

Méta-modélisation de spectres
par réseaux de neurones hybrides
pour accélérer les simulations.

CEA-DAM, DIF
Journée IA

Voir [Kluth et al. «Deep learning for
NLTE spectral opacities » PoP 2020].

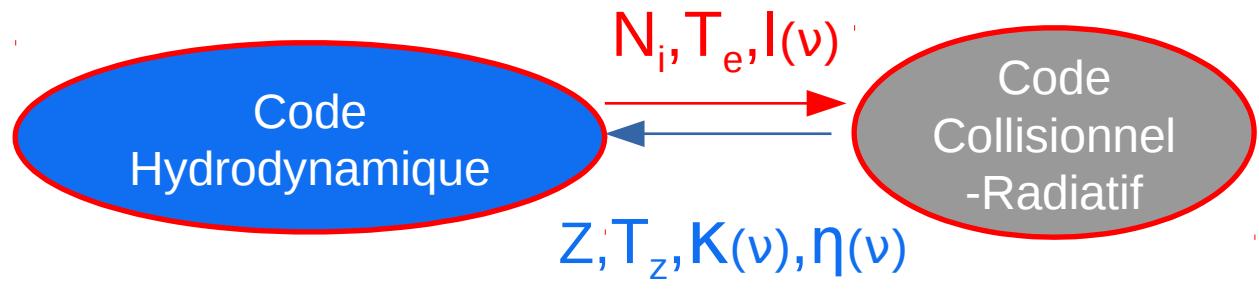
KLUTH Gilles

*GILLERON Franck, LEPAREUR Loic,
MORICE Olivier, PIRON Robin*



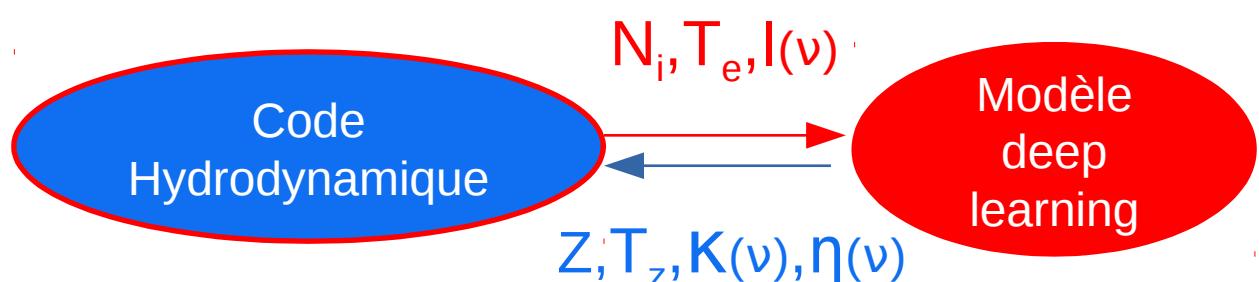
CEA-DAM/DOE-NNSA collaboration

Hydro/HETL pour les simulations de cavité laser



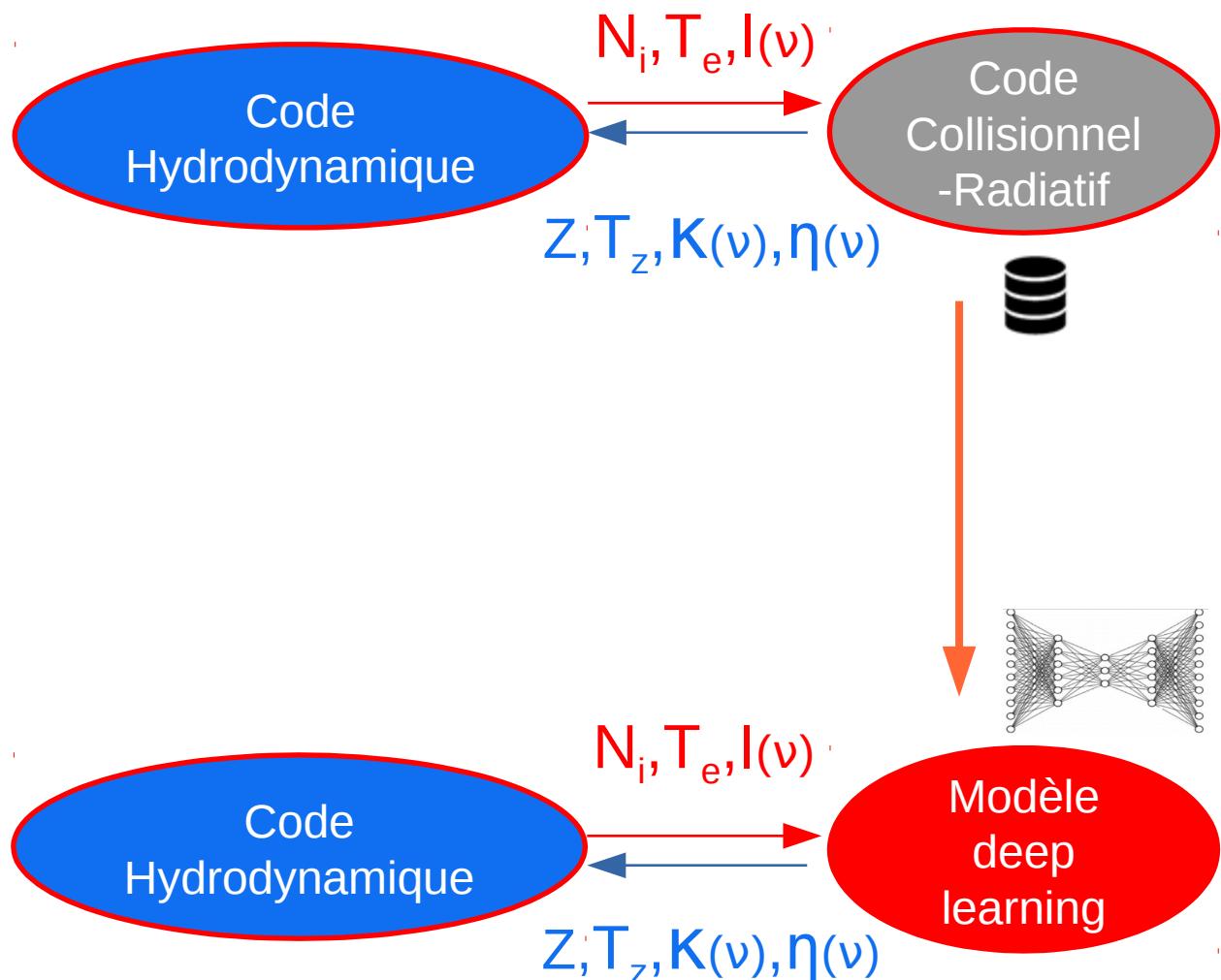
Pour les simulations de cavité laser,
les calculs HETL constituent **une part majeure du coût**
(Cela évolue...)

