



SYNTHÈSE PAR SCENARIOS DE PLUIES DES PRÉVISIONS AROME

Journées de l'IA 2026

*Merci à Benoît Hesse, Léo Mignan (Ex-prévisionnistes à Rennes), Jean Imbert (Ex-prévisionniste à Aix-en-Provence),
Christophe Calas (Chef prévisionniste national à Toulouse),... (liste non exhaustive)*

Toujours plus de données à disposition, mais pas plus de temps pour les explorer

Supposons que nous souhaitons ne regarder qu'une seule variable météorologique...

AROME et AROME-IFS
(51h, 4 fois par jour) =
408 cartes

Prévision d'Ensemble
AROME
(51h, 25 membres, 4
fois par jour) =
5 100 cartes

ARPEGE
(102h ou 114h, 4 fois
par jour) =
420 cartes

Prévision d'Ensemble
ARPEGE
(102h, 35 membres, 4
fois par jour) =
14 280 cartes

échelle



échelle



Prévisionnistes

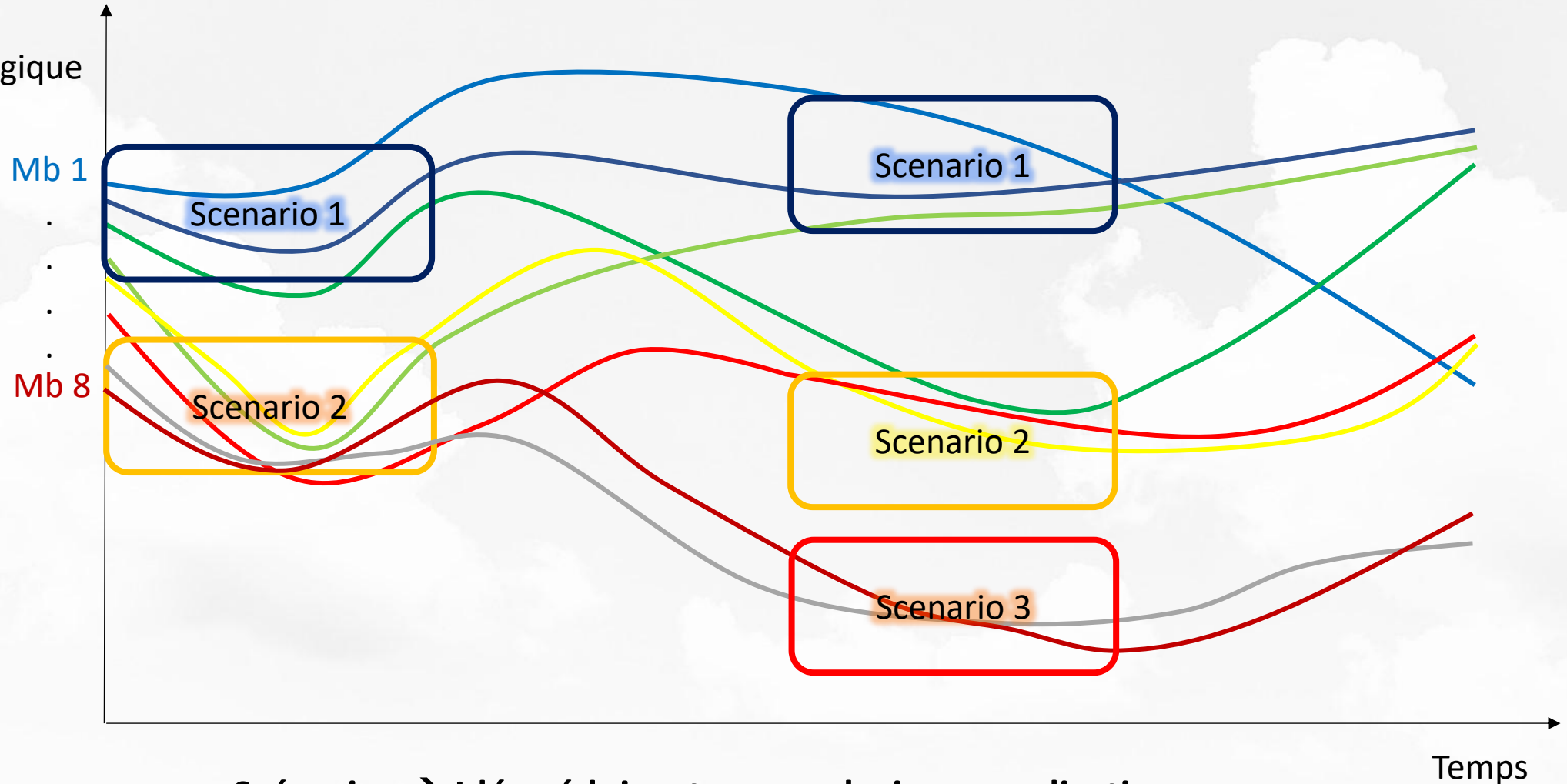


Et à la fin, des prises de décision,
avec une information synthétique et
claire à délivrer

On peut aussi rajouter AROME-PI, les
AROME-OM, les modèles étrangers, le
modèle du CEP,
AIFS, et bientôt AROME-IA ?

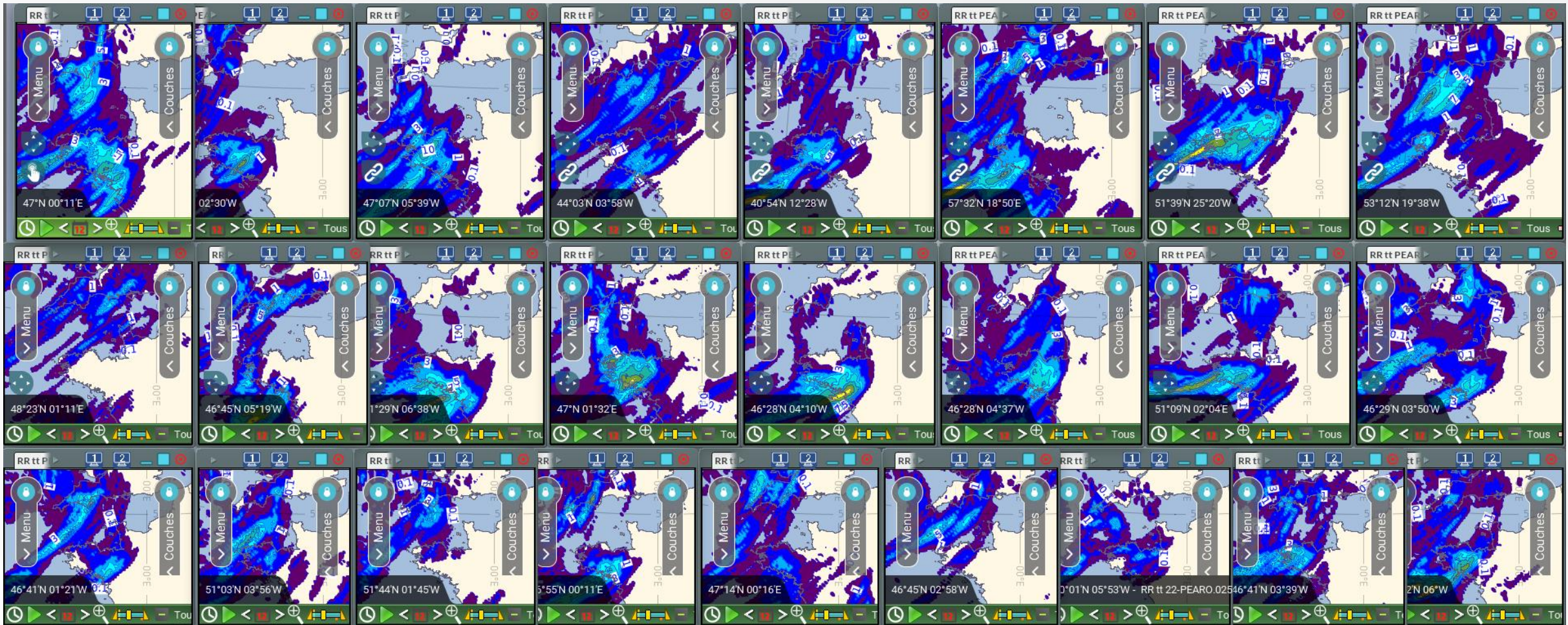
Plusieurs solutions possibles, dont les scénarios

Variable
météorologique



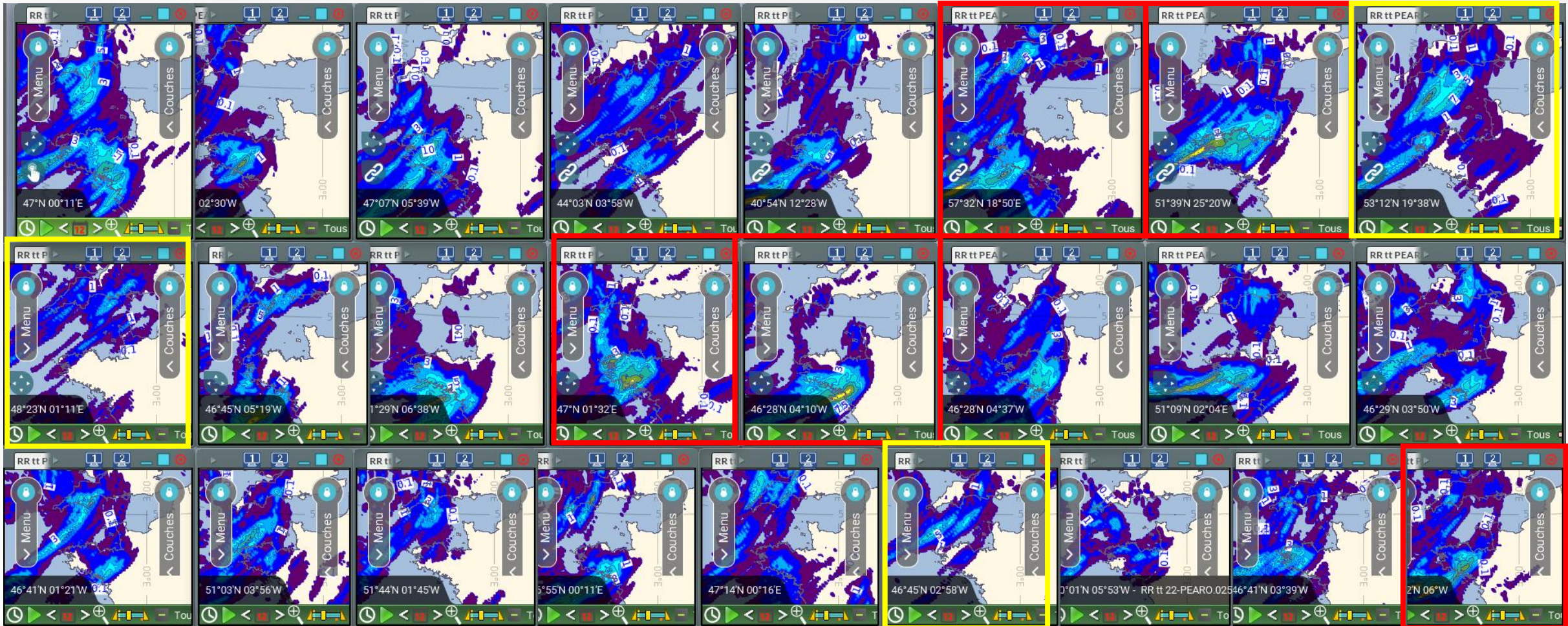
Scénarios → Idée séduisante pour plusieurs applications

Exemple concret sur une PE-AROME



Cumul de pluies sur 3h (RR3) de 25 membres. Comment regrouper les membres entre eux ?

