

Projet de Fin d'Etudes pour l'obtention du diplôme d'Ingénieur d'état en Agronomie
Option Data Science en Agriculture

Évaluation des performances de modèles d'apprentissage profond pour prédire la composition chimique de fourrage à partir de données de spectroscopie en proche infrarouge

Présenté et soutenu publiquement par :

DSSAM Abdelali

Devant le jury composé de :

Pr. HAMOUDA Allal

Pr. BENSIALI Saloua

Dr. LESNOFF Matthieu

Pr. EL AAYADI Soufiane

Président

Rapporteuse

Co-Rapporteur

Examineur

IAV HASSAN II

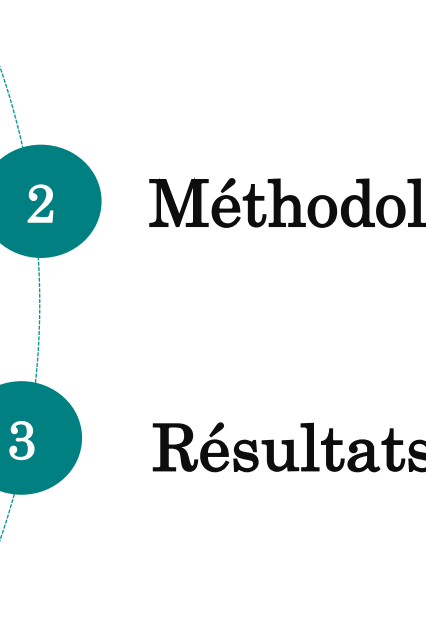
IAV HASSAN II

CIRAD

IAV HASSAN II

31 Juillet 2025

PLAN

- 
- 1 Introduction
 - 2 Méthodologie
 - 3 Résultats et discussion
 - 4 Conclusion et recommandations

1

INTRODUCTION



Qualité des Fourrages



Résilience des systèmes d'élevage



Spectroscopie en proche infrarouge (NIRS)

Non-destructive

Rapidité d'analyse

**Données spectrales en
agronomie**

Haute dimensionnalité

700 variables (longueurs d'onde)

Forte hétérogénéité

- Graminées
- Herbacées
- Légumineuses
- Arbustes

Forte colinéarité

**Approches de Modélisation
en chimiométrie**

kNN-LWPLSR

**Approches de
l'apprentissage profond**

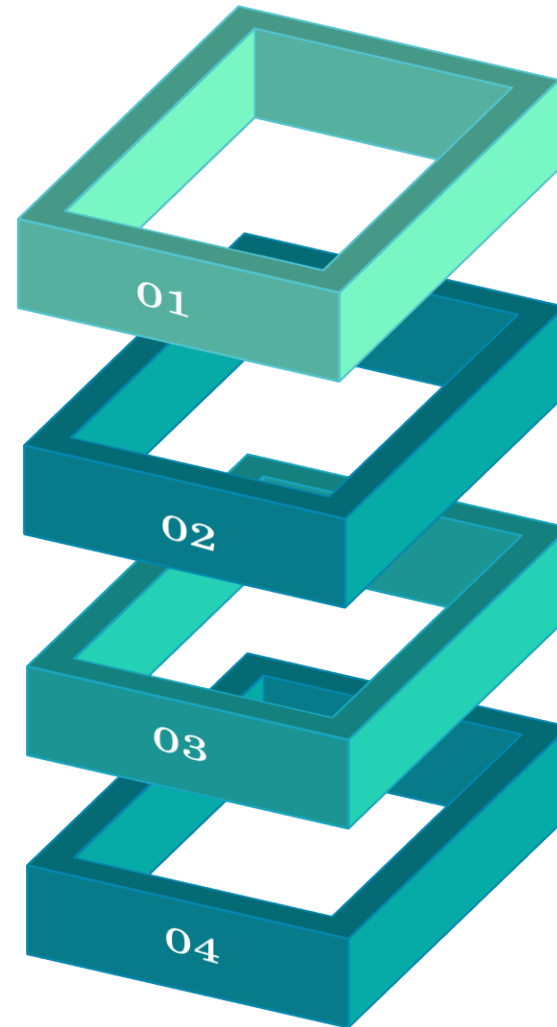
CNN, CAE ...



Sur un jeu de données réel, hétérogène et de taille limitée, les nouvelles architectures de Deep Learning peuvent-elles surpasser une méthode chimiométrique locale comme le kNN-LWPLSR ?

Objectifs spécifiques

- ☐ Développer et optimiser les modèles de référence, le PLSR et le kNN-LWPLSR
- ☐ Établir un protocole expérimental commun pour comparer les performances prédictives des différents modèles



- ☐ Mettre en œuvre trois stratégies de Deep Learning distinctes et adaptées à la problématique
- ☐ Analyser les performances à l'aide de métriques standards

2 | Méthodologie

Données

Protocole Experimental

Métriques d'évaluation

Modèles retenus

Jeu de données

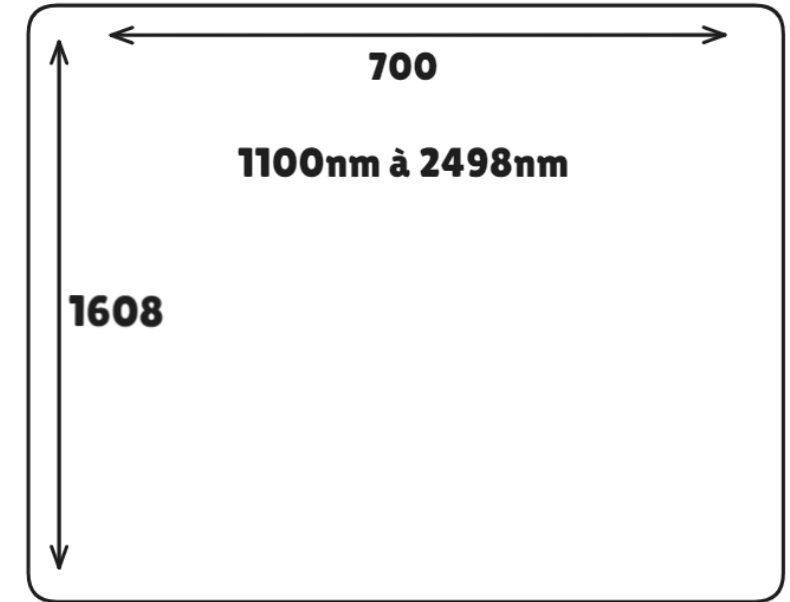
Variables explicatives

X

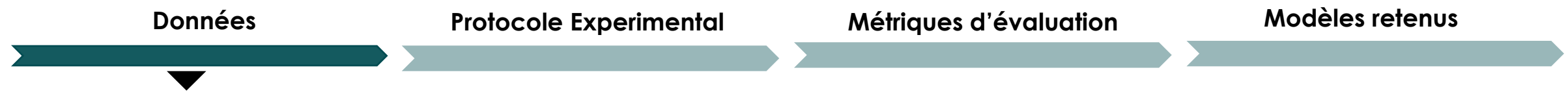
700 longueurs
d'ondes

Les absorbances
(déjà prétraités)

- Normalisation Standard Variante (SNV)
- Lissage et dérivation seconde de Savitzky-Golay.