Même avec un code collisionnel-radiatif rapide (~50 ms),
le **nombre d'appels** (par maille, par pas de temps) est
très grand.



Objectif :
Des simulations rapides avec la meilleure physique
possible, et pour lesquelles on peut garantir le résultat.

Deep learning intégrée dans la simulation



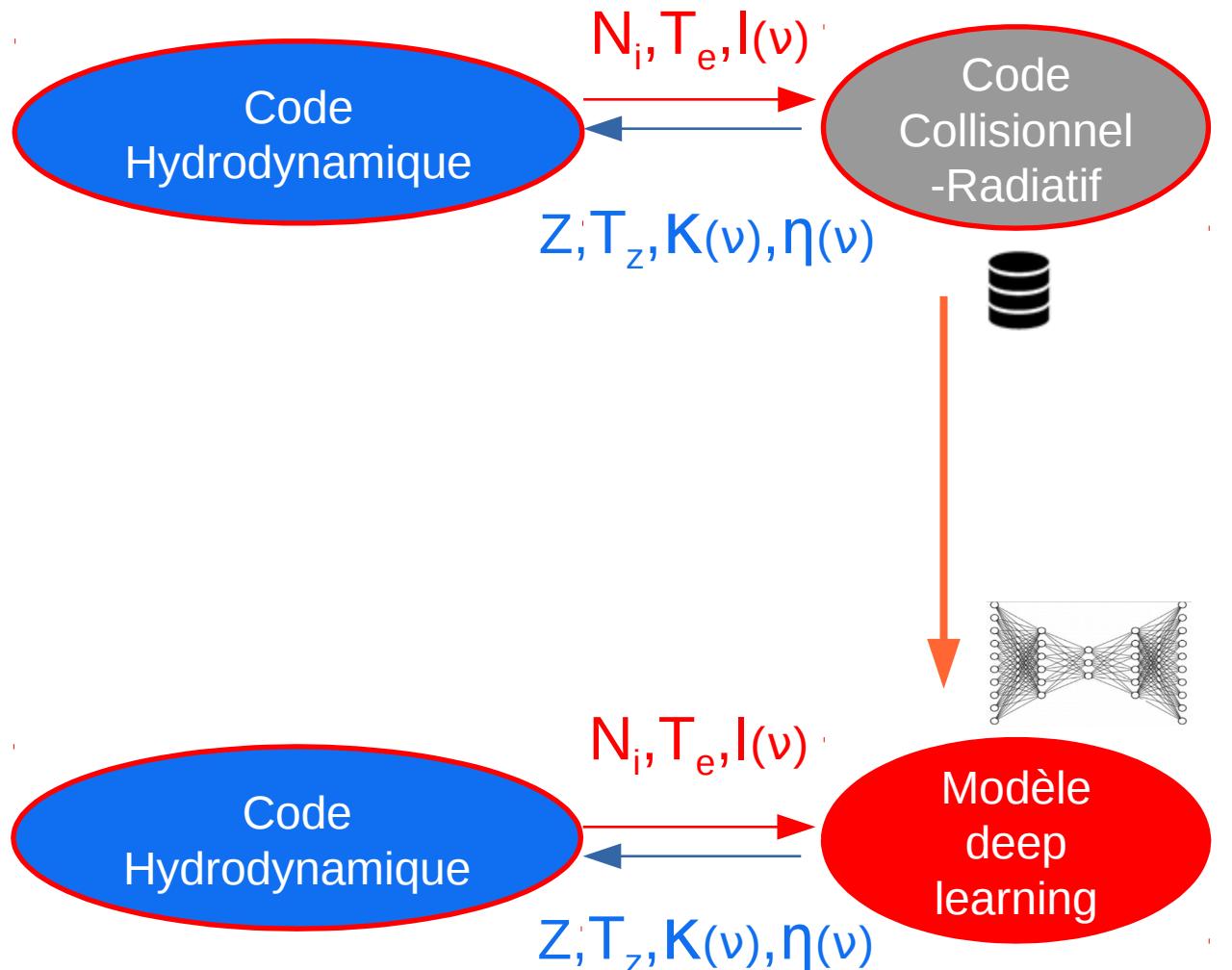
Pour les simulations de cavité laser,
les calculs HETL constituent **une part majeure du coût**
global.
(Cela évolue...)

Même avec un code collisionnel-radiatif rapide (~50 ms),
le **nombre d'appels** (par maille, par pas de temps) est
très grand.

- 1/ Production de données avec le code CR.
- 2/ Entrainement du modèle de DL.
- 3/ Simulation de cavité laser avec la prédition du modèle DL à la place du code CR.

Objectif :
Des simulations rapides avec la meilleure physique
possible, et pour lesquelles on peut garantir le résultat.

Opportunités pour les simulations multi-échelles



En cours
1

Sur le CR actuel.
Simulations accélérées.

