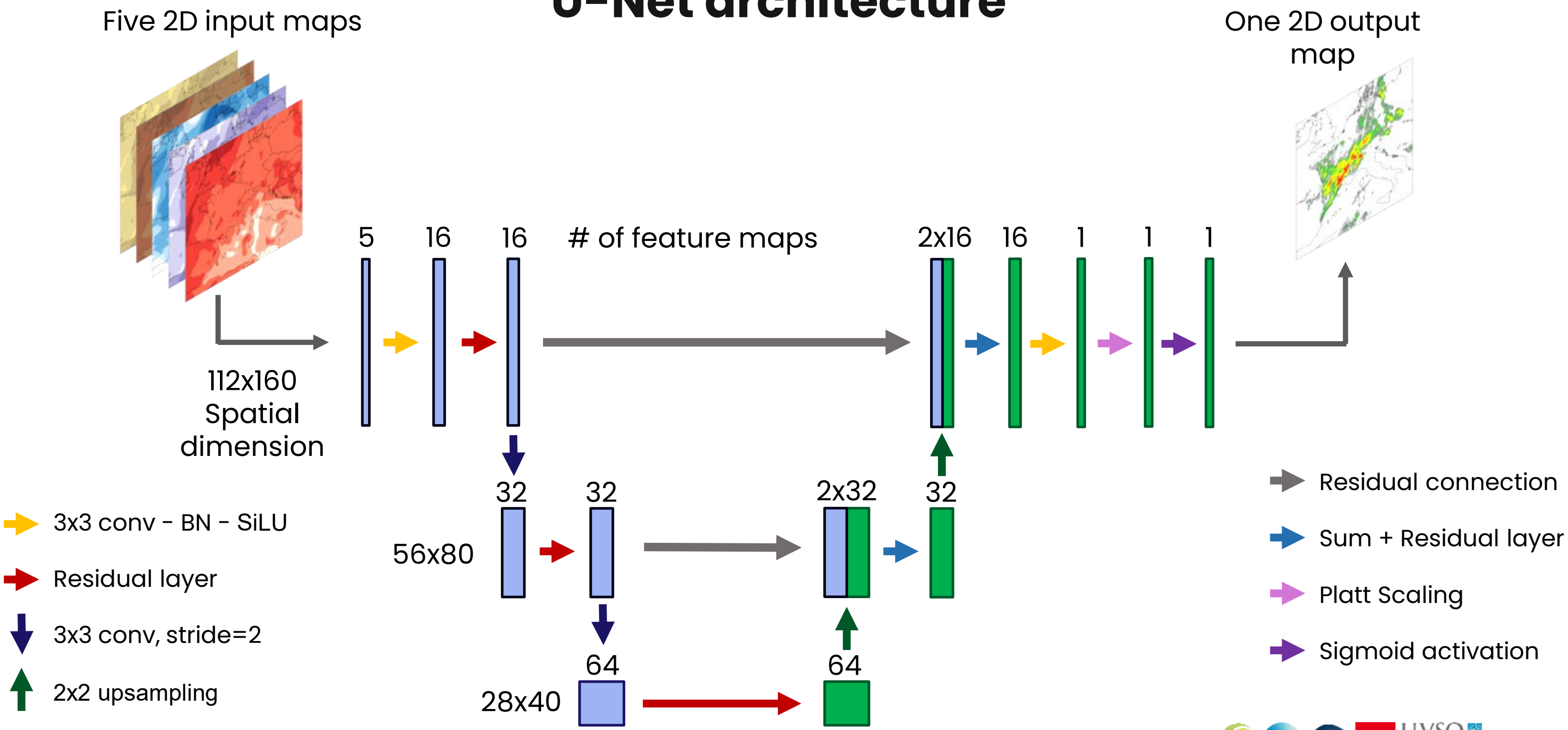
A Deep Learning framework for lightning probabilistic prediction

Supplementary material :

U-Net architecture



A Deep Learning framework for lightning probabilistic prediction

Supplementary material :

Recalibration and optimization process

Recalibration :

→ The U-Net's output is a **map of logits.**

→ The closer the logit is to 0 or 1, the more **confident** the model is in its **classification.**

→ **Not probabilities** yet : we can shift the logits and get the same discriminatory ability.

→ If our loss function is not calibrated, we need to **recalibrate** logits to probabilities.

→ **Platt Scaling** : add a linear layer before the sigmoid : $f(x) = a * x + b$, with a and b being the same parameters for every pixel.

→ **Fit** this linear regression **after** training the rest of the network and **freezing** all other weights

Hyper-parameters tuning with Optuna :

- **DL Training** hyperparameters
- **Data imbalance** handling
- U-Net's **weights**

Loss function : Weighted Binary Cross Entropy with logits.