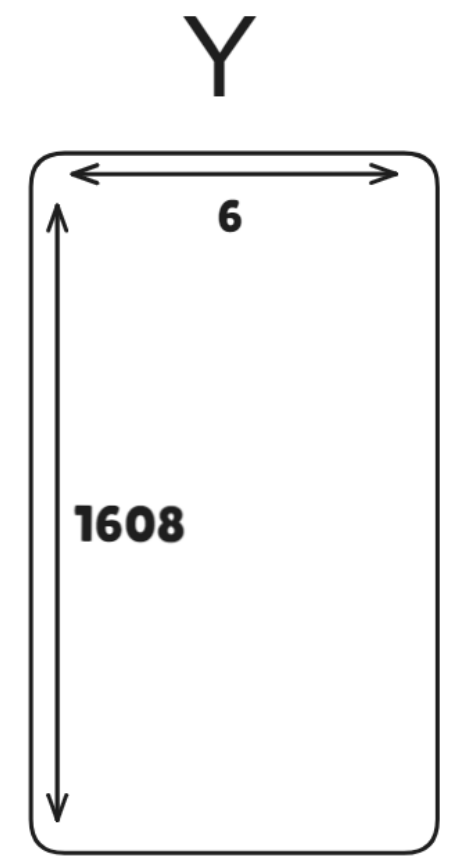
Les absorbances



Jeu de données

Variables à prédire

Variable	Unité	Nom
CP	%MS	Matière azotée totale (Nx6.25)
NDF	%MS	Neutral Detergent Fiber (Méthode Van Soest)
ADF	%MS	Acid Detergent Fiber (Méthode Van Soest)
ADL	%MS	Acid Detergent Lignin (Méthode Van Soest)
CF	%MS	Cellulose brute de Weende
DMDCell	%MS	Digestibilité enzymatique in-vitro de la matière sèche (Méthode pepsine-cellulase Aufrère)



Les Variables à prédire

Données

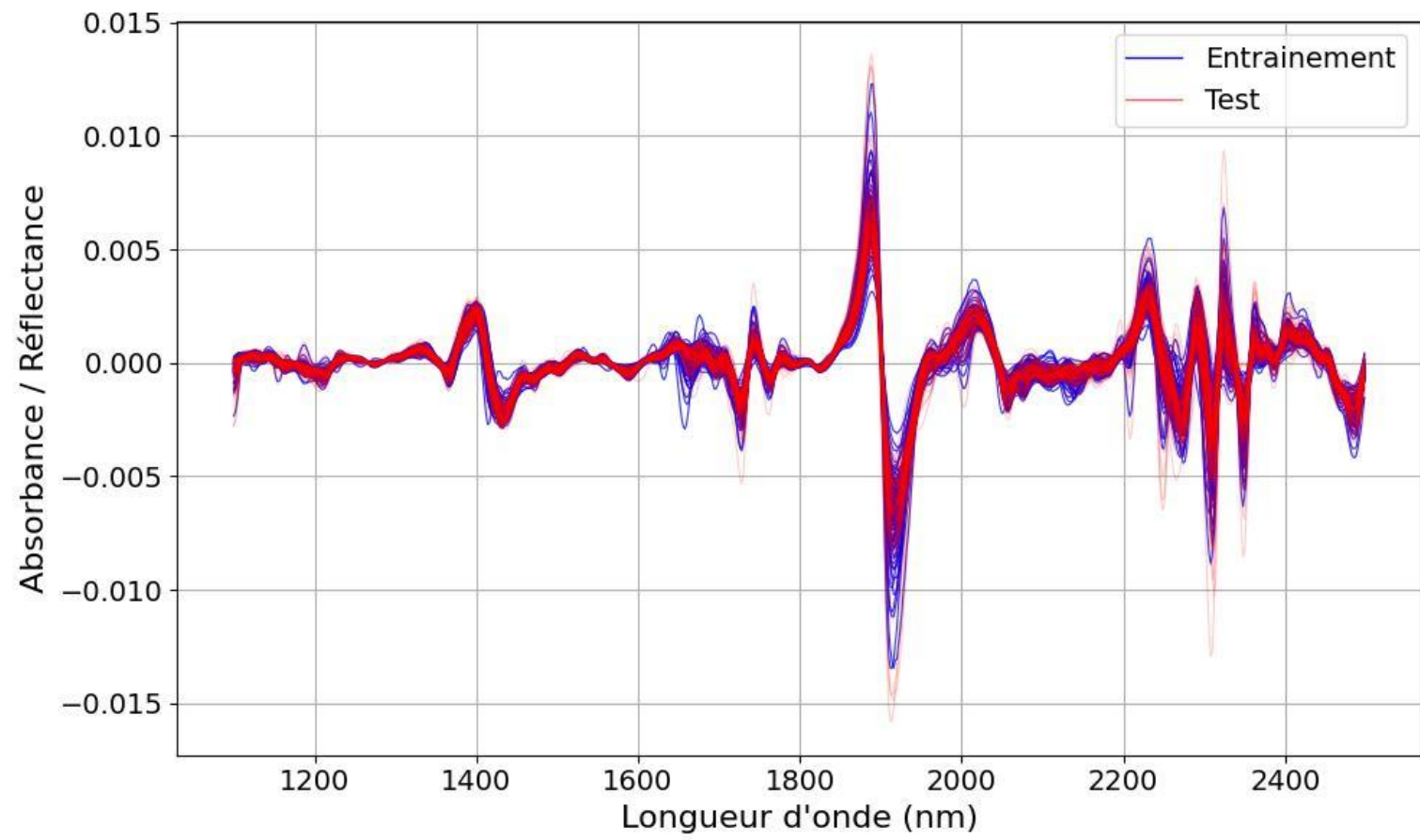
Protocole Experimental

Métriques d'évaluation

Modèles retenus

Qualité de la partition
des données

Pour les spectres X



Superposition des spectres pour un échantillon de 50 observations des ensembles d'entraînement et de test