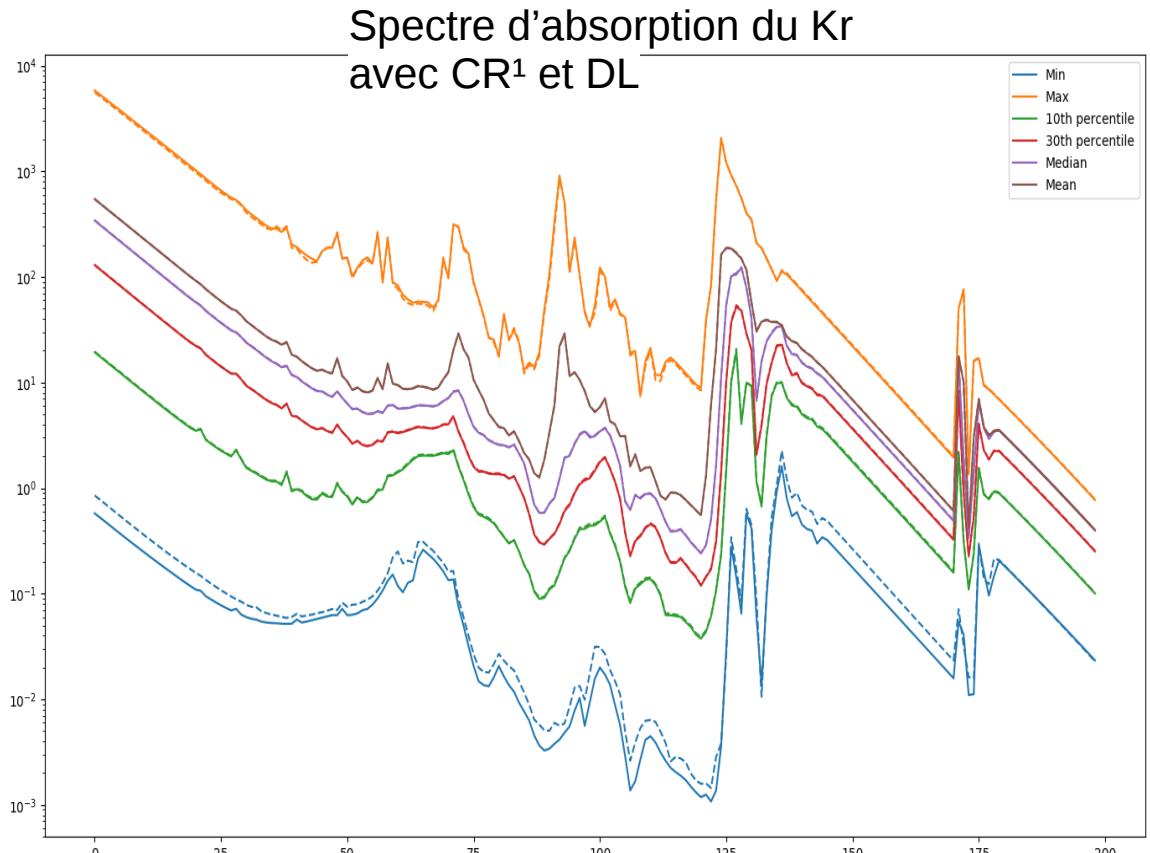
Phase
2

Sur des futurs CR.
Transfert d'apprentissage.
Simulations avec une physique inatteignable par les méthodes classiques.

Phase
3

Sur d'autres physiques.
Simulations multi-échelles.
Flux non-local
[Lamy, Dubroca, Nicolaï. PRE 2022]

Pourquoi utiliser le deep learning ?



Grande dimension : les spectres sont discrétisés sur ~100 groupes.

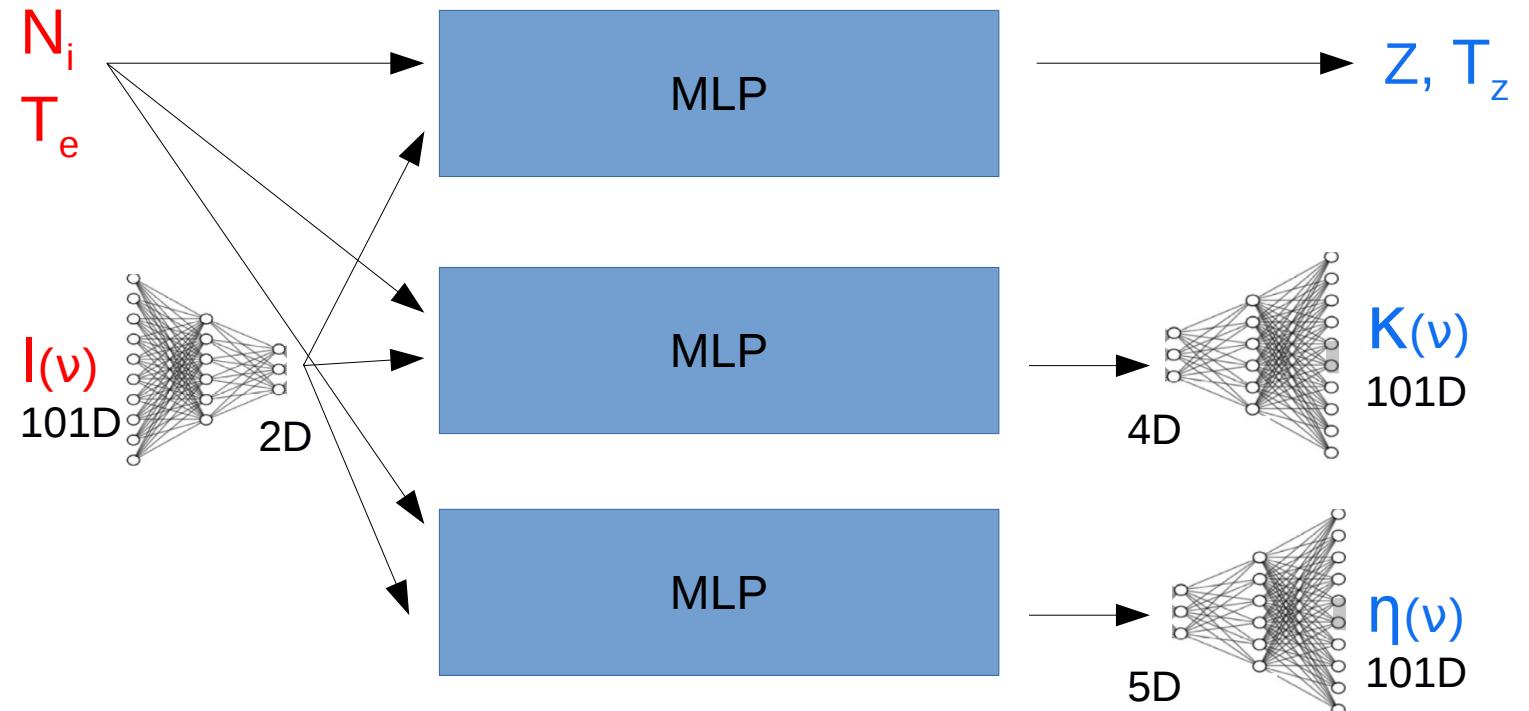
Données massives : nous pouvons produire des millions de données avec le code CR.

Grande précision exigée : à cause de la propagation d'erreur dans les codes hydrodynamiques.

Sur des fonctions “irrégulières” : voir les seuils de photo-ionisation.

