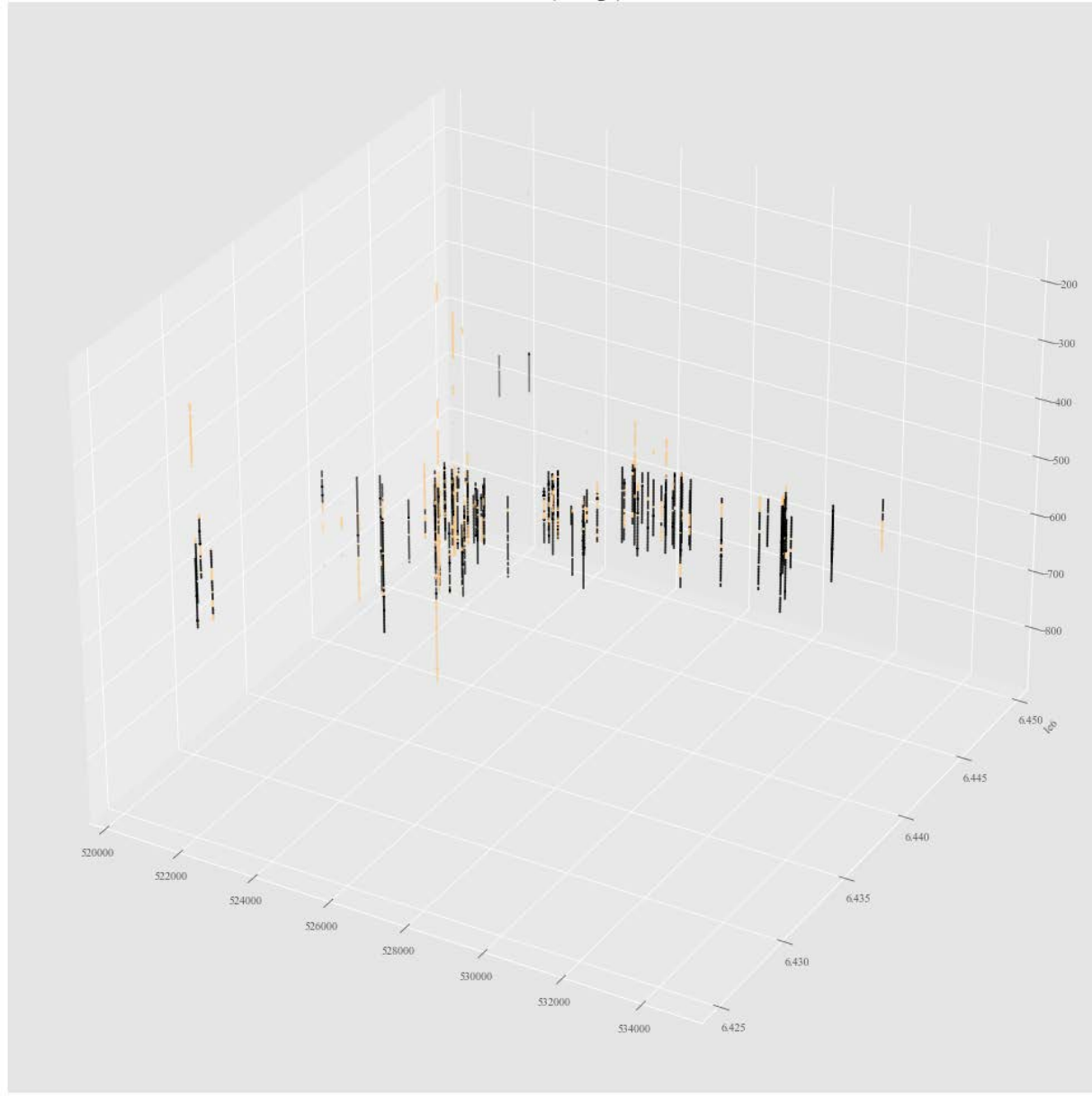


Granite (in logs)