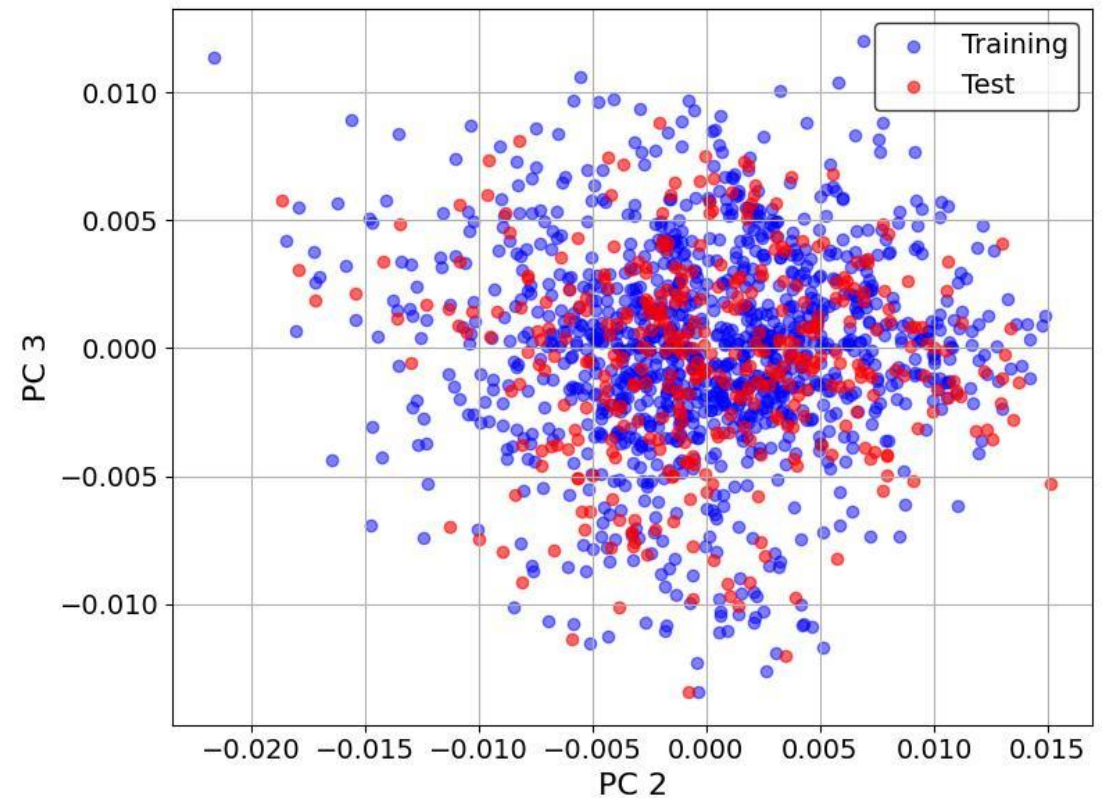
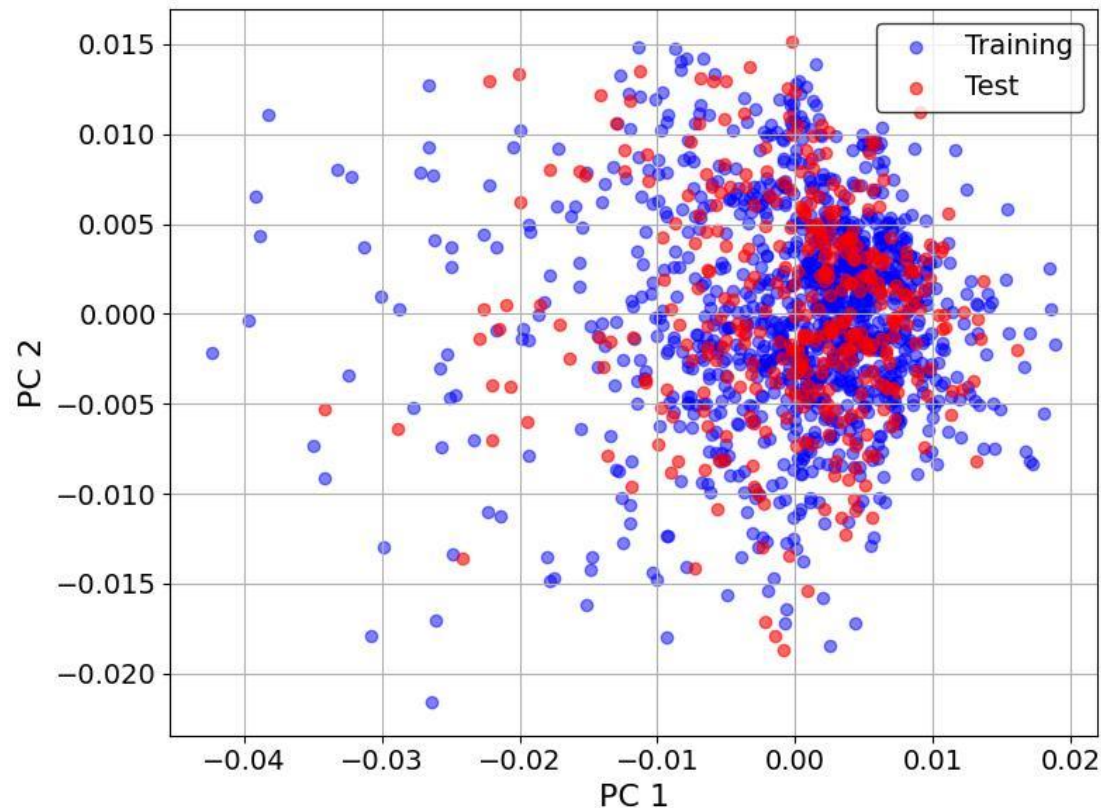
Données

Protocole Experimental

Métriques d'évaluation

Modèles retenus

Qualité de la partition des données



Projections des ensembles d'entraînement et de test sur différents plans des composantes principales pour la variable ADF