Exemple concret sur une PE-AROME

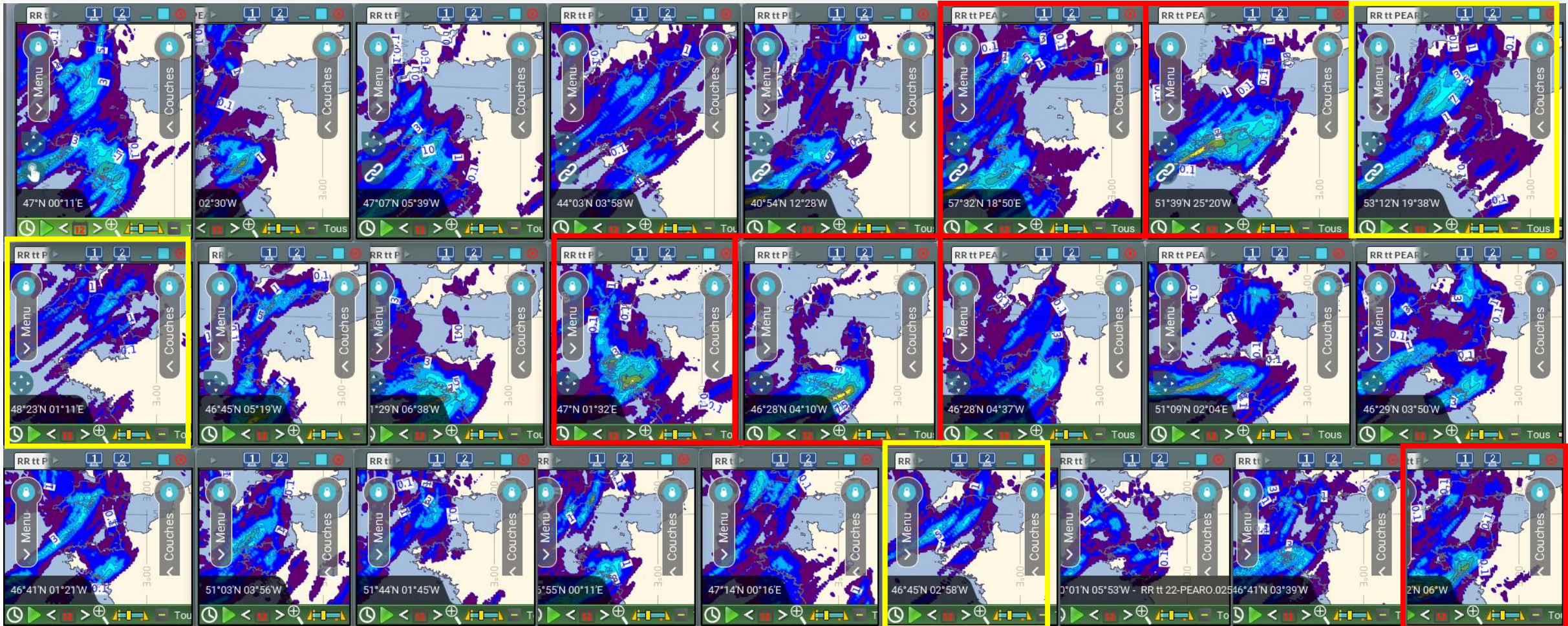


Cumul de pluies sur 3h (RR3) de 25 membres. Comment regrouper les membres entre eux ?

Membres + forts, +faibles

Arnaud Mounier - JIA 2026

Exemple concret sur une PE-AROME



Conclusion: c'est compliqué, subjectif partiellement, dépend des similitudes à mettre en avant ou pas (Position des pluies ? Intensités des pluies ? Certaines zones n'ont-elles pas plus d'intérêts que d'autres ?)

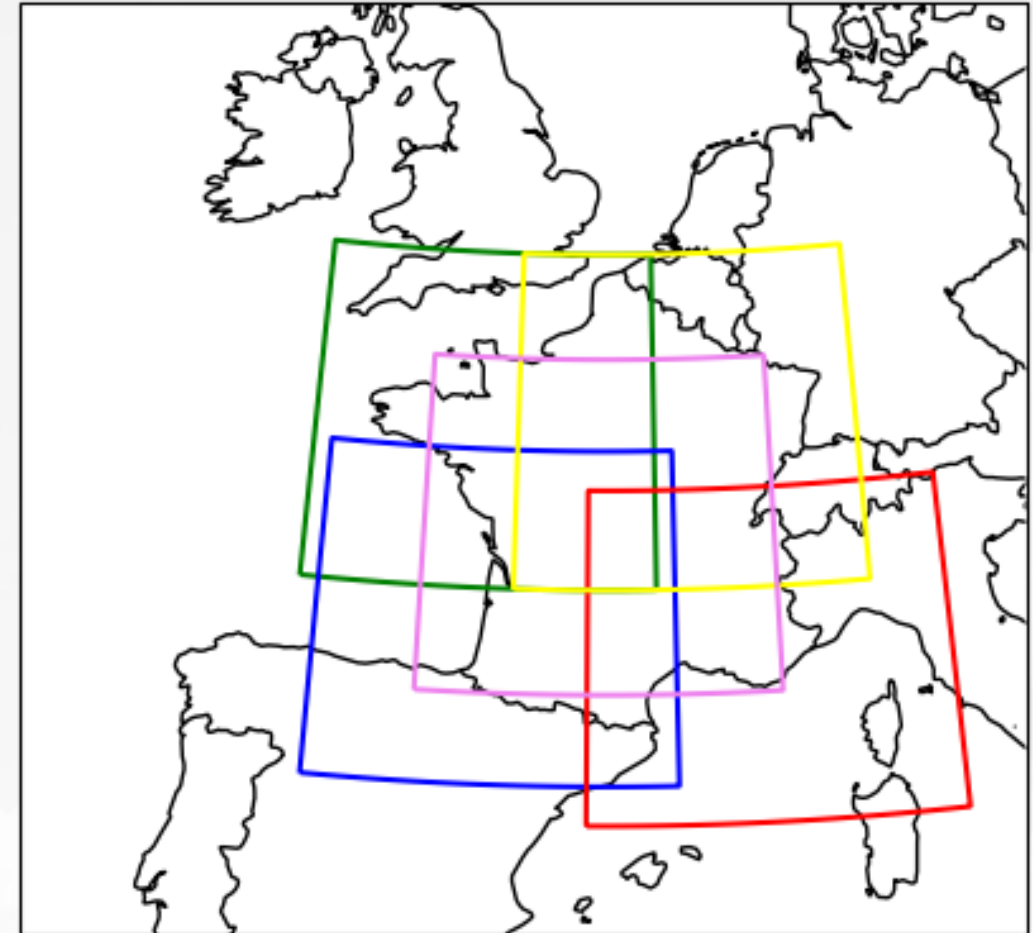
- ❖ Synthèse par scénarios de RR3 incluant:
 - 2 PE-AROME successives (2x25 membres)
 - 2 AROME
 - 2 AROME-IFS

- ❖ Taille caractéristique: \approx quart de France

- ❖ Scénarios définis sur un intervalle de temps: 7h (7 échéances consécutives)

- ❖ On ne représentera pas plus de 4 scénarios différents.

- ❖ Informations explicites sur la sévérité des scénarios et la position des AROME et AROME-IFS



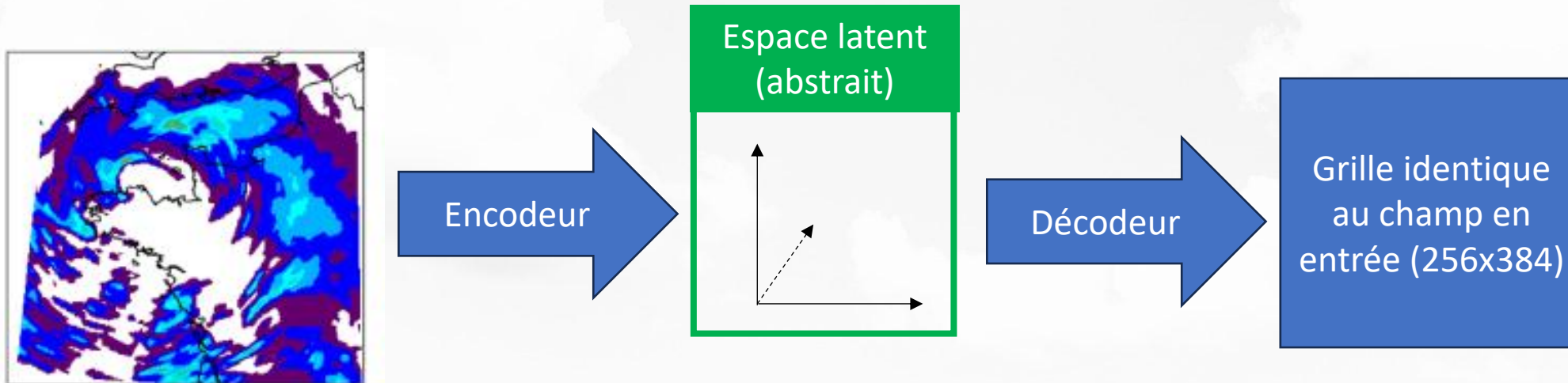
Méthode pour construire les scénarios (étape par étape)

Note: la méthode pour construire les scénarios est longue à expliquer, construite après plusieurs tentatives plus simples mais infructueuses... Je vais tenter d'être le plus clair possible.