¹ [H. Scott « Cretin – a radiative transfer capability for laboratory plasmas ». JQSRT 2001]

Architecture choisie



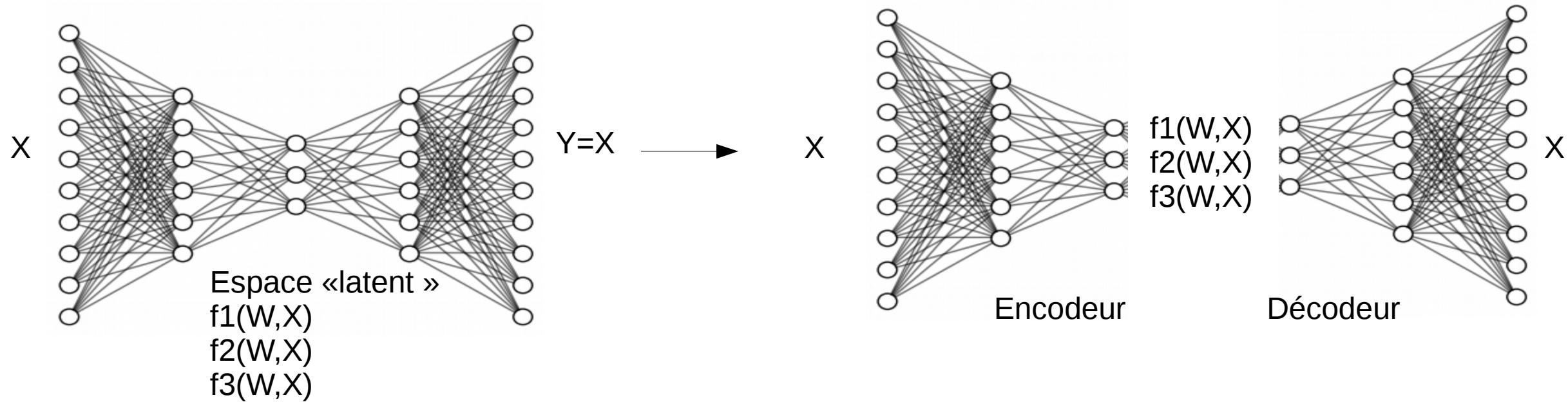
Réseau de neurones hybride :
3 auto-encodeurs et 3 MLP.
1 modèle = 6 entraînements.

Entraînement automatisé :
à architecture donnée.

Environment :
Scikit-learn – Tensorflow

Taille :
~20 couches, $\sim 10^6$ paramètres.

Recherche des hyperparamètres : auto-encodeurs



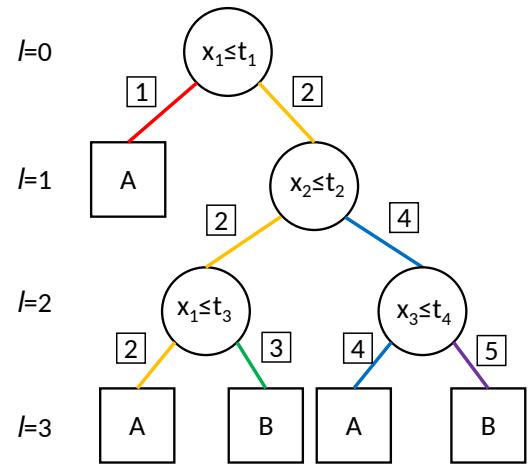
On fige l'architecture, les fonctions d'activation, l'algorithme d'optimisation, l'initialisation, le nombre d'itérations, la taille de batchs.

2 hyperparamètres : la **profondeur** et la **dimension de l'espace latent**.

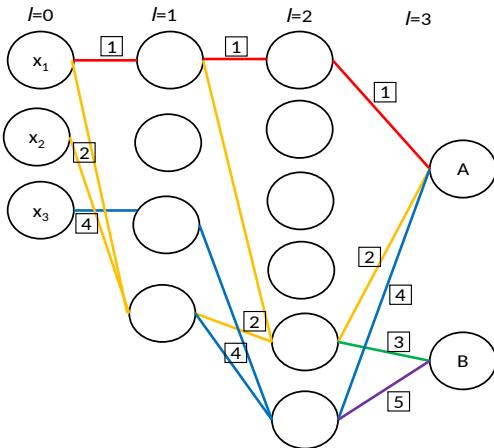
On baisse la dimension de l'espace latent jusqu'à ce que l'on n'arrive plus à encoder-décoder correctement.

Recherche des hyperparamètres : MLP

Arbre de décision



Réseau de neurones initialisé associé.



On construit un arbre de décision binaire (Scikit-learn) : pas d'hyperparamètres; bon premier résultat.

L'architecture est donnée par un mapping entre l'arbre et un réseau de neurones.

On fige les fonctions d'activation, l'algorithme d'optimisation, le nombre d'itérations, la taille de batchs.
[phd Kelli Humbird 2018], [Humbird, Peterson, McClarren IEEE Trans. Plasma Sci. 2020]

1 hyperparamètre : la **profondeur**.

On peut affiner autour de ce premier réseau.

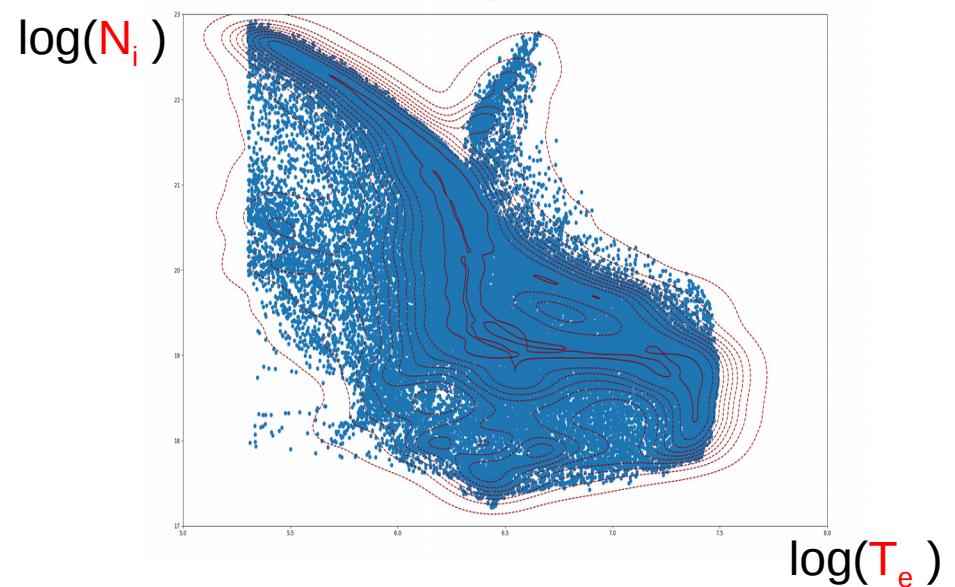
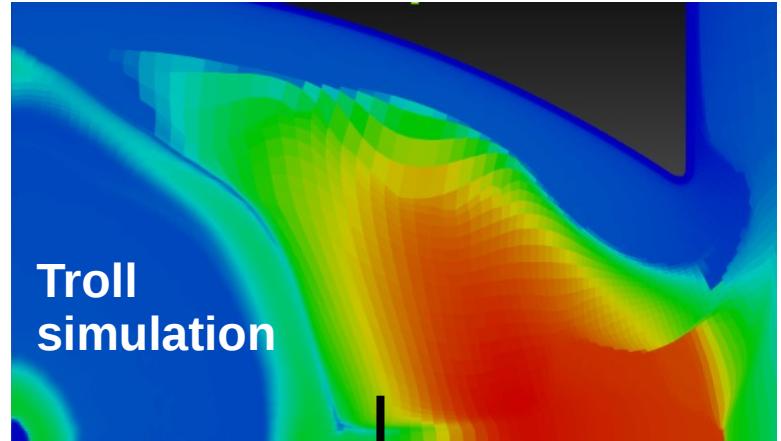


Recherche des hyperparamètres : conclusion

- La construction du **premier réseau** a été fastidieuse :
choix des variables, des unités, des transformations adaptées (logarithmes et normalisation),
choix des données,
choix des hyper-paramètres.
- D'autant plus que la **métrique** d'erreur dans les entraînements n'est pas reliée à celle de nos simulations intégrées : quelle précision faut-il obtenir pendant l'entraînement pour finalement avoir un calcul hydrodynamique correct?
- Le passage à de **nouvelles données** (autre code, autres matériaux, autres simulations), avec les mêmes hyperparamètres, se passe bien.
- L'entraînement complet a été **automatisé**.