Données

Protocole Experimental

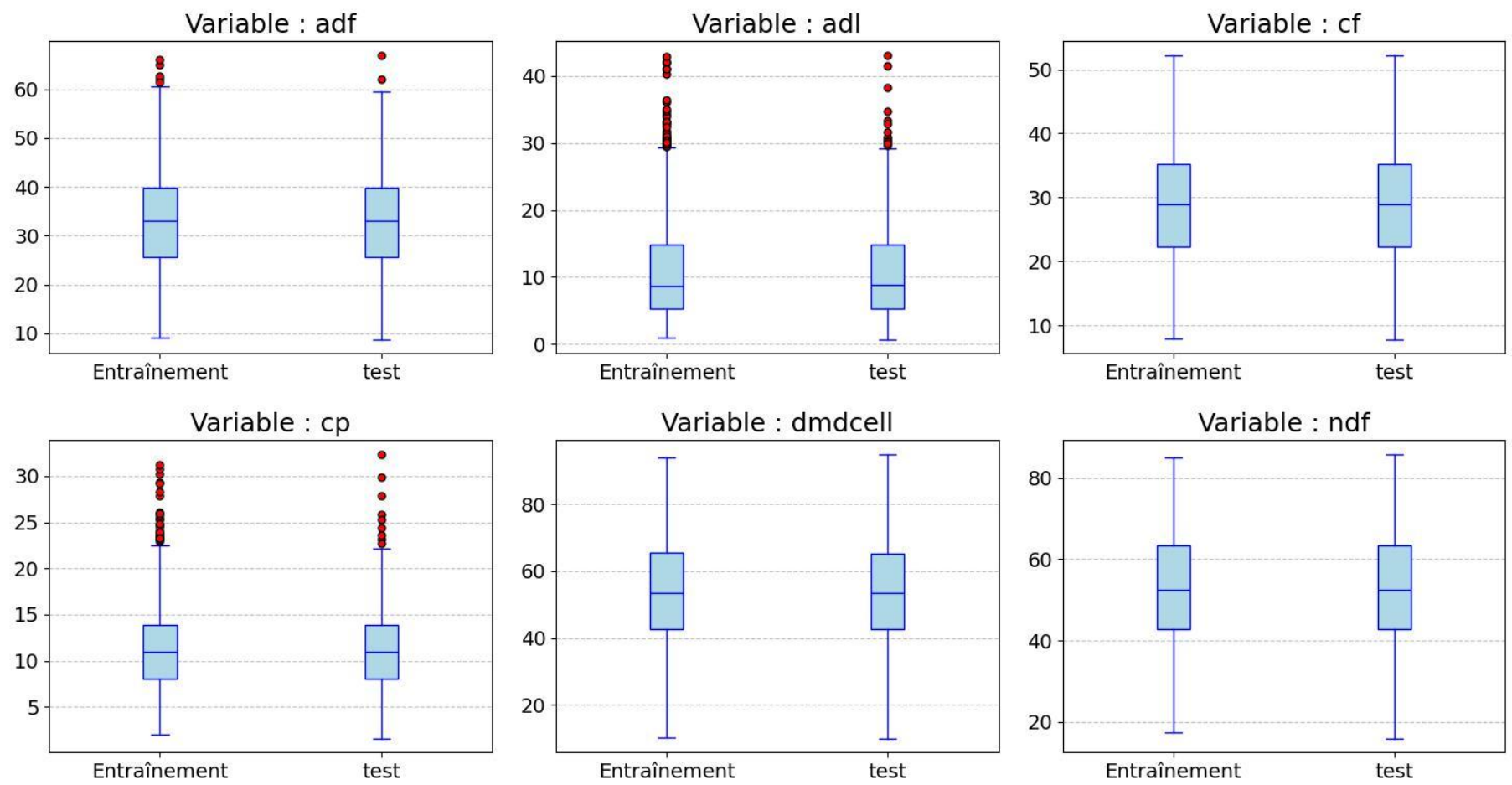
Métriques d'évaluation

Modèles retenus

Qualité de la partition
des données

Pour les variables à
prédire

Distributions statistiques
quasi-identiques entre les
ensembles d'entrainement
et de test



Boîtes à moustache des variables cibles entre les ensembles d'entraînement et de test

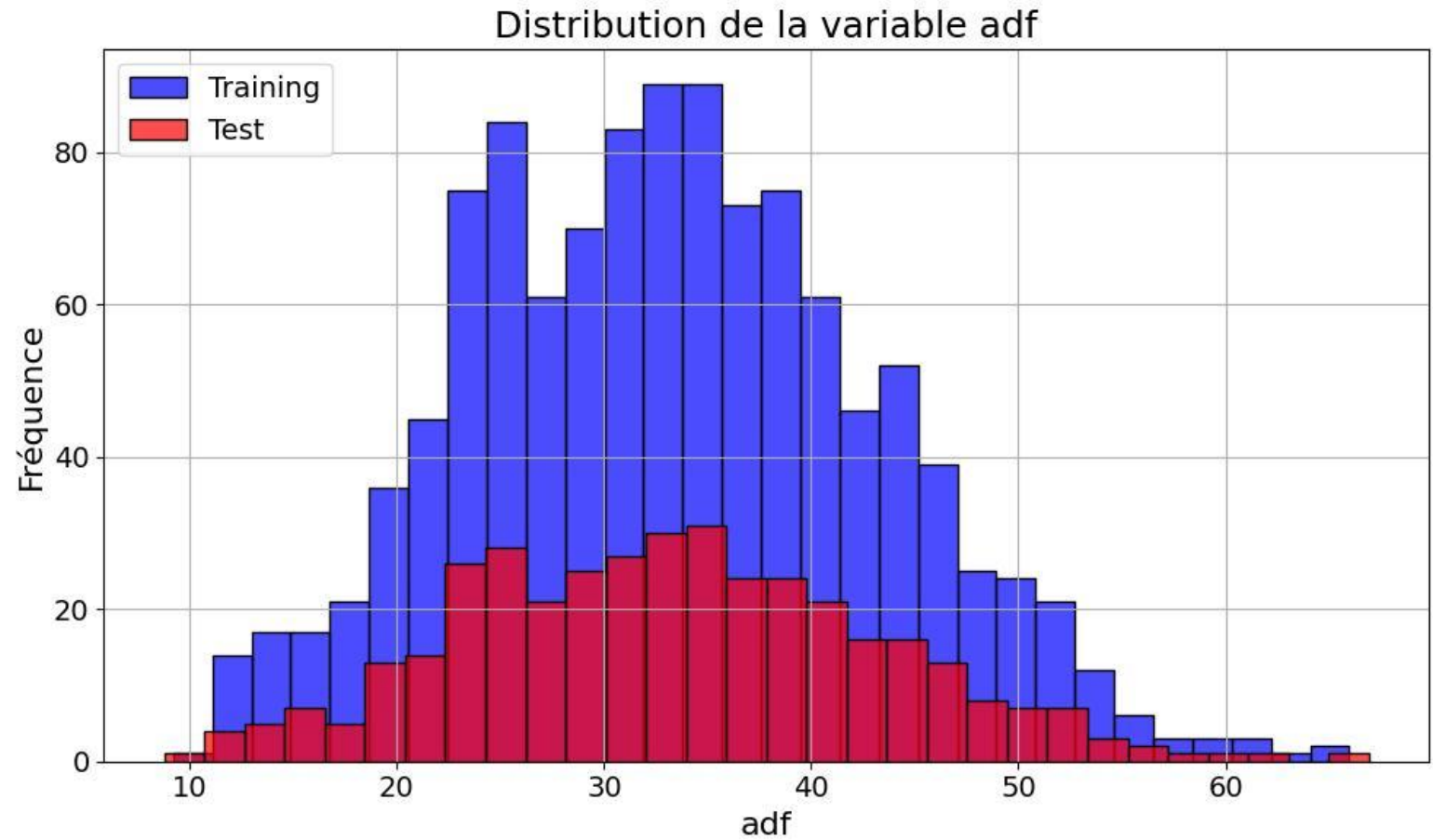
Données

Protocole Experimental

Métriques d'évaluation

Modèles retenus

Pour les variables
à prédire



Comparaison des histogrammes de fréquence entre les ensembles d'entraînement et de test pour la variable ADF



Données

Protocole Experimental

Métriques d'évaluation

Modèles retenus

Erreur quadratique moyenne de prédiction (RMSEP)

RMSEP = $\sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$

- \hat{y}_i représente la valeur prédite
- y_i représente la valeur réelle
- n est le nombre total d'échantillons dans l'ensemble de test

Une valeur faible de RMSE indique une meilleure capacité prédictive du modèle

Données

Protocole Experimental

Métriques d'évaluation

Modèles retenus

Erreur relative (RE) (RMSEP normalisée)

