

# DÉVELOPPEMENT D'UN ALGORITHME DE « SCORING » PAR APPRENTISSAGE PROFOND POUR L'ÉVALUATION DE CARTES DE PRÉCIPITATIONS

Encadrement : L. Barthes (LATMOS), C. Mallet (LATMOS), P. Lepetit (Météo-France).

### Problématique

**But:** Inter-comparaison du produit de pluies IMERG et Mosaïque (Météo-France(MF))

**Hypothèses prises par IMERG:**

- Advection des cellules de pluies pour boucher « les trous »
- Cellules supposées indéformables

**Limites d'IMERG:**

- Surestimation des intensités de pluies aux frontières mer/terre  
→ Une extraction sur une zone de 700 km x700 km excluant la mer et les zones côtières a été effectuée au dessus de la France

**Approche:** Créer un auto-encodeur et comparer les cartes dans l'espace latent

### Caractérisation des Classes

Classes 1, 3, 5, 6, 13, 16, 18 → Pas de pluies

Classes 4, 7, 8, 9, 11, 14 → Pluies faible

Classes 2, 10, 12, 14 → Pluies Moyennes

Classes 15, 17, 19 → Pluies Très fortes

**La classe 19 représente les fortes pluies à l'Est de la France vu la présence des montagnes**

### Auto-encodeur

**Base de données**

- Données météo années 2017 → 2020
- Images 70x70 pixels Résolution pixel 10kmx10km
- Repartis en 19 classes météorologiques

**Architecture du réseau (Auto-encodeur Variationnel)**

**Validation 1/2 2020: 9,500 cartes MF**

**Test 1/2 2020: 9,500 cartes MF**

**Train 2017 → 2019: 48,140 cartes MF**

### Auto-encodeur

**Base déséquilibrée:** Le nombre de cartes pluvieuses est faible comparé au nombre de cartes où il ne pleut pas

Train dataset (2017 → 2019)		Val/Test dataset (2020)	
Classes Pluie Forte	% de Nombre d'Images	Classes Pluie Forte	% de Nombre d'Images
15	1.96%	15	3.24%
17	1.18%	17	1.43%
19	0.31%	19	0.74%

Test	Tirage avec remise	Ajout du bruit	Erreur de Test
1	Non	Non	0.0754
2	Non	Oui	0.0744
3	Oui	Non	0.0623
4	Oui	Oui	0.0621

**Augmentation de données:**

- Shift: décaler un peu les cellules pluvieuses
- Superposer deux cartes pluvieuses
- Ajouter du bruit aux cartes

**Privilégier les classes rares:**

- Durant la phase d'apprentissage, le tirage se fait avec remise et donnant une probabilité plus forte aux classes de pluies intenses qui sont rares

**La plus faible valeur de l'erreur de test correspond au Test 4**

### Etude des résidus

**Comparaison des Histogrammes de Résidus du Test 1 et du Test 4**

**Le résidus augmente dans les classes pluvieuses**

**Le résidus des classes non pluvieuses ~ 0**

### Perspective des prochaines étapes

Comparer dans les espaces latents les cartes MF et IMERG à l'aide de méthodes de machine learning (clustering, classification, ...)

Etudier la cohérence entre MF et IMERG

**Cohérence** (thumbs up)

**Cohérence** (thumbs down)