Un des problèmes majeurs pour faire des scénarios: il faut définir une métrique pour calculer une distance entre les membres deux à deux. Or, on sait que les distances « basiques » ne fonctionnent pas (problème lié à la dimension).



Idée: **compresser les champs de pluies dans un espace abstrait** d'un réseau de neurones puis calculer les distances à l'intérieur de cet espace. Comment ? **Avec un autoencodeur.**



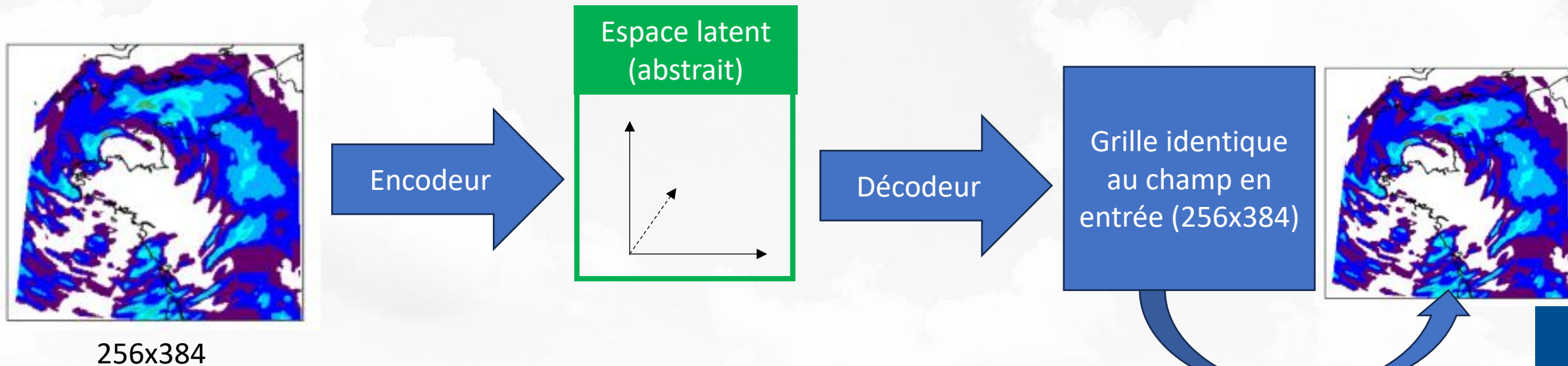
256x384

Méthode pour construire les scénarios (étape par étape)

Un des problèmes majeurs pour faire des scénarios: il faut définir une métrique pour calculer une distance entre les membres deux à deux. Or, on sait que les distances « basiques » ne fonctionnent pas (problème lié à la dimension).



Idée: **compresser les champs de pluies dans un espace abstrait** d'un réseau de neurones puis calculer les distances à l'intérieur de cet espace. Comment ? **Avec un autoencodeur.**



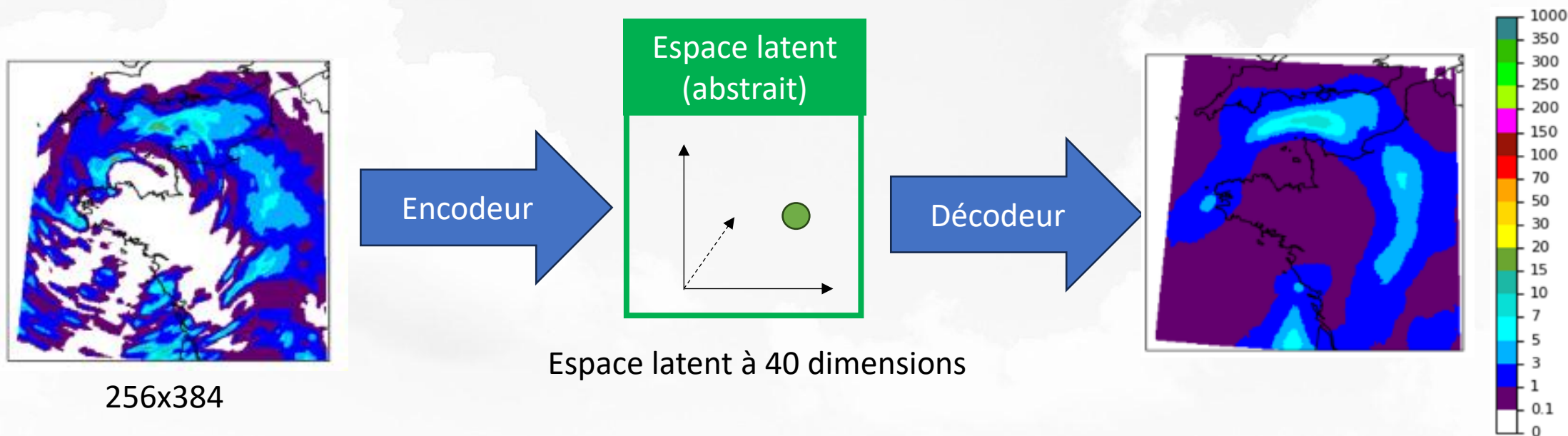
Objectif: reproduire (au mieux) le champ en entrée

Méthode pour construire les scénarios (étape par étape)

Un des problèmes majeurs pour faire des scénarios: il faut définir une métrique pour calculer une distance entre les membres deux à deux. Or, on sait que les distances « basiques » ne fonctionnent pas (problème lié à la dimension).



Idée: **compresser les champs de pluies dans un espace abstrait** d'un réseau de neurones puis calculer les distances à l'intérieur de cet espace. Comment ? **Avec un autoencodeur.**



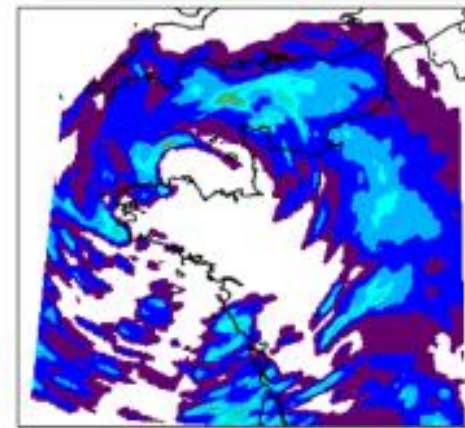
256x384

Espace latent à 40 dimensions

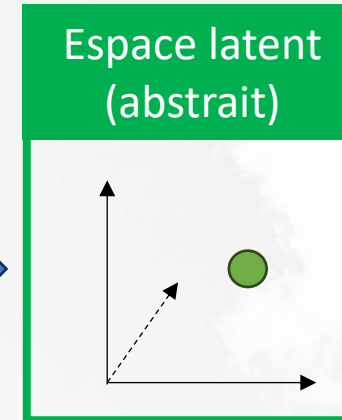
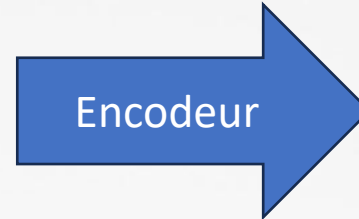
Après entraînement de l'IA

Méthode pour construire les scénarios (étape par étape)

Maintenant que nous avons un moyen de compresser la donnée (passage de 256x384 points de grille à 40 dimensions), on ne garde que l'**encodeur**.



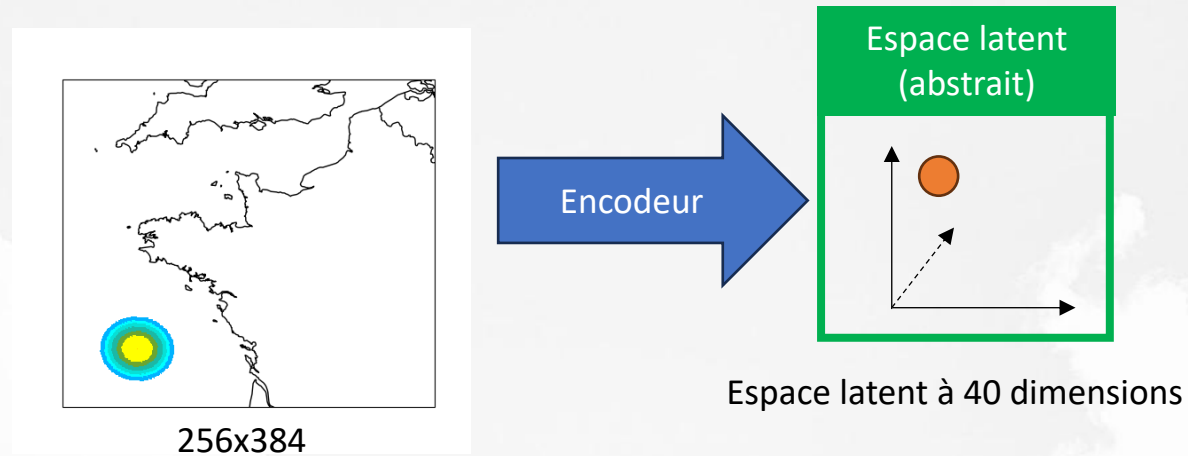
256x384



Espace latent à 40 dimensions

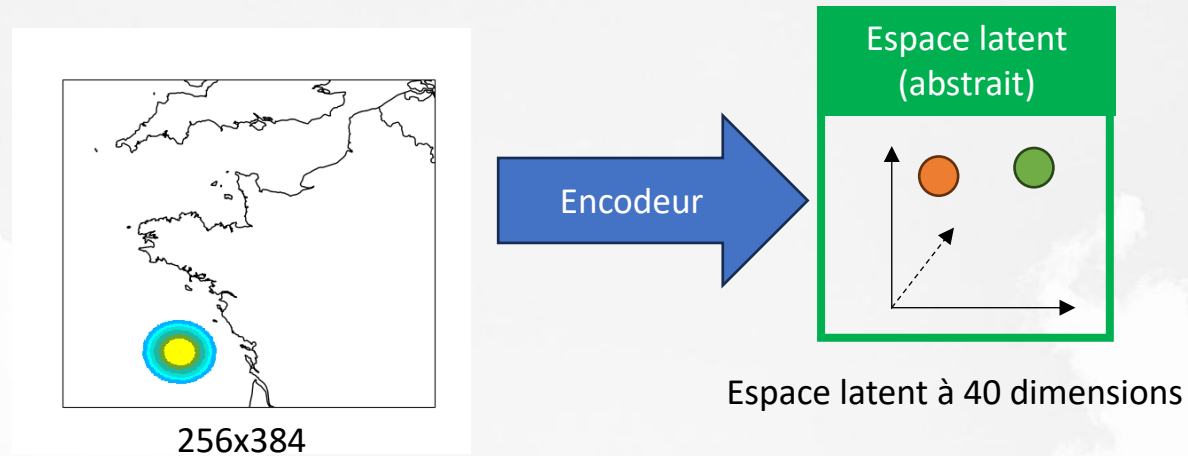
Méthode pour construire les scénarios (étape par étape)

Nous allons maintenant créer des patterns idéalisés de pluies et les projeter dans l'espace abstrait avec l'encodeur.