$$RE = \frac{RMSEP}{\mu}$$

- μ est la moyenne des valeurs observées y_i

Une valeur faible de RE indique une meilleure capacité prédictive du modèle

Données

Protocole Experimental

Métriques d'évaluation

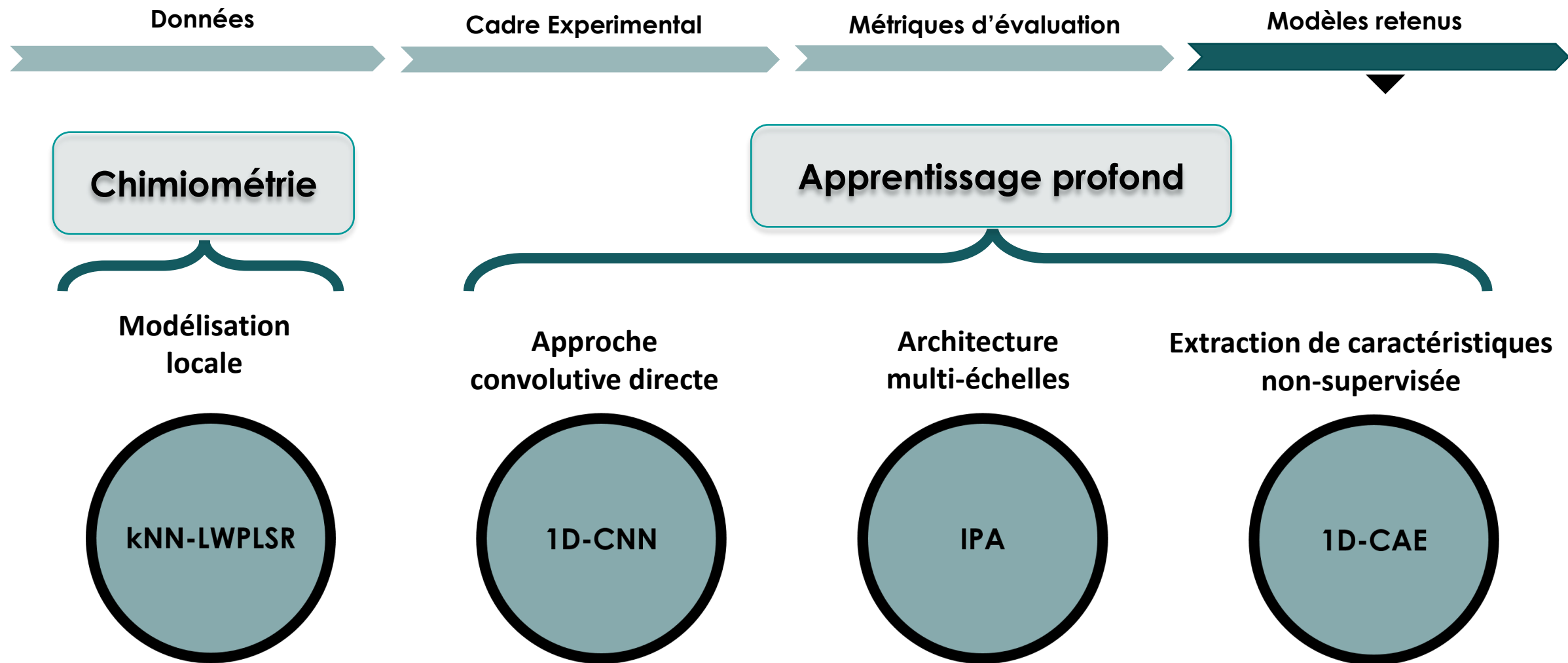
Modèles retenus

Rapport de performance de la prédiction (RPD)

$$\text{RPD} = \frac{\sigma_y}{\text{RMSEP}}$$

- σ_y est l'écart-type des valeurs observées y_i

Une valeur élevée de RPD indique une meilleure capacité prédictive du modèle



Données

Cadre Experimental

Métriques d'évaluation

Modèles retenus

kNN-LWPLSR

1D-CNN

IPA

1D-CAE



kNN-LWPLSR

Données

Cadre Experimental

Métriques d'évaluation

Modèles retenus

kNN-LWPLSR

PLSR

Crée un seul modèle
linéaire pour toutes les
données

LWPLSR

Pour chaque nouvelle
observation

Construit un modèle en
donnant plus de poids aux
échantillons similaires
(proches)

kNN-LWPLSR

Pour chaque nouvelle
observation

Sélectionne uniquement un
voisinage de k plus proches
voisins (kNN)

Construit un modèle LWPLSR
sur ce voisinage



1D-CNN

Données

Cadre Experimental

Métriques d'évaluation

Modèles retenus

1D-CNN

Extraire des caractéristiques avec des convolutions

Concaténer les sorties des différents filtres

Faire la régression

Entré

Convolution

Concaténation

Une ou plusieurs couches denses

Sortie

Architecture simplifiée du modèle 1D-CNN

Données

Cadre Experimental

Métriques d'évaluation

Modèles retenus

1D-CNN

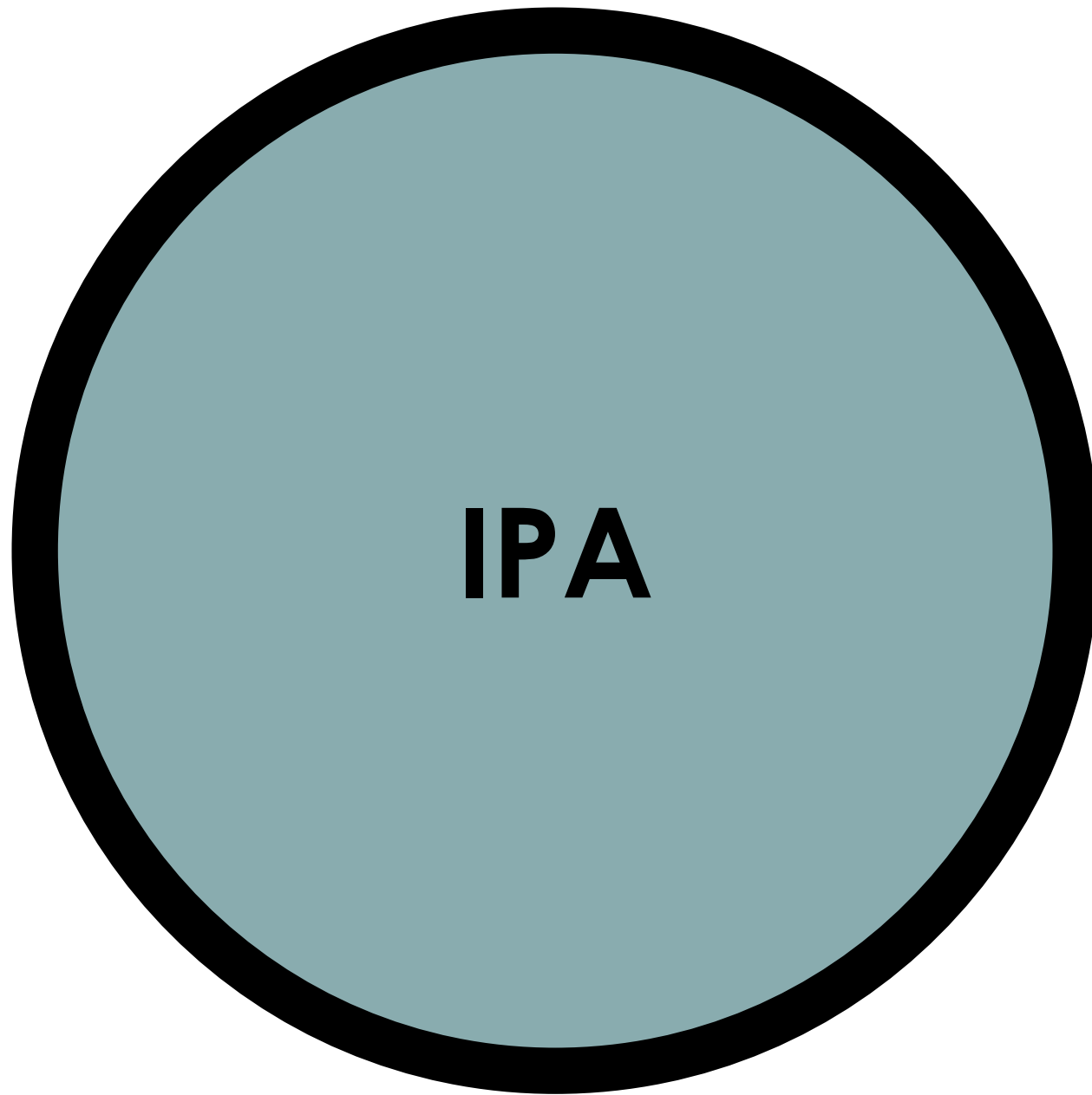
Deux variantes

1D-CNN_v1E

- Un seul grand filtre de la taille du spectre (700)
- Plusieurs couches denses pour la régression