- Nous sommes désormais dans une phase de simplification de l'architecture et d'amélioration des métriques.

Entraînement sur une simulation 2D : LMJ2019



Sur LMJ2019 :

Troll² avec Dedale¹ en ligne.

~20k mailles, 64 CPUs, 1 jour, 101 groupes.

~50.10⁶ d'appels HTL Dédale.

Extraction de données :

~ 1.10⁶ données ($N_i, T_e, I(v)$) pour l'or.

==> **Echantillonnage du spectre**

Calculs Dedale hors ligne :

~ 1.10⁶ données ($Z, T_z, \kappa(v), \eta(v)$).

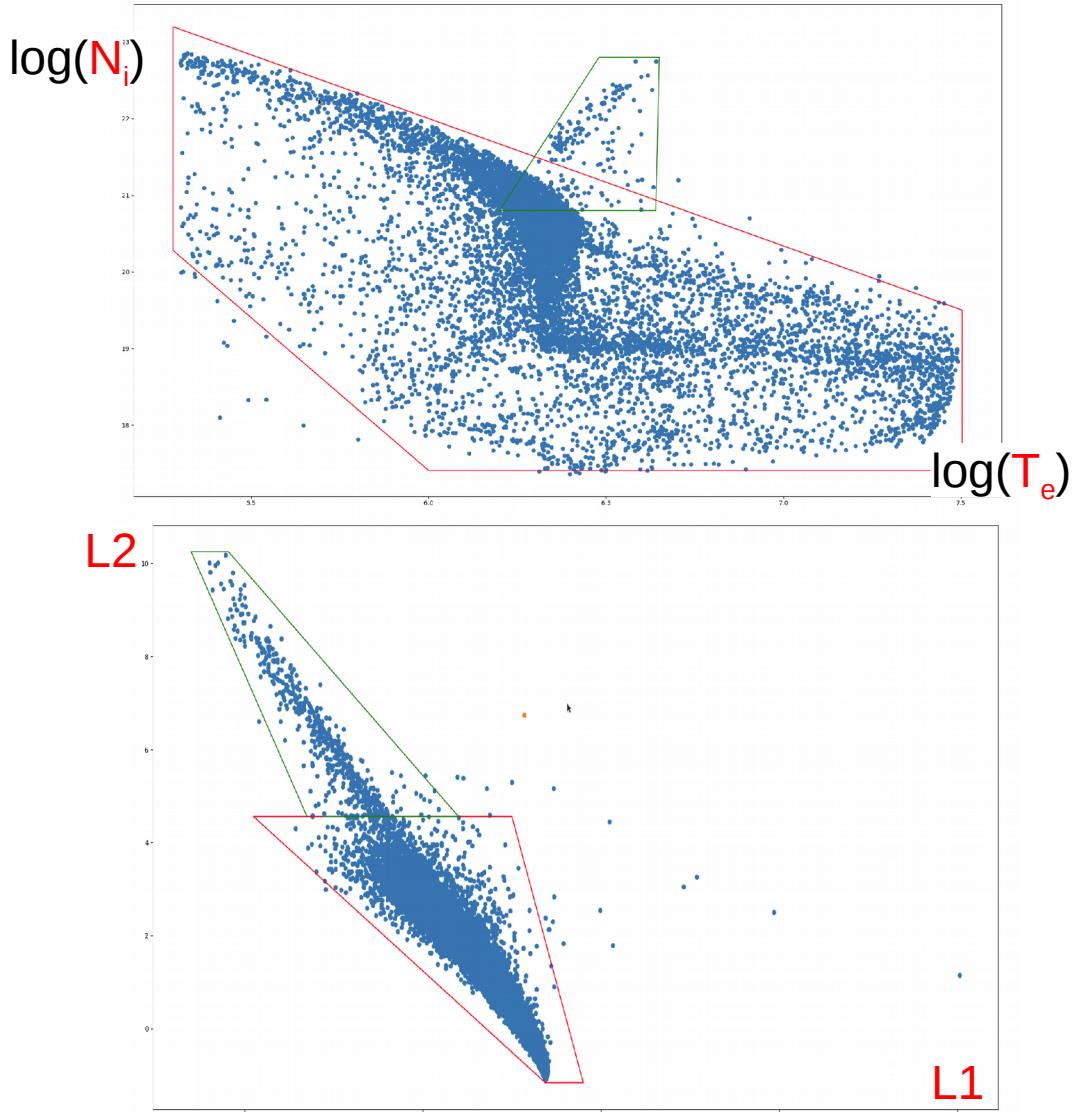
Entraînement :

avec 1.10⁵ données.

¹ [Gilleron, Piron. «The fast non-LTE code Dédale » HEDP 2015]

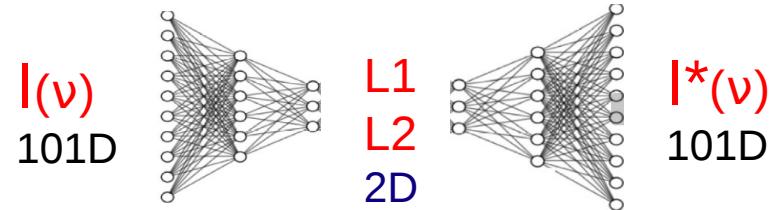
² [Lefebvre, Bernard, Esnault, Gauthier, Grisollet, Hoch, Jacquet, Kluth, Laffite, Liberatore, Marmajou, Masson-Laborde, Morice, Willien. « Development and validation of the troll radiation-hydrodynamic code for 3D hohlraum calculations. » Nucl. Fusion 2019]

Domaine de validité



Dans l'espace latent de dimension 4 :

Nous choisissons des polygônes
dans les plans $(\log(N_i), \log(T_e))$ et (L_1, L_2) .