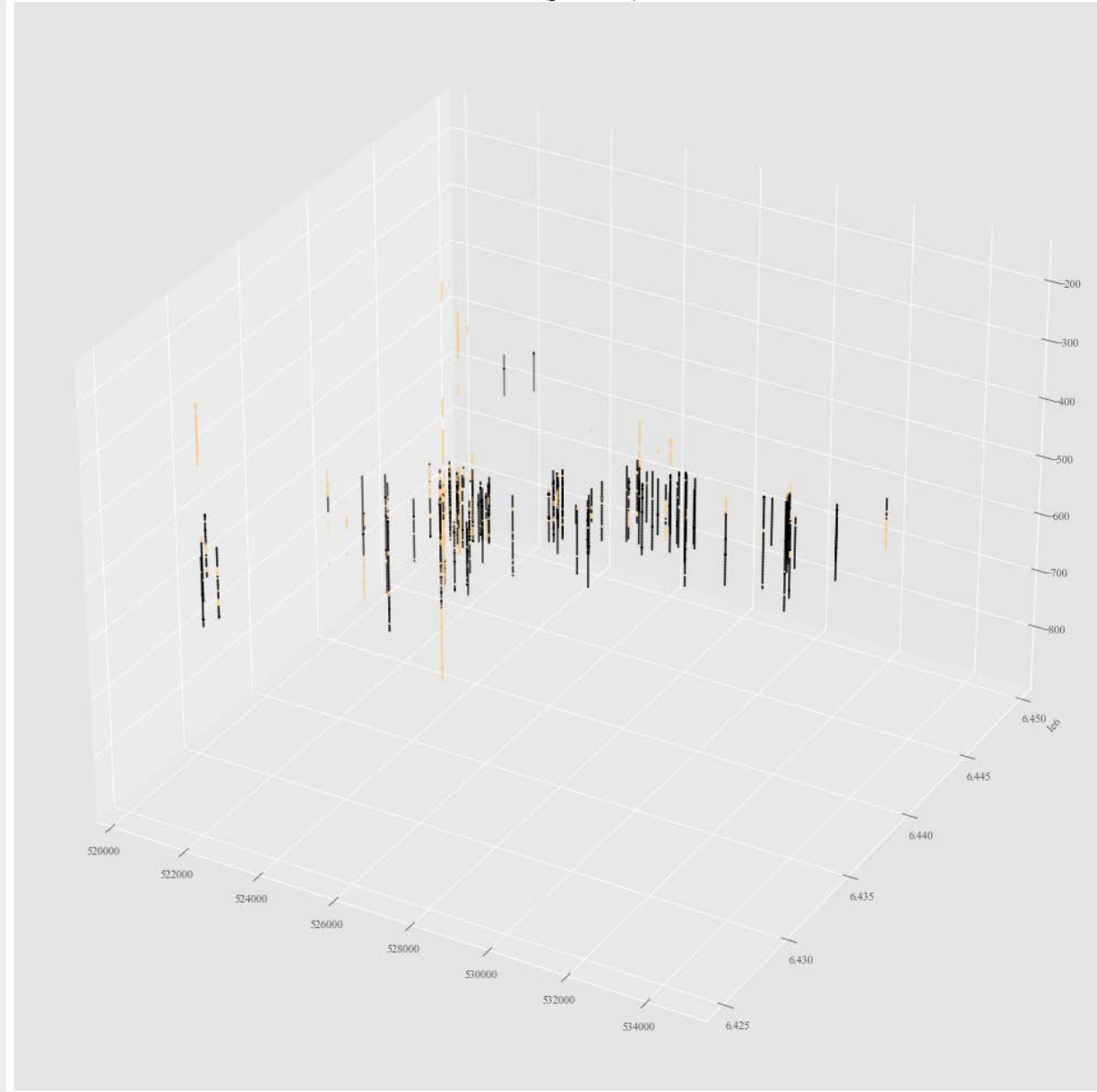


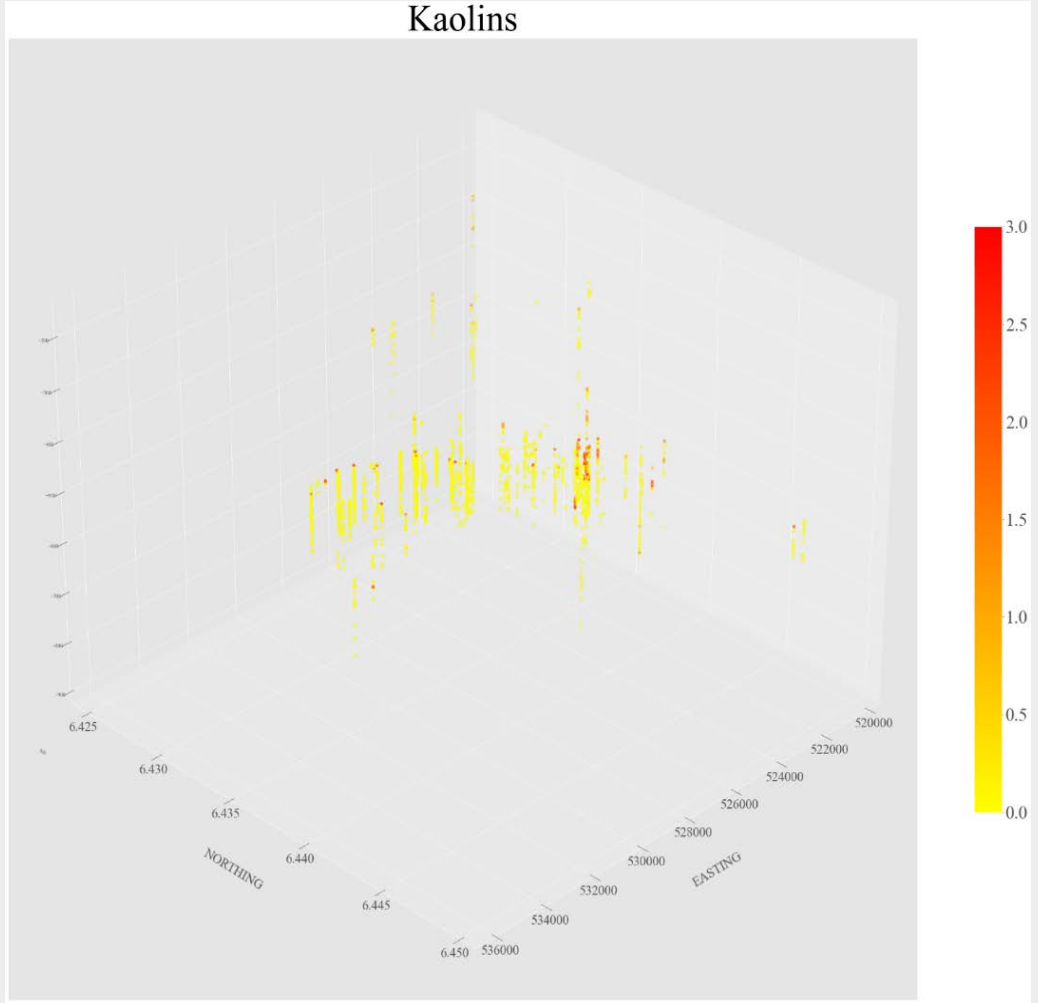