1D-CNN_v1D

- Plusieurs filtre de petite taille
- Une seule couche dense pour la régression



Données

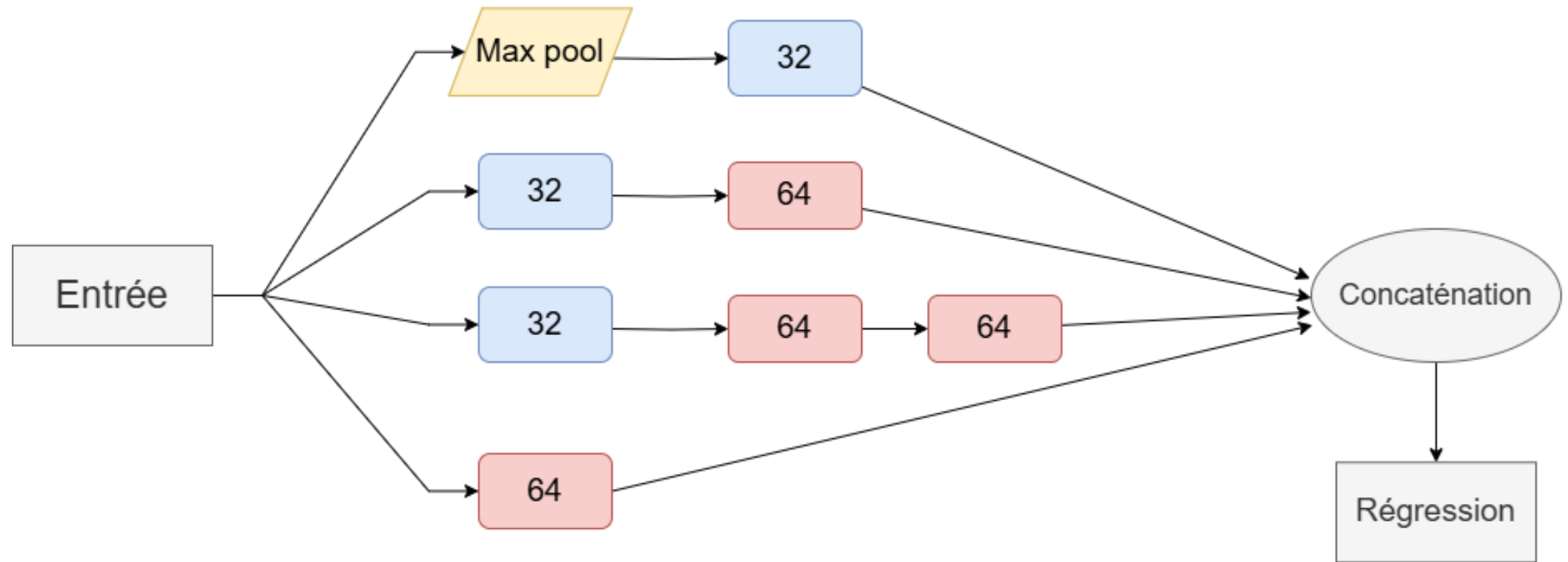
Cadre Experimental

Métriques d'évaluation

Modèles retenus

IPA

Plusieurs branches de convolution opérant en parallèle



Architecture simplifiée du modèle IPA



1D-CAE

Données

Cadre Experimental

Métriques d'évaluation

Modèles retenus

1D-CAE + MLR

Spectres d'absorbance
(X)