En dehors du domaine → appel Dédale
Dans le domaine → Appel réseau de neurones

Rq : on choisit ici un périmètre assez large (nous privilégions la performance à la précision pour ce premier essai).

Prédiction dans le code Troll :

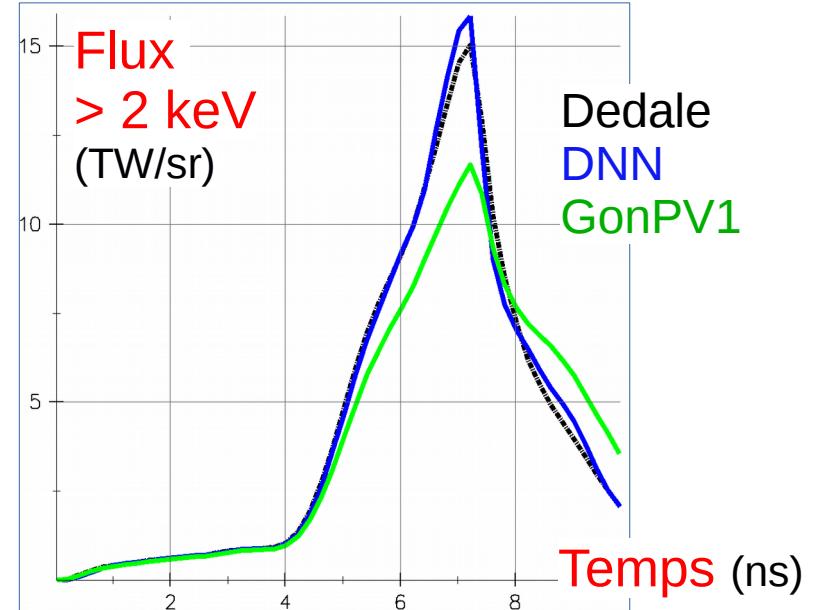
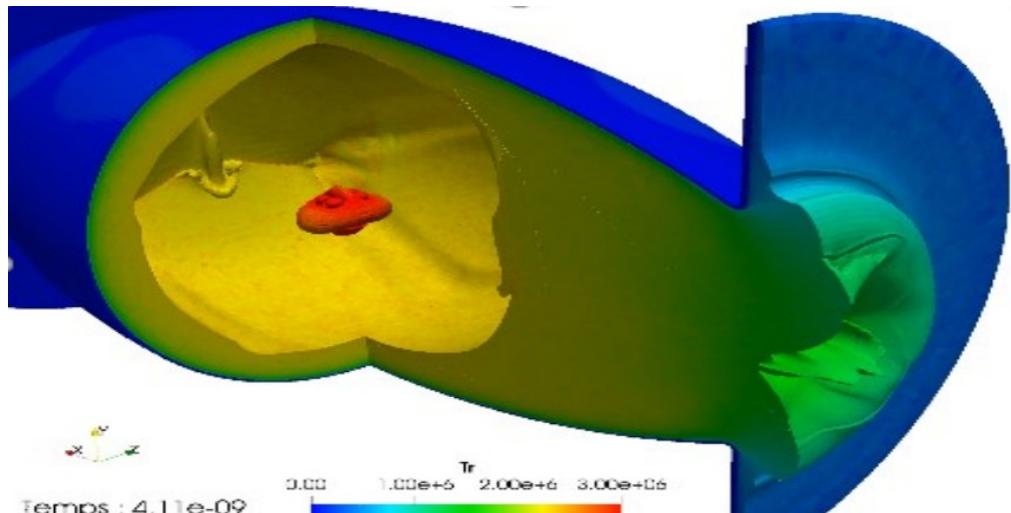
Le résultat de l'entraînement et le domaine de validité sont enregistrés dans un fichier au format **hdf5** (~20Mo).

La prédiction est codée en **C++ avec OpenBlas**.

Résultats des prédictions



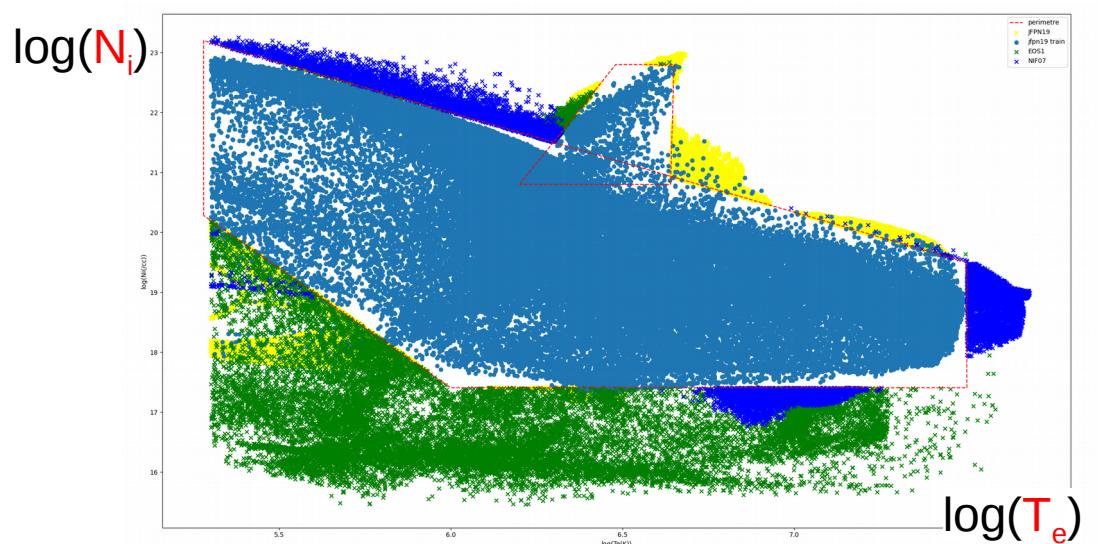
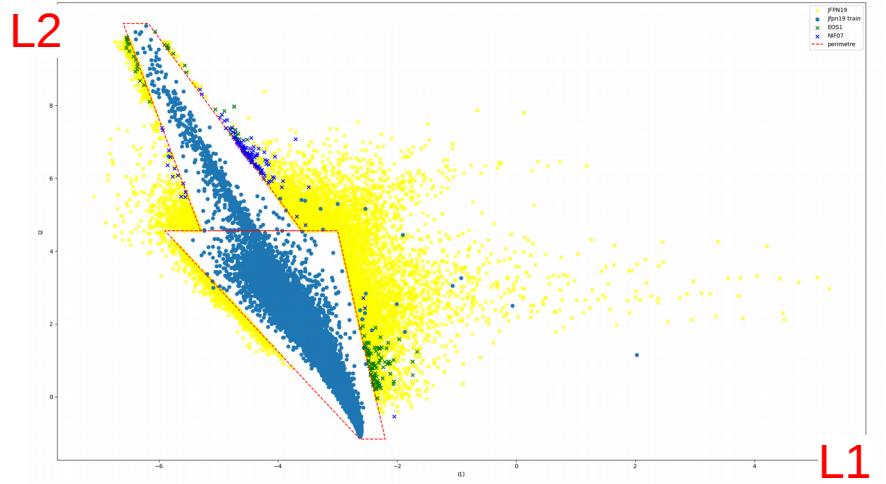
4 simulations variées, avec le même modèle.
2D/3D, cavité gaz/vide, puissances laser variées...
Avec sorties des diagnostics d'intérêt.