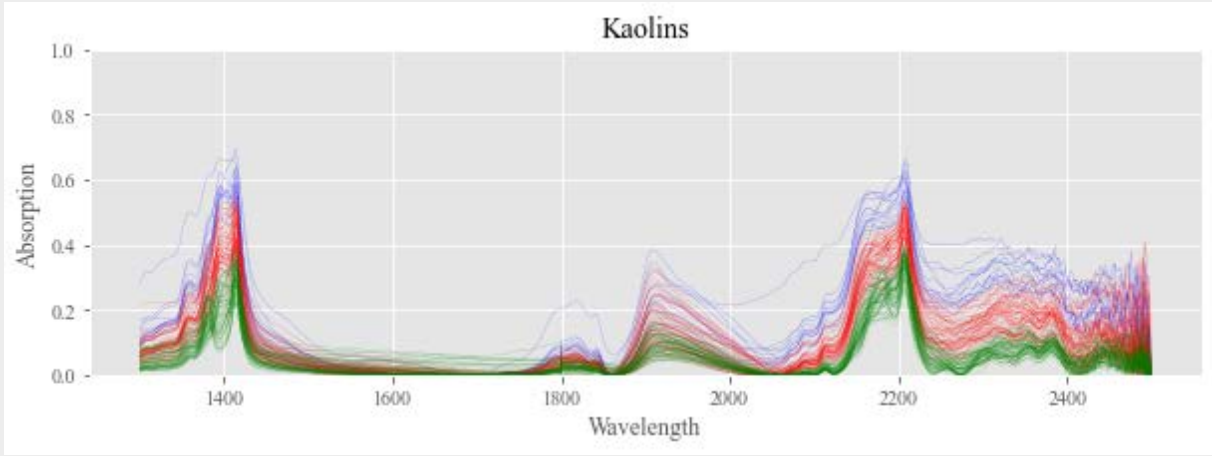
Lithology logs

Gneiss (in logs)