Entraîner un auto-
encodeur

Utiliser l'encodeur
entraîné

Prédictions

MLR

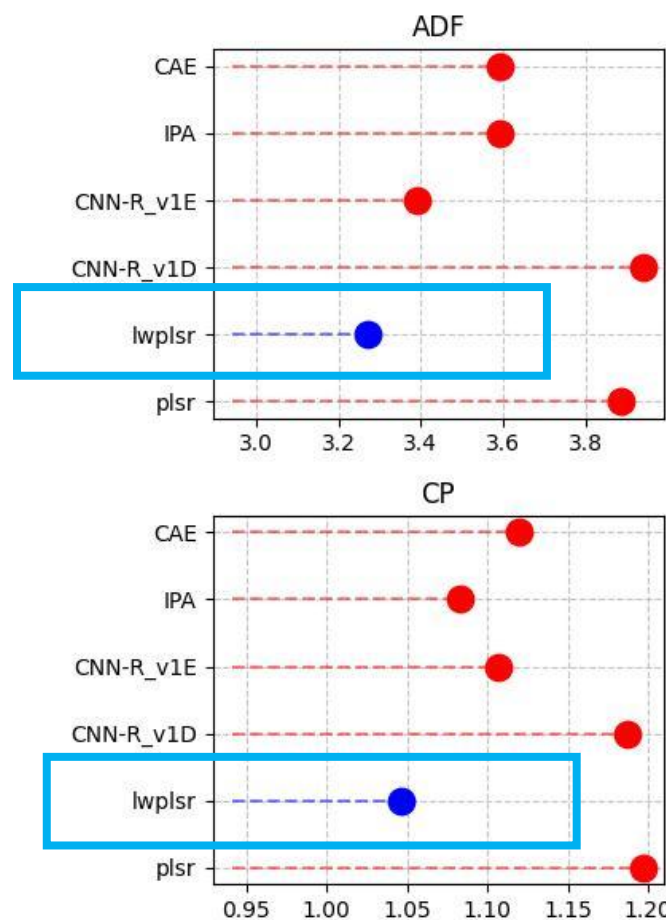
Variables
latentes

3 | Résultats et discussion

Résultats

Discussion

RMSEP

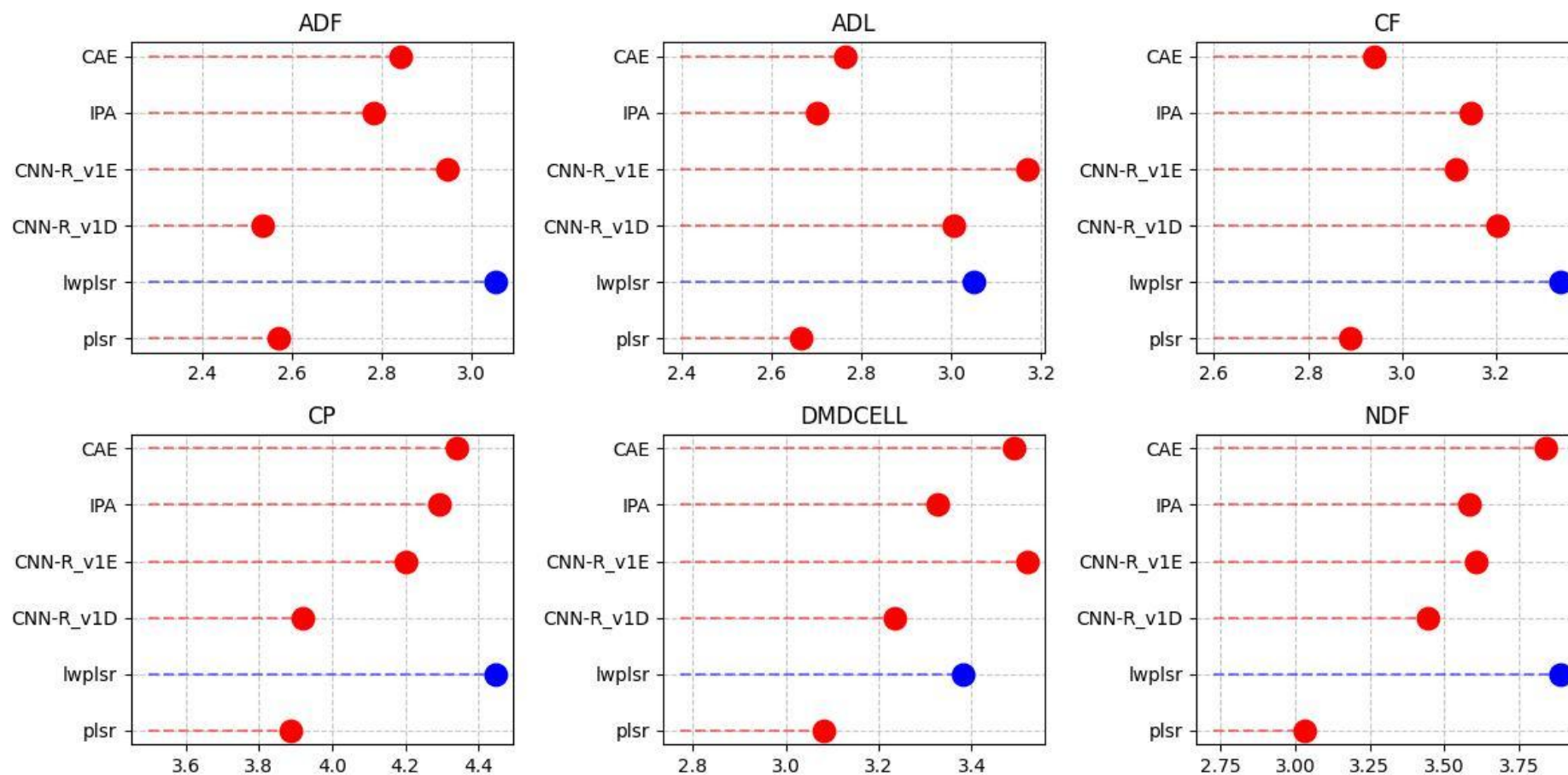


Comparaison de l'erreur quadratique moyenne de prédiction

Résultats

Discussion

RPD

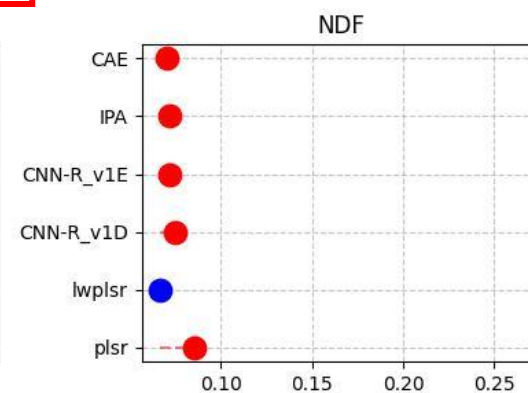
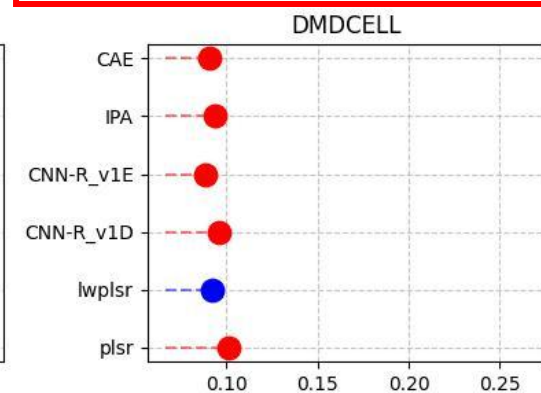
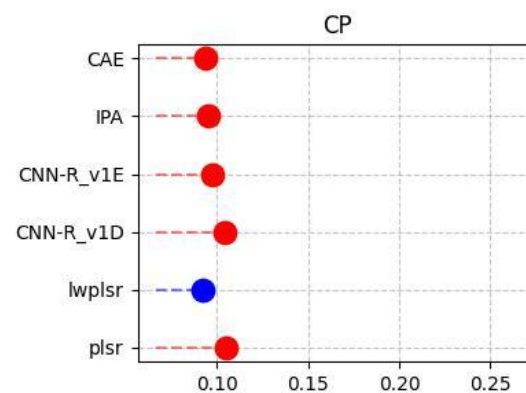
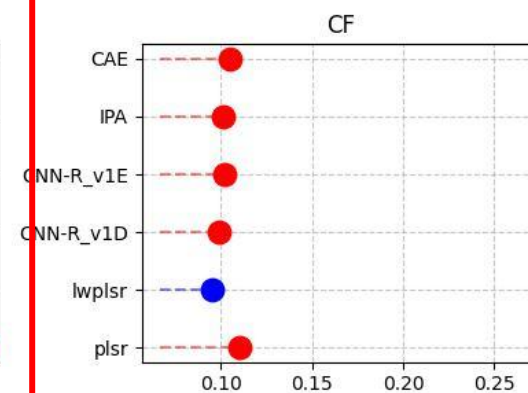
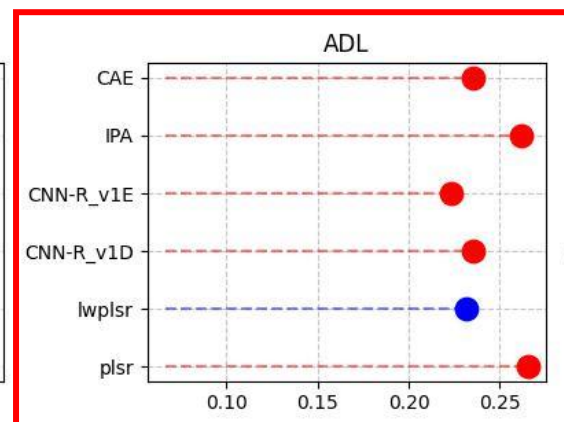
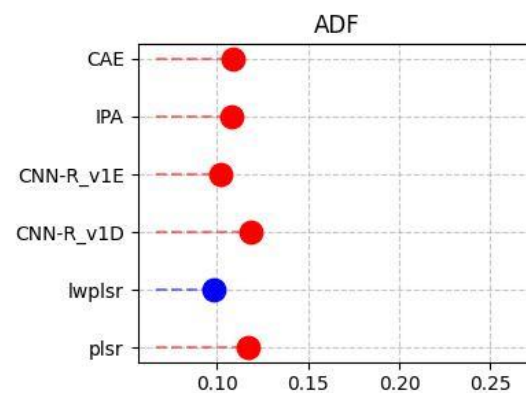