Metrics used for optimisation : **ROC-AUC** and **Deviance** (to follow Battaglioli et al.), **Fractionnal Brier Score, Precision-Recall AUC, Dice Loss**

Ensemble metrics : Spread-to-skill ratio, Continuous Ranked Probability Score

A Deep Learning framework for lightning probabilistic prediction

Supplementary material :

Other models

Climatology :

Pixel-wise and season-wise climatology :

→ At each pixel and for each season, get the proportion of hours with lightning in the training set.

Logistic Regression :

$$g(y) = \sigma(\beta_0 + a_1 * x_1 + a_2 * x_2 + \cdots + a_m * x_m)$$

→ The x_i are : **most unstable lifted index / mixing ratio, convective precipitation, average relative humidity** and **elevation**.

→ σ is the sigmoid function.

Generalized Additive Model (GAM), following Battaglioli et al., 2023 : $g(y) = \beta_0 + f_1(x_1) + f_2(x_2) + \cdots + f_m(x_m)$

→ Here the $\boldsymbol{f_i}$ are belong to the exponential family.

XGBoost :

→ 100 trees, max depth of 5, 0.1 learning rate, BCE oss func

Multi Layer Perceptron :

→ 3 layer model of size [5, 16, 32, 16, 1]

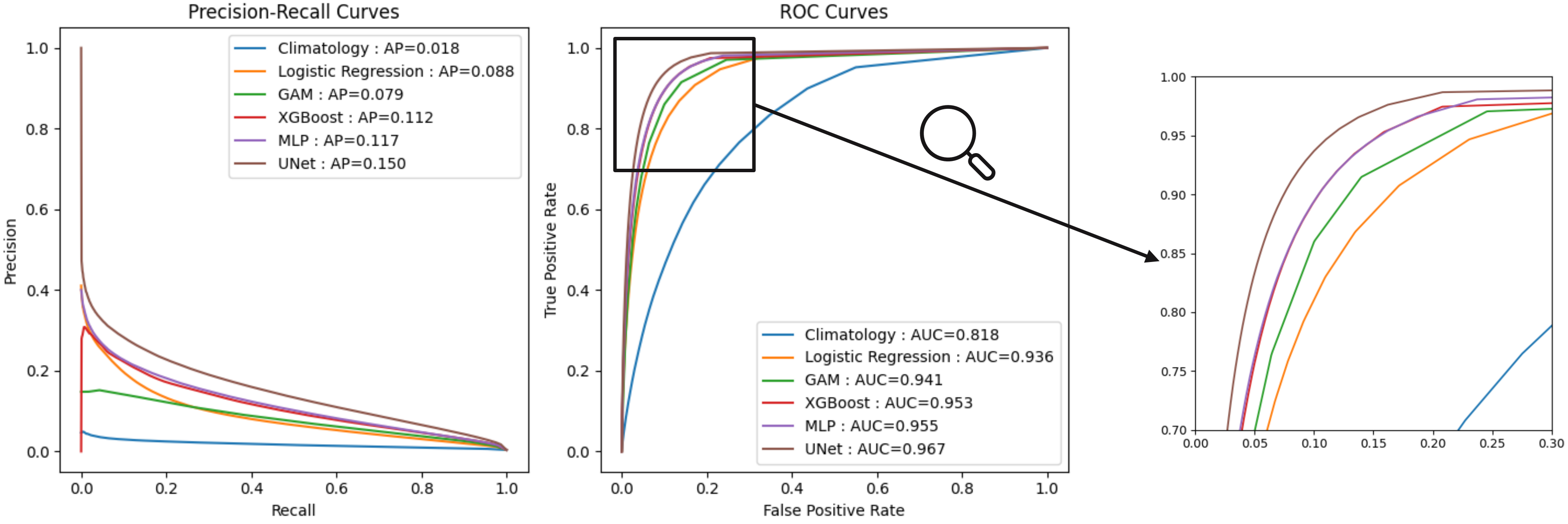
→ Weighted BCE loss like U-Net

→ Platt scaling recalibration layer like the U-Net

A Deep Learning framework for lightning probabilistic prediction

Supplementary material :

ROC and Precision-Recall curves


Precision-Recall Curves
Climatology : AP=0.018
Logistic Regression : AP=0.088
GAM : AP=0.079
XGBoost : AP=0.112
MLP : AP=0.117
UNet : AP=0.150
Precision
Recall
ROC Curves
Climatology : AUC=0.818
Logistic Regression : AUC=0.936
GAM : AUC=0.941
XGBoost : AUC=0.953
MLP : AUC=0.955
UNet : AUC=0.967
True Positive Rate
False Positive Rate