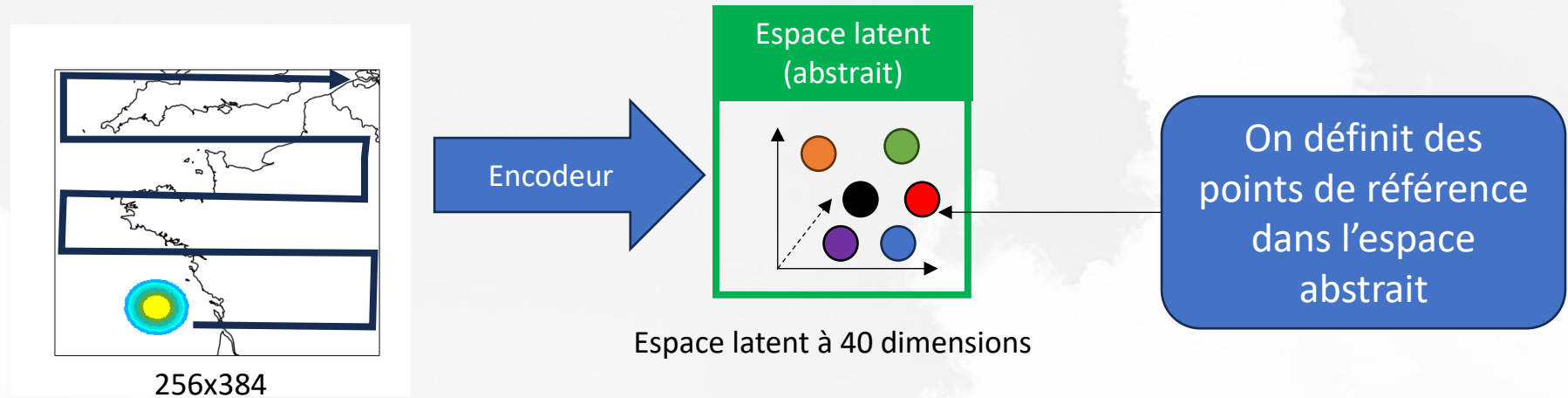
Méthode pour construire les scénarios (étape par étape)

Nous allons maintenant créer des patterns idéalisés de pluies et les projeter dans l'espace abstrait avec l'encodeur.



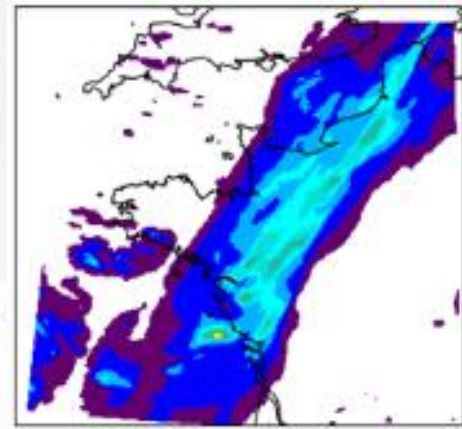
Méthode pour construire les scénarios (étape par étape)

Nous allons maintenant créer des patterns idéalisés de pluies et les projeter dans l'espace abstrait avec l'encodeur.

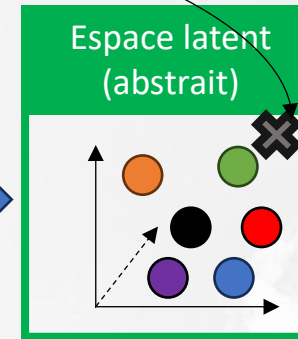
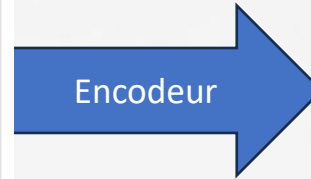


Méthode pour construire les scénarios (étape par étape)

Maintenant, que fait-on avec un champ de RR3 d'une vraie prévision ?



256x384

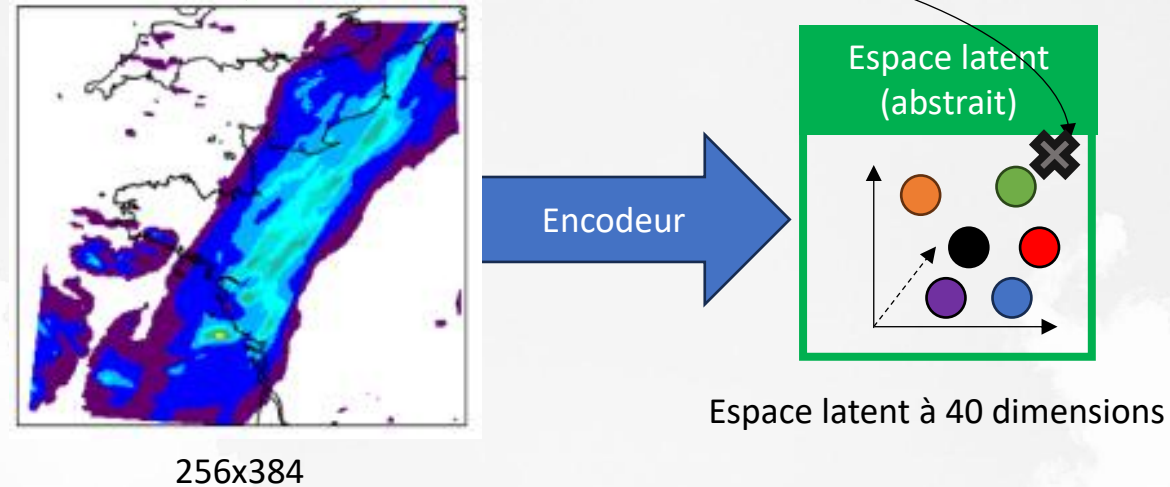


Espace latent à 40 dimensions

On cherche parmi les points de référence celui qui est le plus proche (au sens de la distance euclidienne). On dit alors que le champ de RR3 est assigné à ce point de référence (ici le vert).

Méthode pour construire les scénarios (étape par étape)

Maintenant, que fait-on avec un champ de RR3 d'une vraie prévision ?

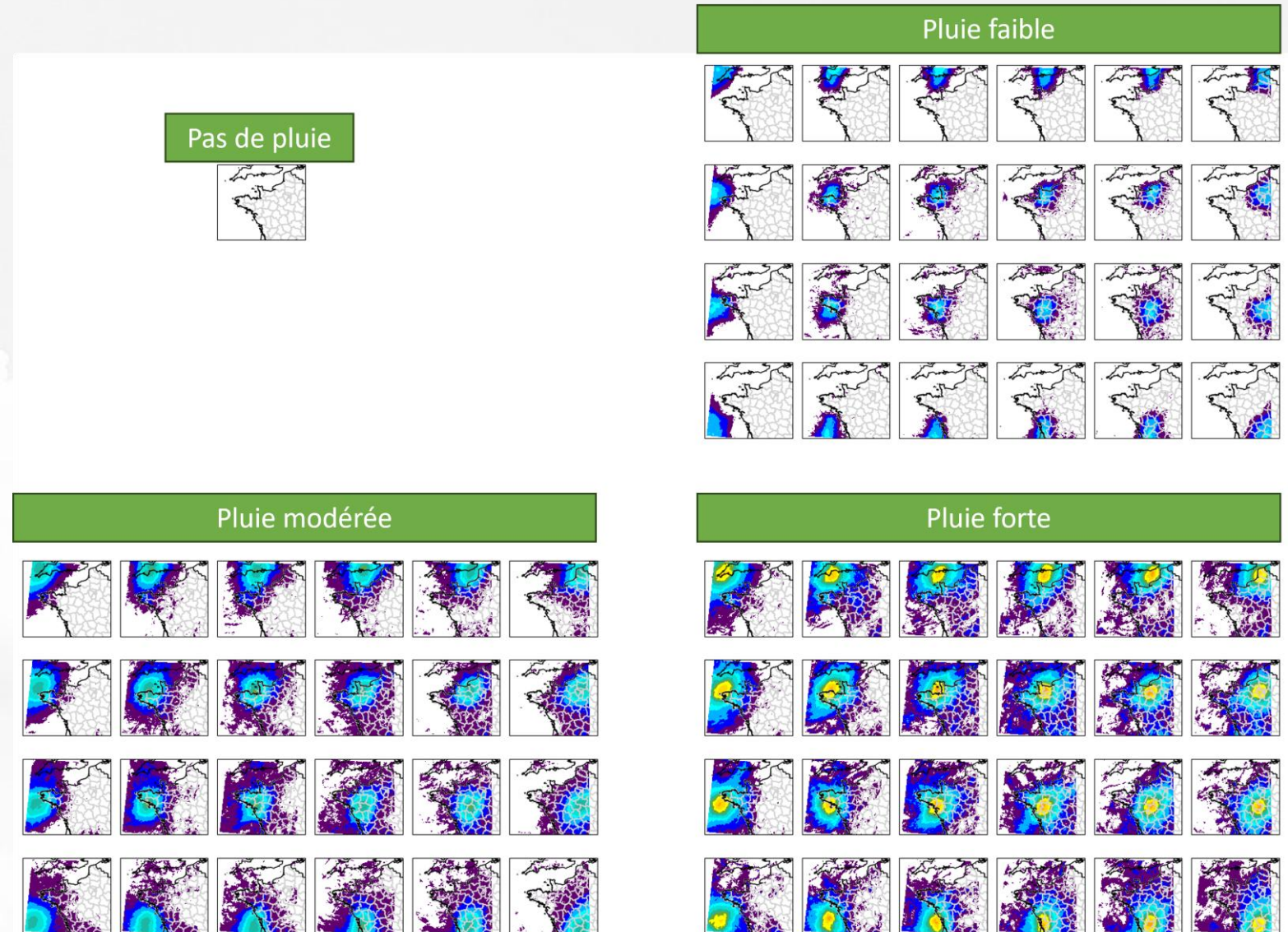


On cherche parmi les points de référence celui qui est le plus proche (au sens de la distance euclidienne). On dit alors que le champ de RR3 est assigné à ce point de référence (ici le vert).