Comparaison des RPD de différents modèles

Résultats

Discussion

ER



Comparaison des performances avec l'erreur relative

Résultats

Discussion

Le Verdict : Le kNN-LWPLSR s'impose

Le kNN-LWPLSR est le plus performant pour la plupart des variables

Modélisation locale

Hétérogénéité du jeu de données

Résultats

Discussion

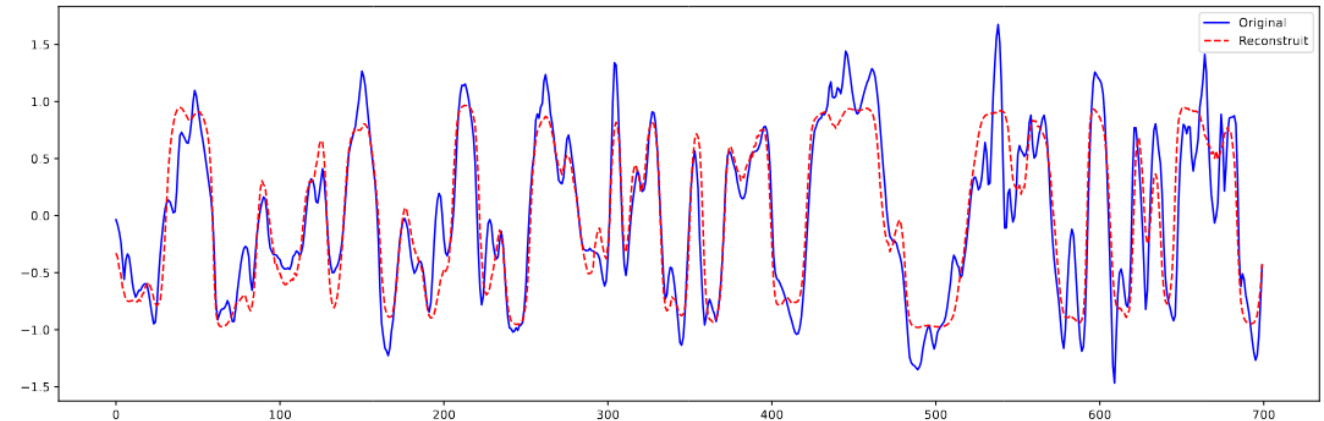
Le potentiel du Deep Learning

Le 1D-CNN_v1E est le meilleur pour l'ADL et la DMDCell

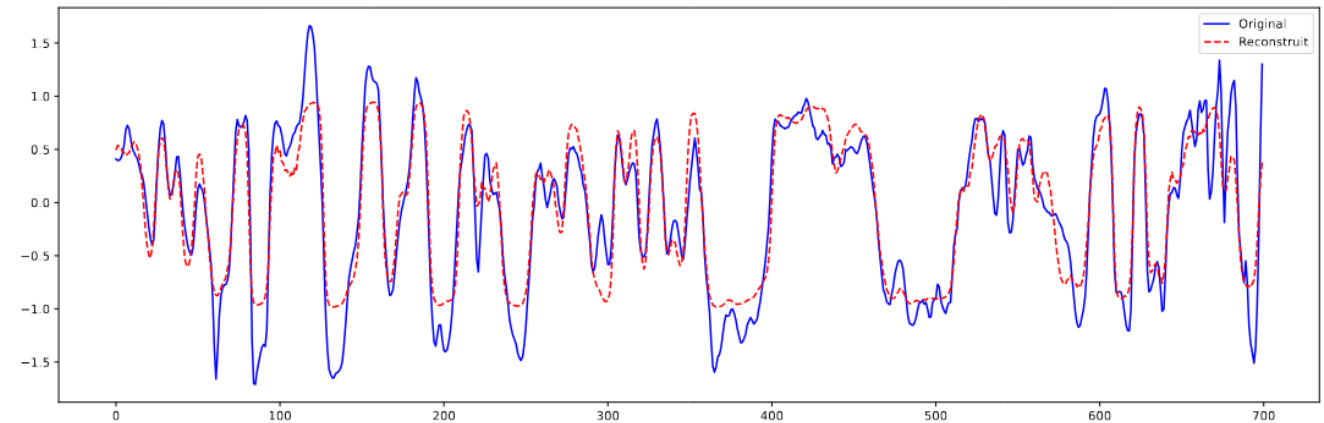
**Informations spectrales
globales**

Analyse des limites (Ex: CAE + MLR)

La régression linéaire est une
approche excessivement
simple



(a) Échantillon 220.



(b) Échantillon 249.

Exemple de deux spectres: original (en bleu) et reconstruit (en rouge) de l'ensemble du test

4 | Conclusion et Recommandations

Conclusion

Recommandations



Le kNN-LWPLSR s'avère le modèle le plus performant pour la majorité des variables



Le modèle 1D-CNN_v1E s'adapte mieux pour deux variables



Tous les modèles non-linéaires ont surpassé la référence linéaire PLSR

Conclusion

Recommandations

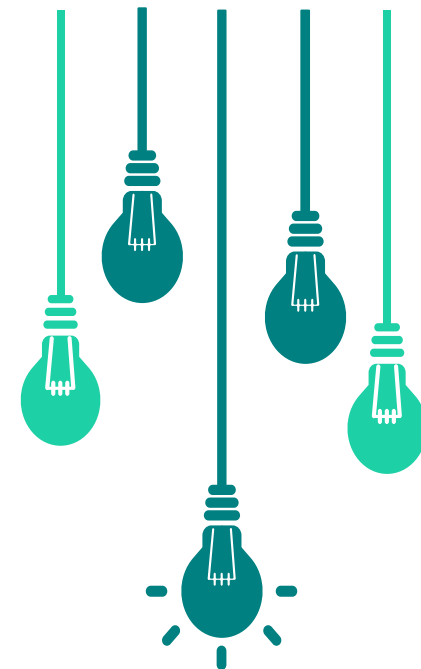
Pistes d'amélioration

Modèles de régression multivariée

Techniques d'augmentation de données

Architectures de bout en bout

Agrégation des modèles (*ensembling*)



Merci de votre aimable attention

Projet de Fin d'Etudes pour l'obtention du diplôme d'Ingénieur d'état en Agronomie
Option Data Science en Agriculture

Évaluation des performances de modèles d'apprentissage profond pour prédire la composition chimique de fourrage à partir de données de spectroscopie en proche infrarouge

Présenté et soutenu publiquement par :

DSSAM Abdelali

Devant le jury composé de :

Pr. HAMOUDA Allal

Pr. BENSIALI Saloua

Dr. LESNOFF Matthieu

Pr. EL AAYADI Soufiane

Président

Rapporteuse

Co-Rapporteur

Examineur

IAV HASSAN II

IAV HASSAN II

CIRAD

IAV HASSAN II

31 Juillet 2025