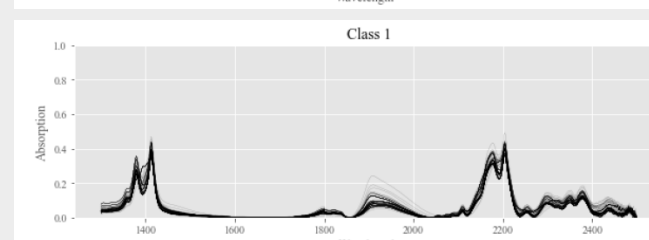
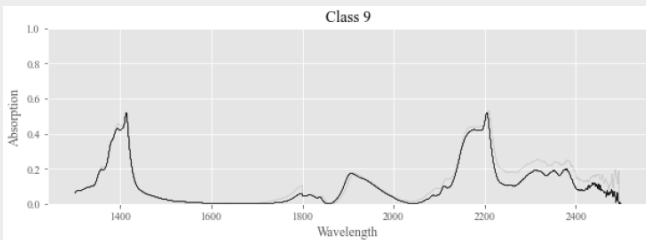
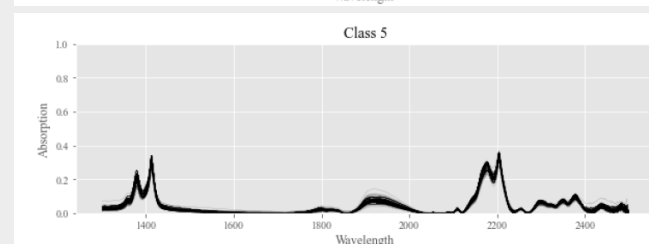
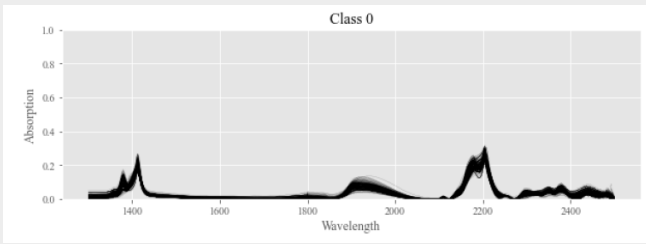
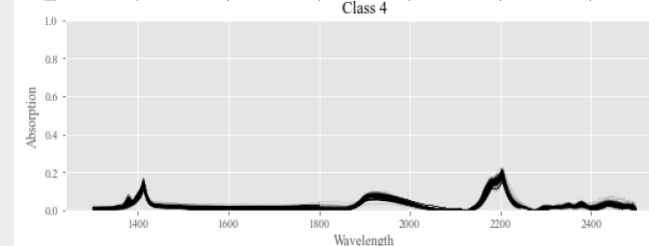
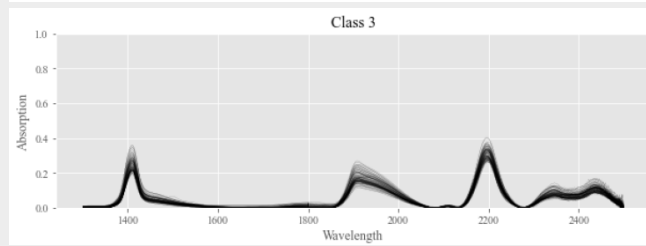
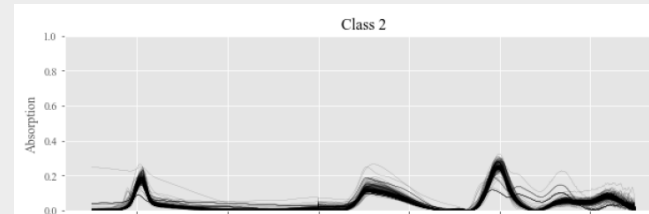
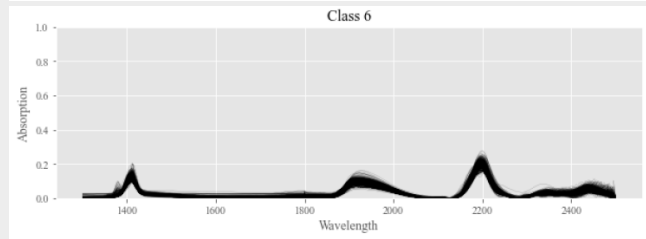
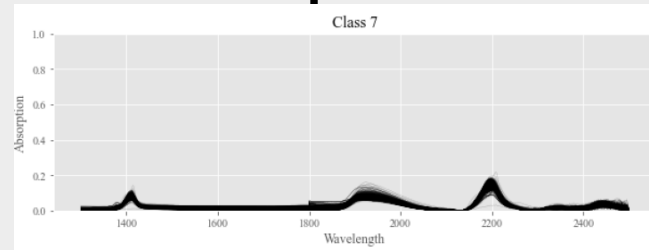
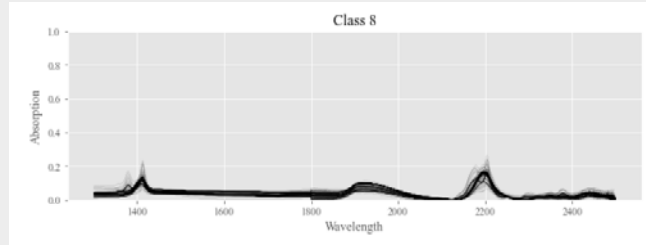
Gneiss (predicted)





IR-spectra

Unsupervised clustering



- Improving the number of clusters gradually allows for distinction between kaolins and white-micas, then chlorites
- Isolating kaolins allows for a distinction between kaolinite and dickite
- Possible to retrieve semi-quantitative scales from these classes