On peut alors reproduire ce schéma sur des milliers de champs de RR3 pour construire une climatologie. On peut regarder par exemple la moyenne des champs de RR3 pour toutes les croix assignées au même point de référence.

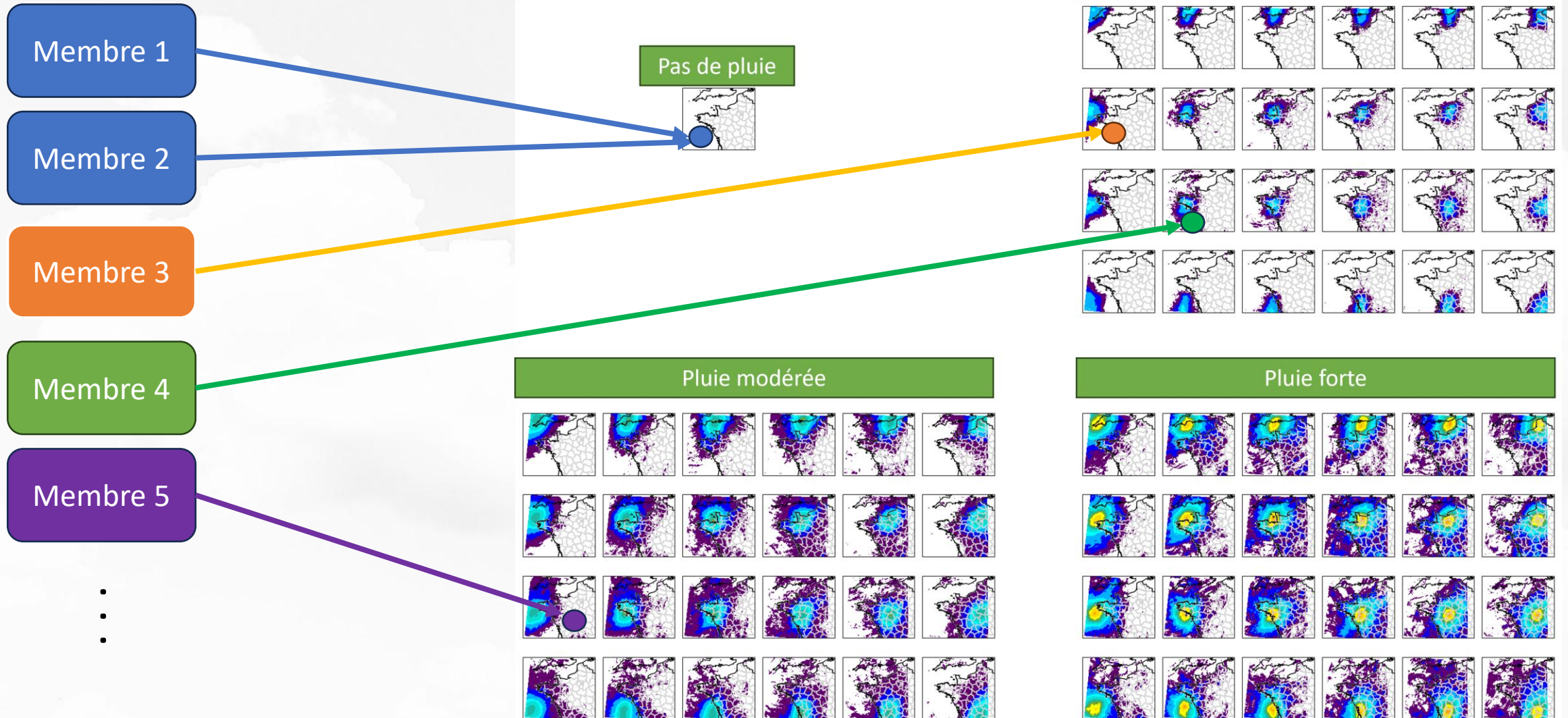
Méthode pour construire les scénarios (étape par étape)

Champ moyen obtenu pour tous les patterns idéalisés



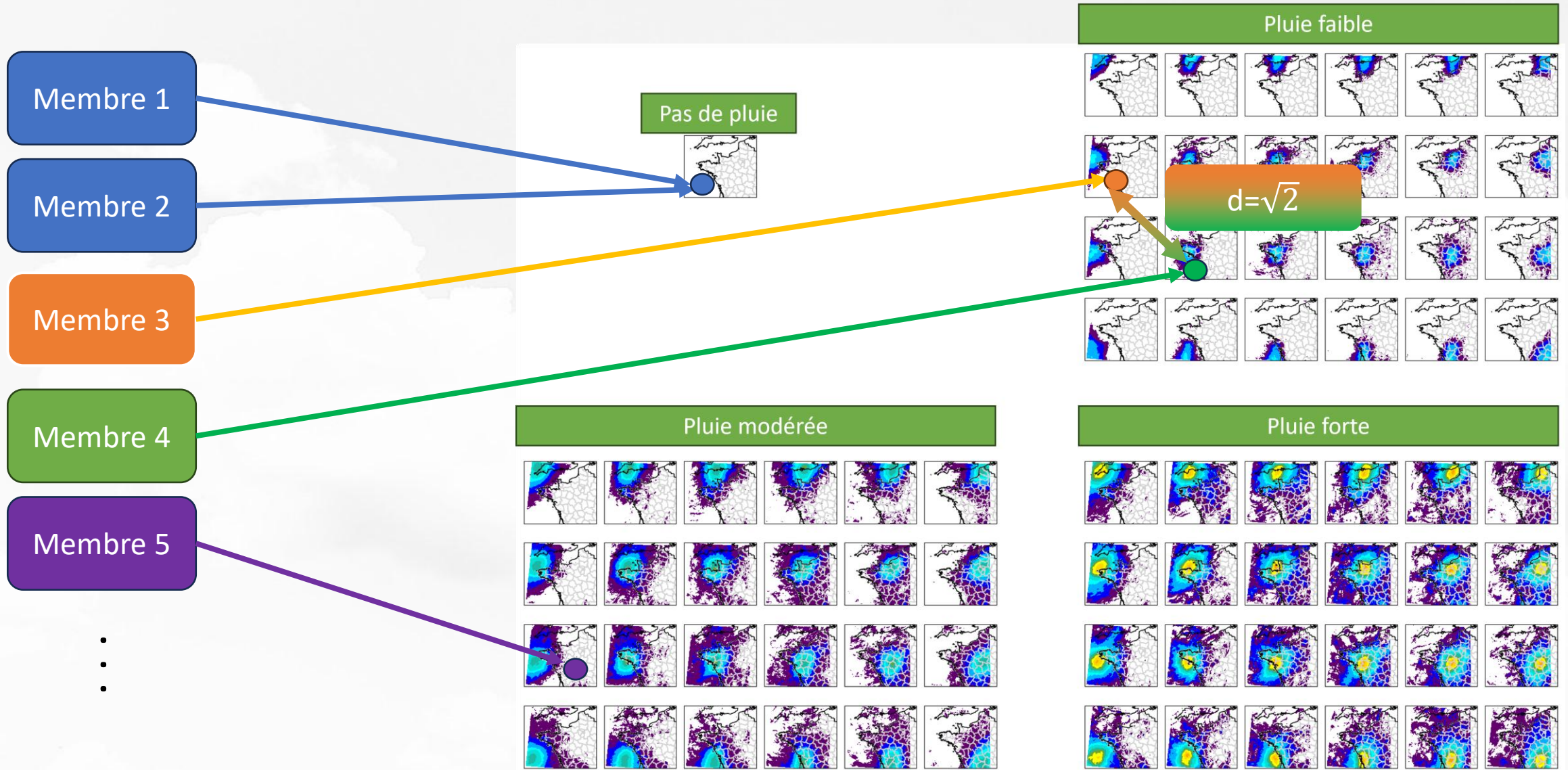
Méthode pour construire les scénarios (étape par étape)

On va (enfin) calculer les distances entre membres



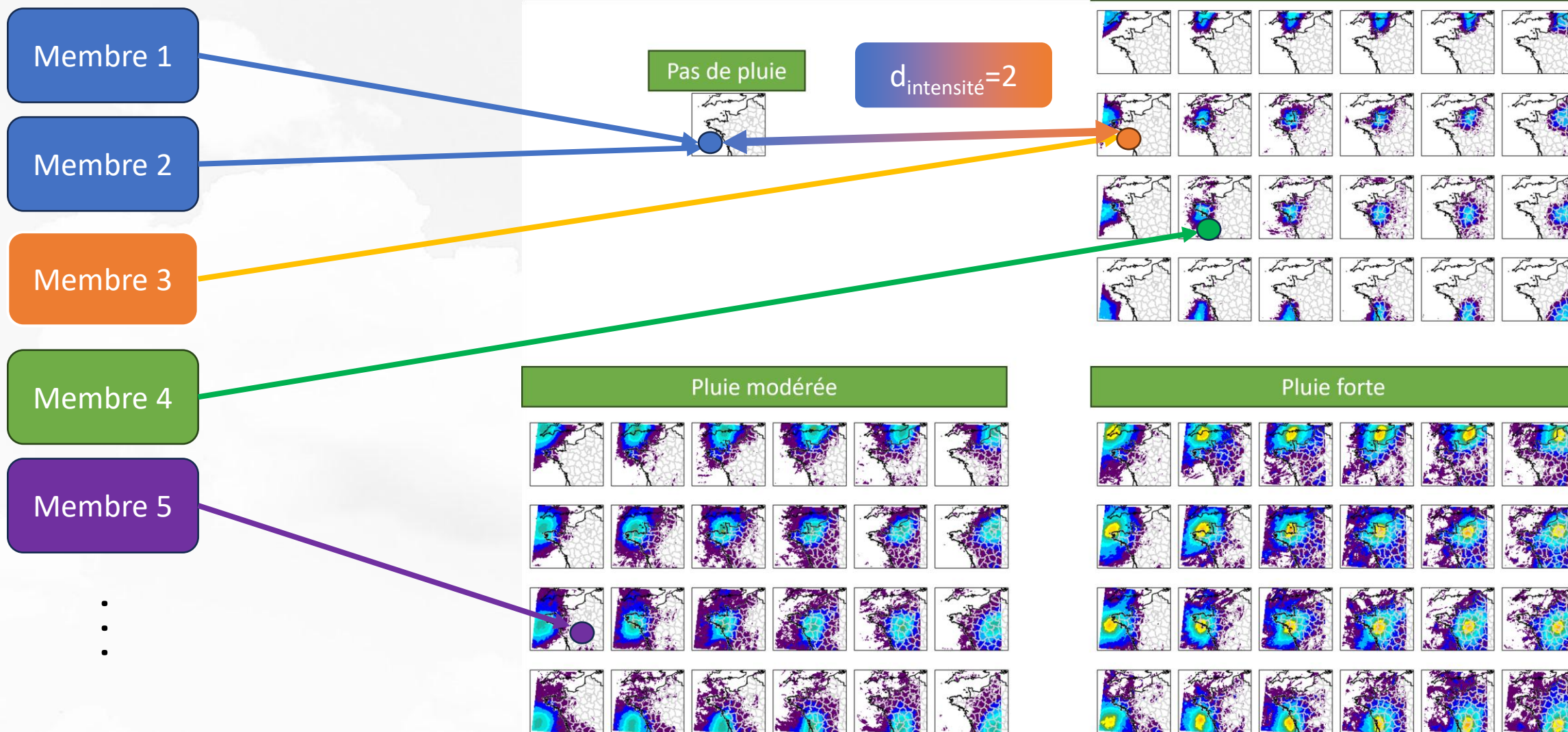
Méthode pour construire les scénarios (étape par étape)