Conclusions :

- L'**extrapolation** est à éviter absolument.
- Le **coût calcul** HETL disparaît lorsque l'on utilise uniquement le réseau de neurones.
- Les résultats des diagnostics sont très **proches** mais dans certains cas, pas superposés.

Assimilation de nouvelles données : problème



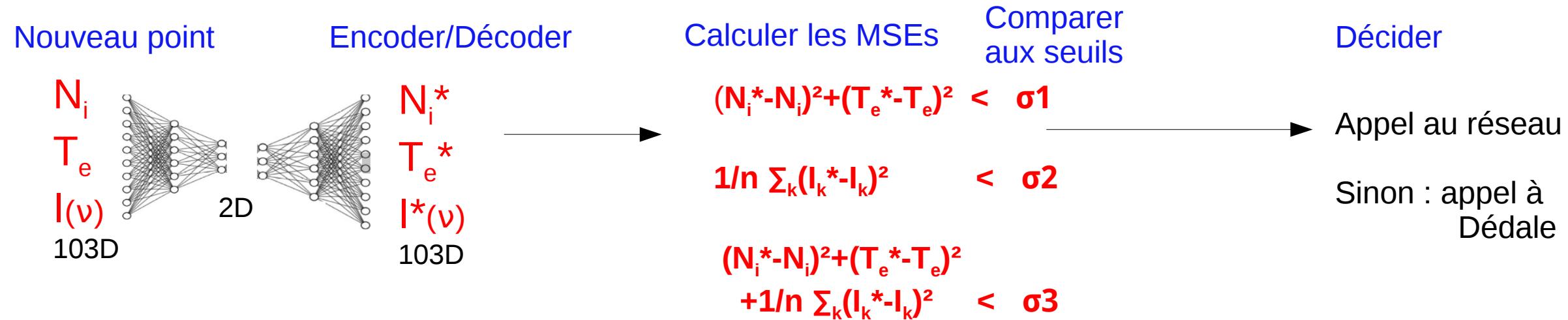
On veut :

- **Une détection des nouveaux points (hors domaine)** automatique, rapide, préservant les corrélations.
→ Lorsque l'on ré-apprend sur de nouvelles données, le domaine de validité est re-défini.
- La possibilité de **ré-échantillonner** en préservant les corrélations.
→ on peut améliorer la précision à domaine fixé.

Assimilation de nouvelles données : solution envisagée

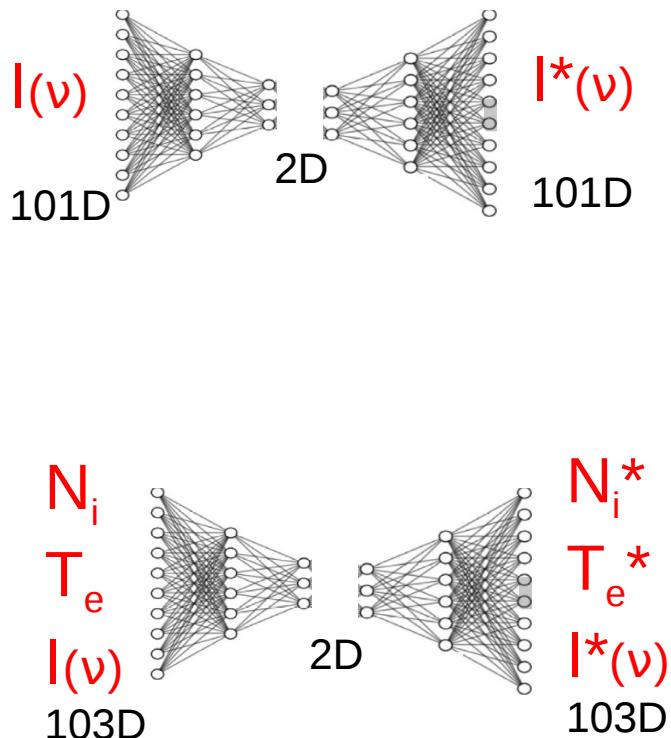
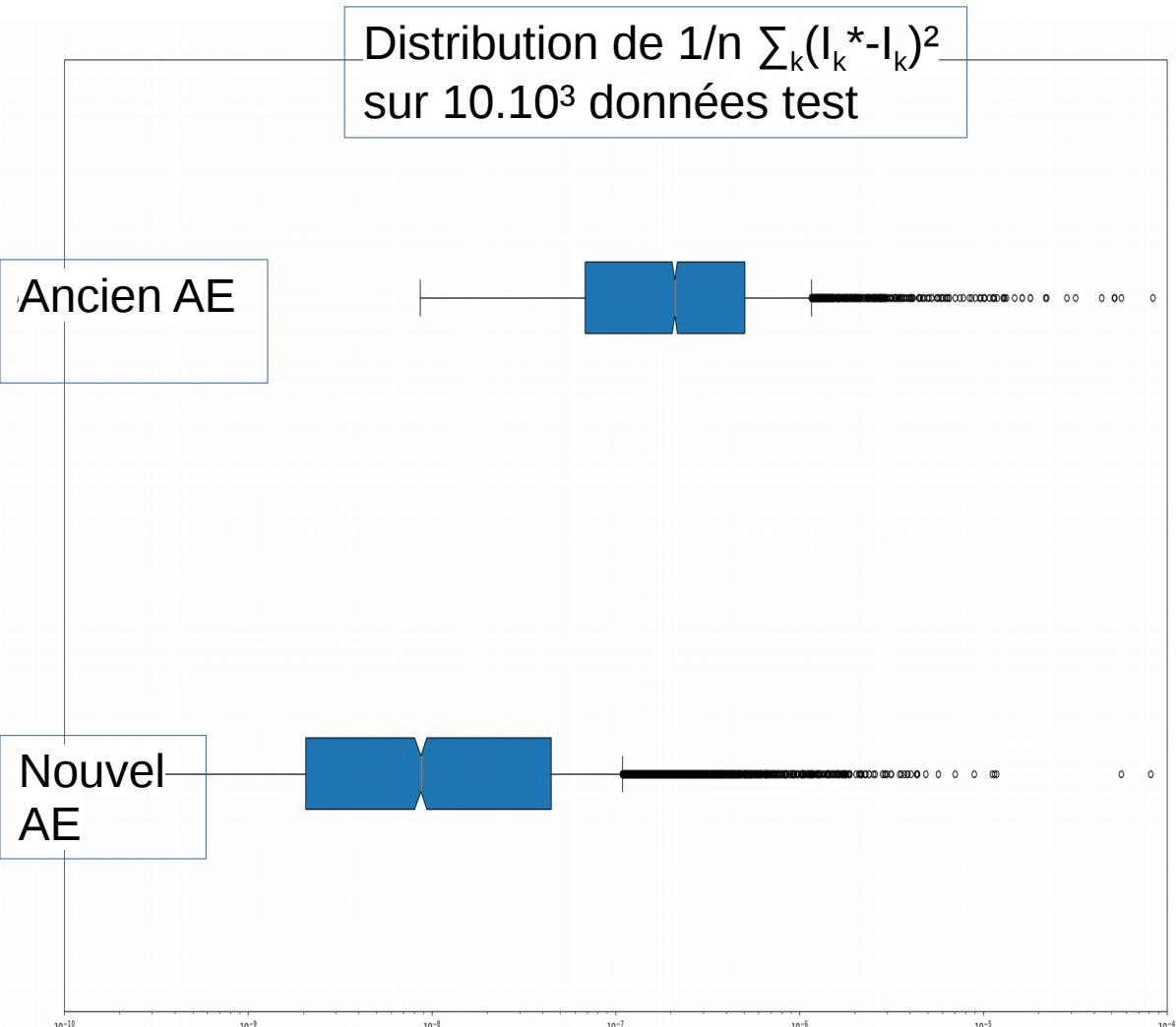
Auto-encodeur pour l'ensemble des entrées.