On va (enfin) calculer les distances entre membres



Méthode pour construire les scénarios (étape par étape)

On va (enfin) calculer les distances entre membres



Méthode pour construire les scénarios (étape par étape)

On peut alors calculer la distance moyenne sur une période donnée:

$$d(mb_i, mb_j) = \frac{1}{\Delta t} \sum_1^{\Delta t} d(mb_i(t), mb_j(t))$$

Matrice des distances

| | Mb 1 | Mb 2 | Mb 3 | Mb 4 | Mb 5 | ... |
|----------|------|------|------------|------------|------------|-----|
| Membre 1 | 0 | 0 | 2 | 2 | 4 | |
| Membre 2 | 0 | 0 | 2 | 2 | 4 | |
| Membre 3 | 2 | 2 | 0 | $\sqrt{2}$ | $\sqrt{5}$ | |
| Membre 4 | 2 | 2 | $\sqrt{2}$ | 0 | $\sqrt{5}$ | |
| Membre 5 | 4 | 4 | $\sqrt{5}$ | $\sqrt{5}$ | 0 | |

⋮

Méthode pour construire les scénarios (étape par étape)

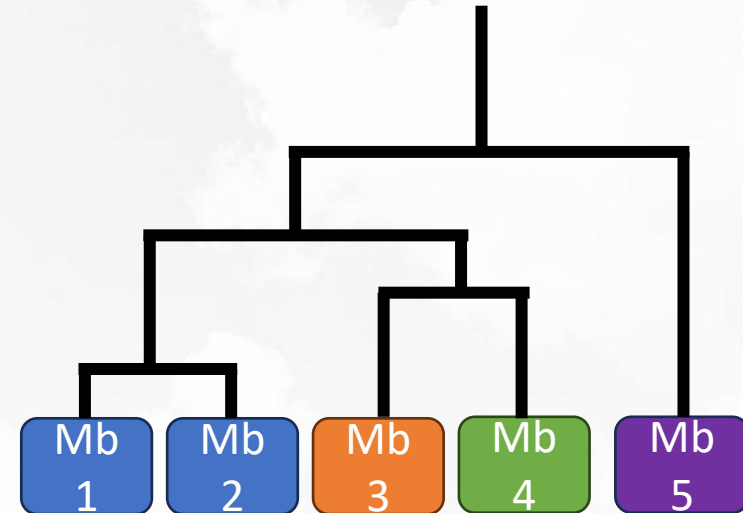
On peut alors calculer la distance moyenne sur une période donnée:

$$d(mb_i, mb_j) = \frac{1}{\Delta t} \sum_1^{\Delta t} d(mb_i(t), mb_j(t))$$

Matrice des distances

| | Mb 1 | Mb 2 | Mb 3 | Mb 4 | Mb 5 |
|----------|------|------|------------|------------|------------|
| Membre 1 | 0 | 0 | 2 | 2 | 4 |
| Membre 2 | 0 | 0 | 2 | 2 | 4 |
| Membre 3 | 2 | 2 | 0 | $\sqrt{2}$ | $\sqrt{5}$ |
| Membre 4 | 2 | 2 | $\sqrt{2}$ | 0 | $\sqrt{5}$ |
| Membre 5 | 4 | 4 | $\sqrt{5}$ | $\sqrt{5}$ | 0 |

Calcul dendrogramme (méthode classification classique)



⋮

Méthode pour construire les scénarios (étape par étape)

On peut alors calculer la distance moyenne sur une période donnée:

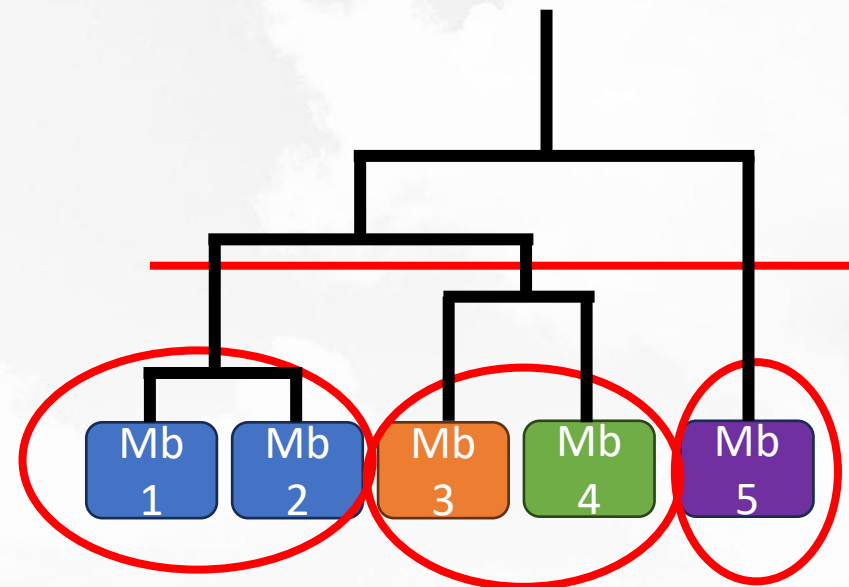
$$d(mb_i, mb_j) = \frac{1}{\Delta t} \sum_1^{\Delta t} d(mb_i(t), mb_j(t))$$

Matrice des distances

| | Mb 1 | Mb 2 | Mb 3 | Mb 4 | Mb 5 |
|----------|------|------|------------|------------|------------|
| Membre 1 | 0 | 0 | 2 | 2 | 4 |
| Membre 2 | 0 | 0 | 2 | 2 | 4 |
| Membre 3 | 2 | 2 | 0 | $\sqrt{2}$ | $\sqrt{5}$ |
| Membre 4 | 2 | 2 | $\sqrt{2}$ | 0 | $\sqrt{5}$ |
| Membre 5 | 4 | 4 | $\sqrt{5}$ | $\sqrt{5}$ | 0 |

...

Calcul dendrogramme (méthode classification classique) + **seuil coupure** :
On obtient 3 scénarios: 1+2, 3+4 et 5



⋮

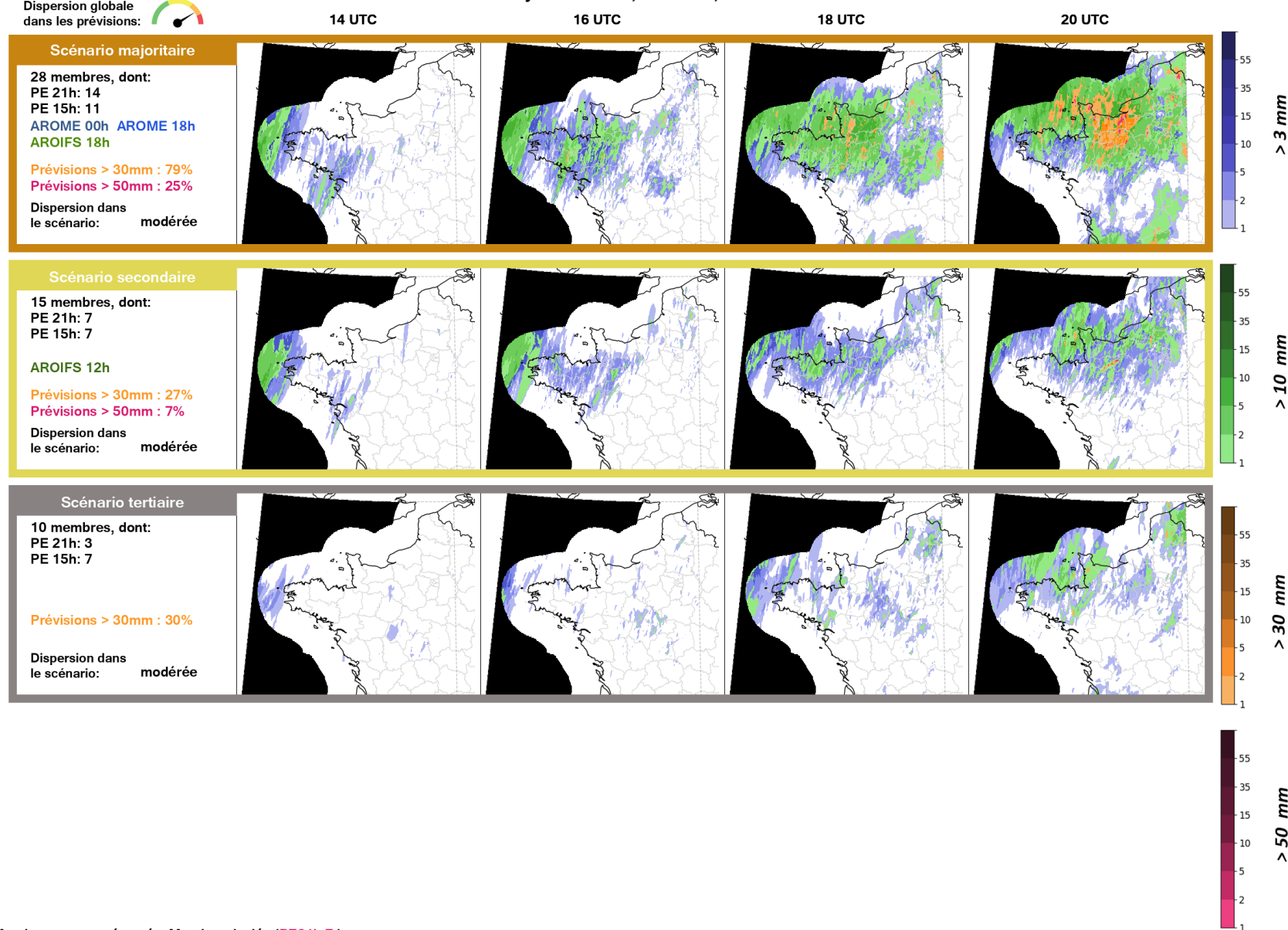
J'espère que vous avez réussi à suivre... On va regarder maintenant ce que ça donne sur un cas réel

Visualisation des scénarios

Issue de longues discussions dans le temps avec PG LABO et des prévisionnistes, il y a beaucoup d'informations:

Dispersion globale dans les prévisions: 

Synthèse RR3, zone NO, du 13/06/2025 - 14Z au 13/06/2025 - 20Z




Nombre d'individus dans un scénario dépassant un seuil de RR3 (en %)

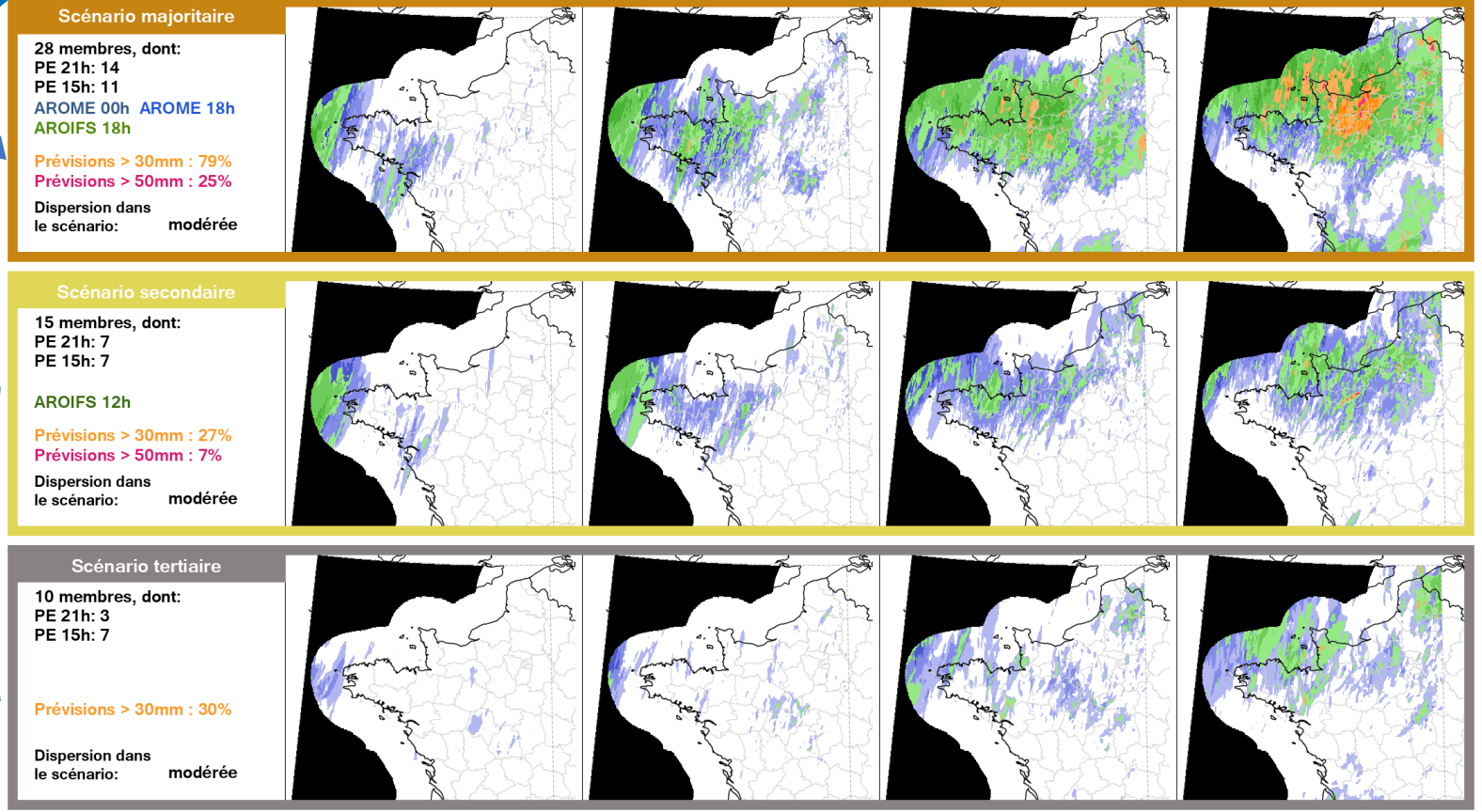
Visualisation des scénarios

Période considérée: 13 juin 2025 de 14h à 20h UTC

Une image toutes les 2h

Scénarios classés du plus peuplé au moins

Dispersion globale dans les prévisions:  Synthèse RR3, zone NO, du 13/06/2025 - 14Z au 13/06/2025 - 20Z



Membres non représentés: Membres isolés (PE21h 7)

Nombre d'individus dans un scénario dépassant un seuil de RR3 (en %)

Visualisation des scénarios

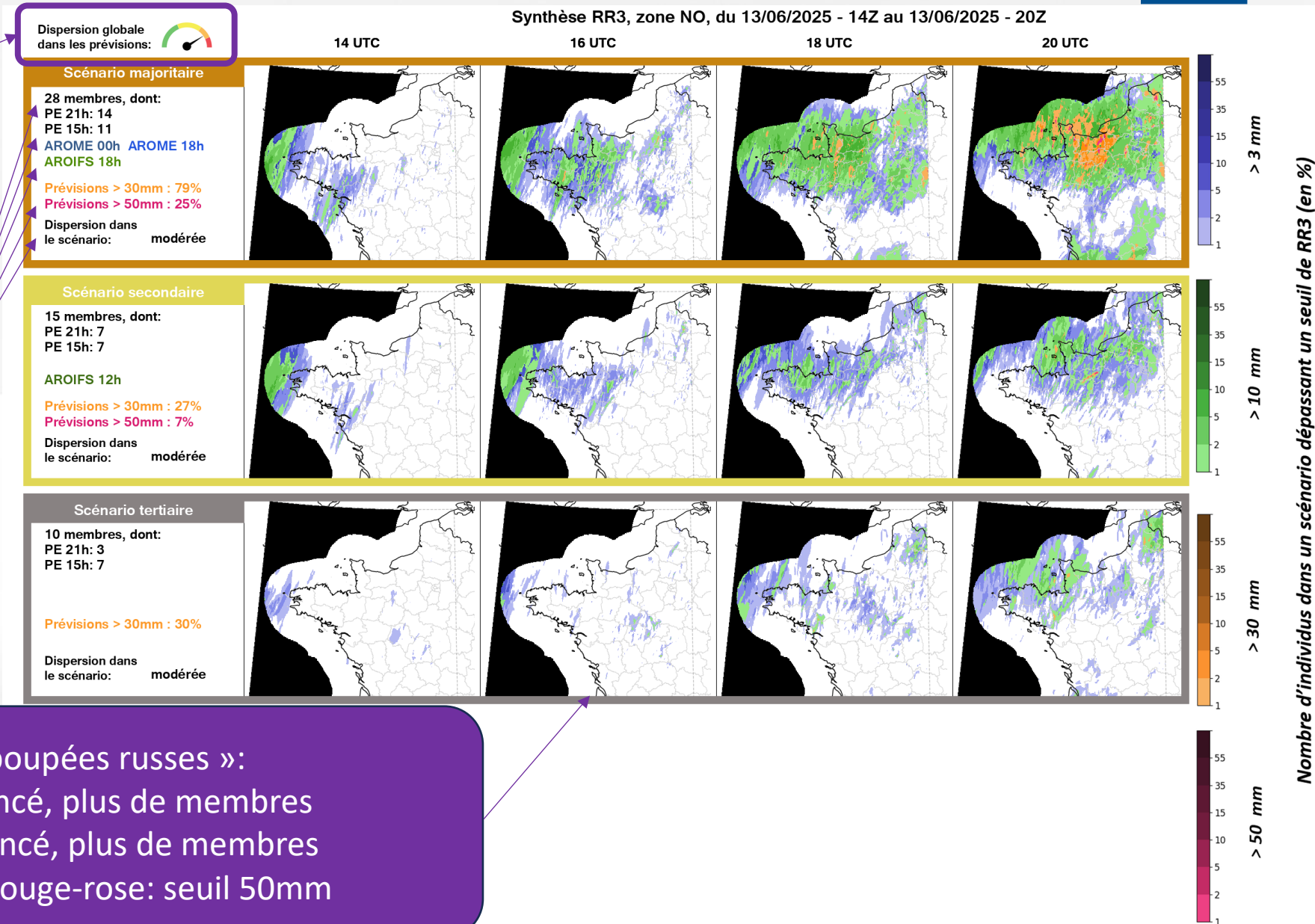
Indicateur global de la dispersion

Informations sur le scénario:

- Nombre total d'individus
- Détail par PE/déterministe
- Pourcentage d'individus dans le scénario dépassant des seuils prédéfinis (sévérité)
- Indicateur dispersion dans le scénario

Cartes avec logique « poupées russes » :

- bleu: seuil 3mm, plus foncé, plus de membres
- vert: seuil 10mm, plus foncé, plus de membres
- orange: seuil 30mm / → rouge-rose: seuil 50mm