Calcul de **seuil d'erreur** à partir des données d'entraînement.



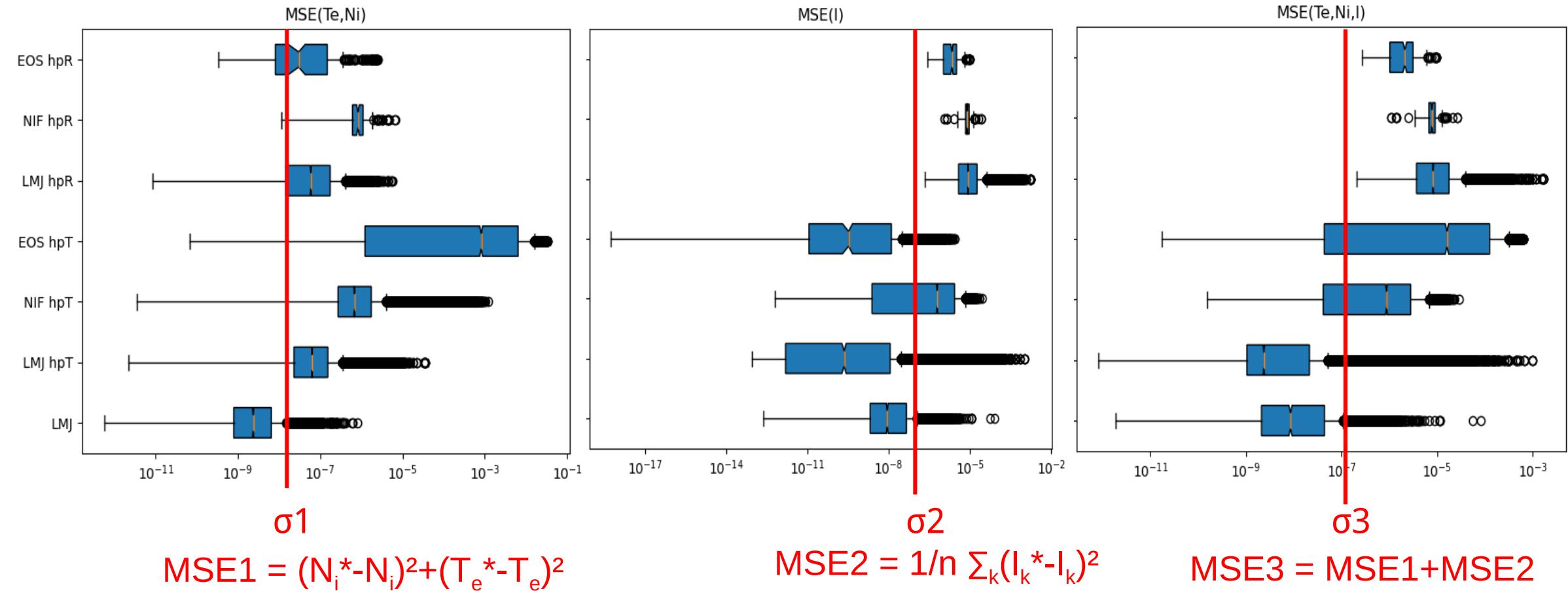
- on retrouve les corrélations de nos données d'entrée
- le test est rapide
- ici on obtient un espace latent de dimension 2, que l'on peut facilement ré-échantillonner.

Autoencoder pour l'ensemble des entrées



Nous avons un meilleur résultat avec le nouvel AE, avec un espace latent de même dimension !

Détection de Nouveauté/Anomalie



Nous montrons ici qu'il est possible avec cette méthode de détecter les points hors périmètre (donnés comme tels avec le domaine de validité initial) comme des nouveaux points.



Perspectives

Deep-Learning

1

MODELES KERAS

Afin d'utiliser les GPUs, le ré-entraînement, le transfert d'apprentissage...

2

1 AUTO-ENCODEUR POUR TOUTES LES SORTIES

Réseaux plus légers et plus précis.

3

RE-ENTRAÎNEMENT DE TOUT LE RÉSEAU

Améliorer la précision

4

TRANSFERT D'APPRENTISSAGE

Code CR plus complexe.

5

METHODES D'ENSEMBLE

Améliorer la précision.

Applications

1

ROBUSTESS DU CODE CR

2

ETENDRE LES SIMULATIONS 2D/3D AVEC LE DOMAINE DE VALIDITÉ AUTOMATIQUE

3

ETENDRE LE DOMAINE DE VALIDITÉ

Peut-on avoir un modèle unique pour toutes les simulations ? Pour un périmètre de simulations ?

4

AMELIORER LA PRÉCISION

En ré-échantillonnant le domaine de validité.

5

DIFFERENTIATION AUTOMATIQUE

Utilisation des dérivées de T_z dans les solveurs.



Merci pour votre attention.

Des questions ?