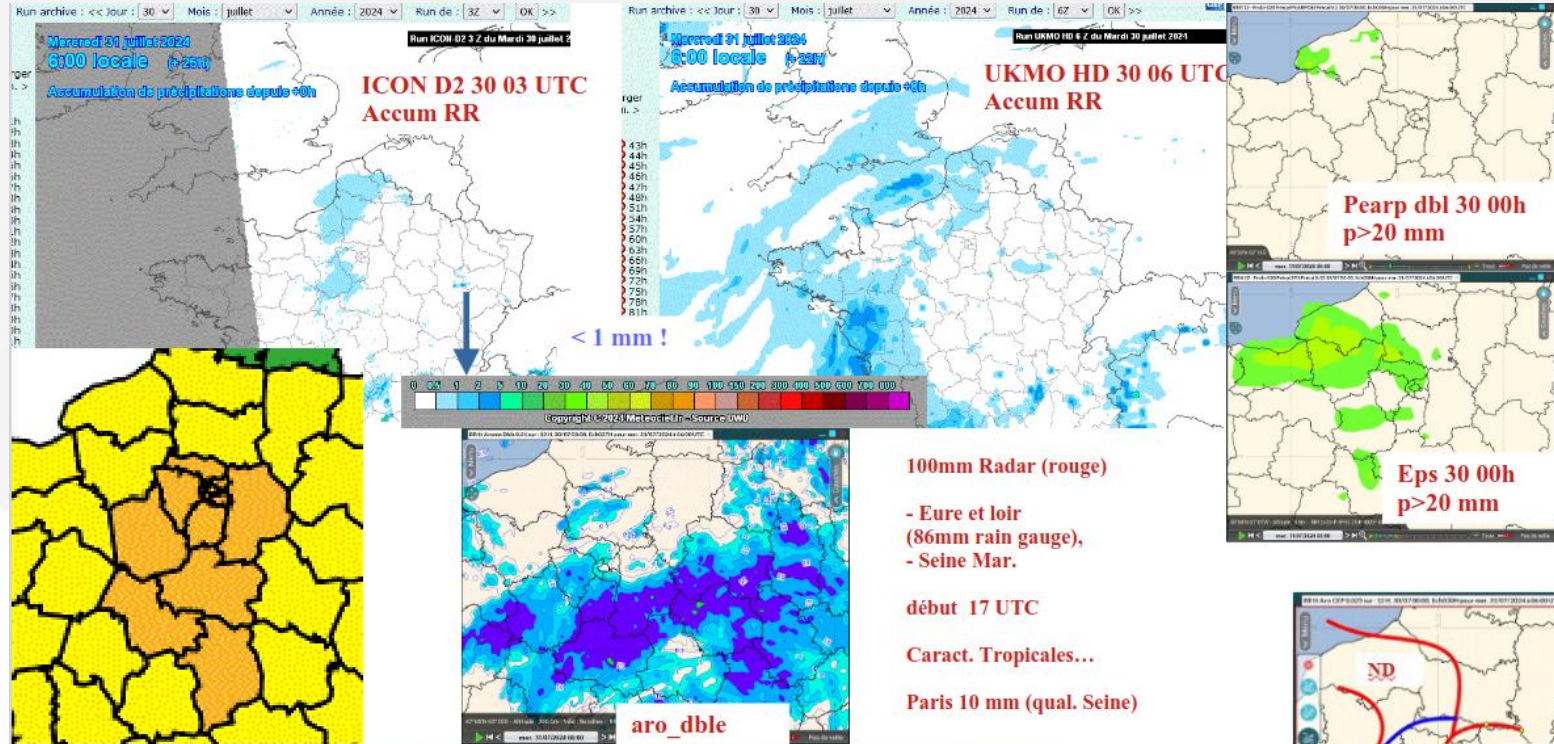


Et sur toutes les échéances d'un réseau PE-AROME ?

Deux exemples:

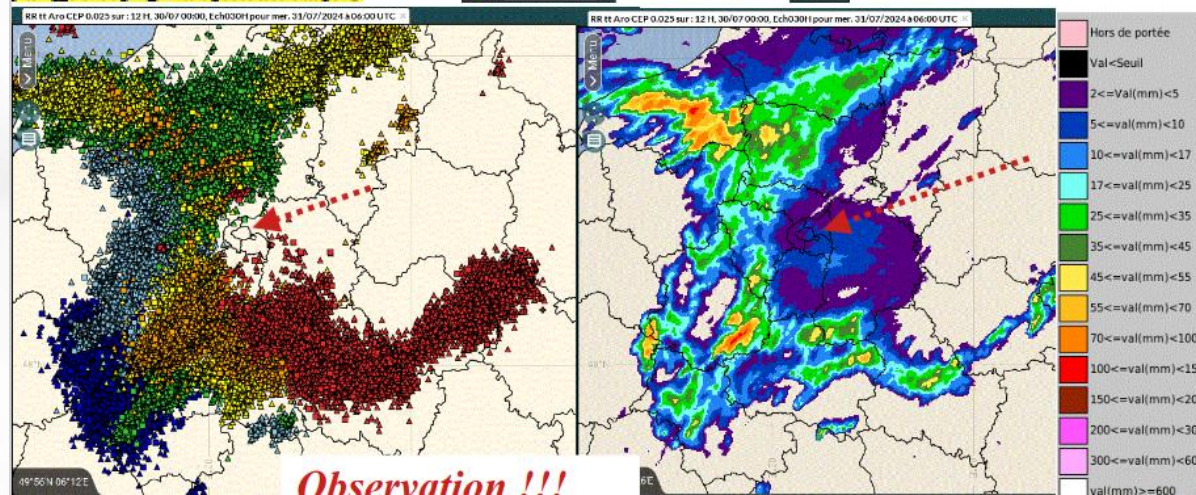
1. Cas « JO » (30 juillet 2024, zone Nord-Ouest/IDF)
2. Nuit et journée du 1^{er} septembre 2025 (zone Sud-Est)

1. Cas « Jeux Olympiques »



PEARO
DBLE ????
Prise de
Décision
délicate !

Slide de Bruno Gillet-Chaulet
 (Chef Prévisionniste National) aux
 ateliers Prévi-GMAP de janvier
 2025



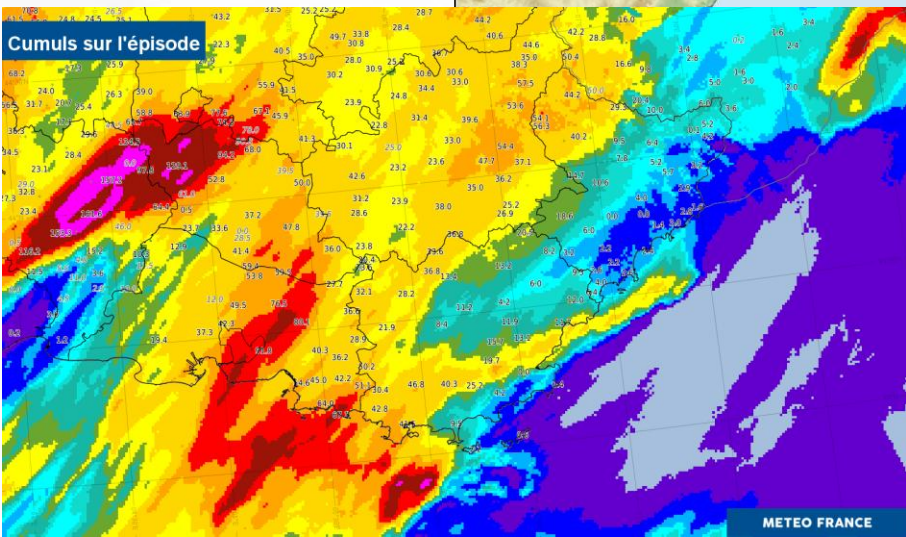
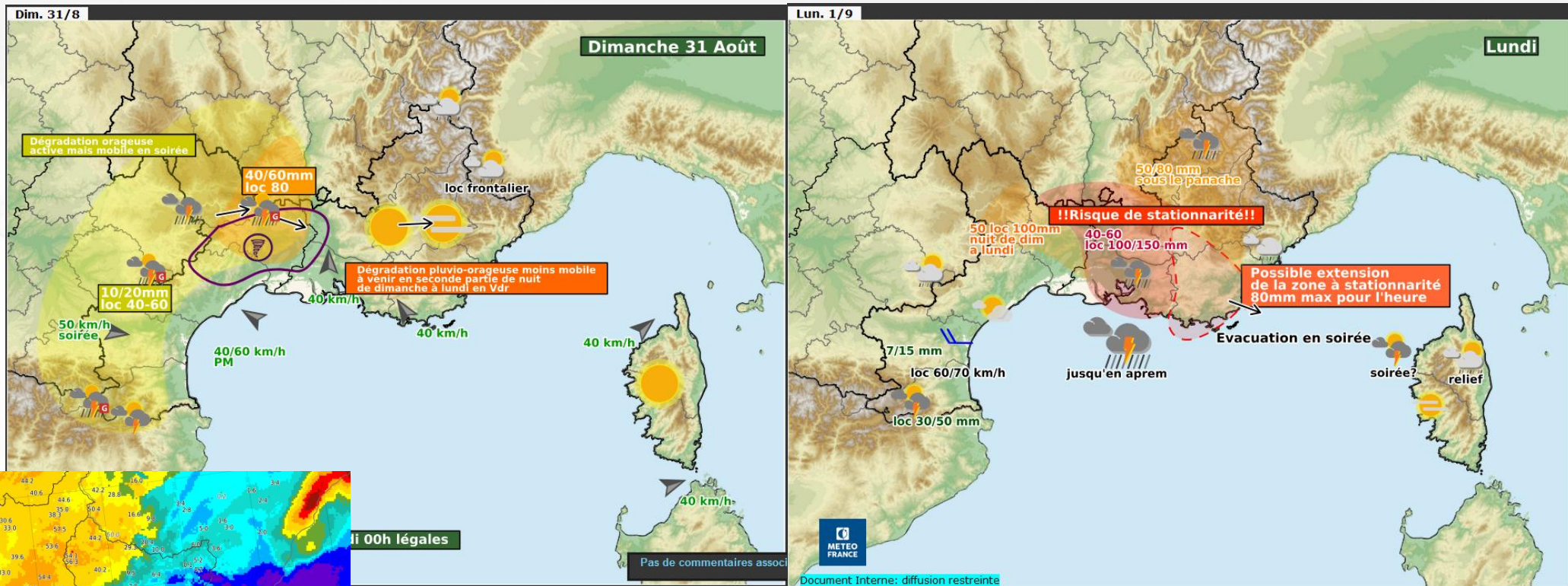
AROME
 - RR ok
 - Conv. Trop précoce
 - RR trop à l'est Convergence GE ?
 Effet Urbain ?
 Et
 Seine Maritime

1. Cas « Jeux Olympiques »

On passe sur le plot interactif !

2. 31 août/1^{er} septembre 2025

Extrait RETEX de la Direction Inter-Régionale Sud-Est (espace confluence)



2. 31 août/1^{er} septembre 2025

On passe sur le plot interactif !

Quelques éléments de conclusion

- Travail sur plusieurs années (avec plusieurs tentatives qui n'ont pas fonctionné). On espère que c'est la bonne !
- La méthode se veut adaptative selon les retours et la plus explicable possible (pourquoi ce regroupement et pas un autre ?)
- L'autoencodeur permet de compresser efficacement les données, on peut aussi orienter les caractéristiques qu'on souhaite plus ou moins garder avec la fonction de coût.
- Etape suivante: peut-être le retour d'un prototype en temps réel prochainement. D'abord sur le Nord-Ouest, puis Sud-Est ?
- Merci à celles et ceux qui ont contribué à ce travail d'une façon ou d'une autre. Les avis extérieurs sont précieux.



Merci pour votre attention ! Questions ?