

Prévision des pluies extrêmes avec la prévision d'ensemble AROME

F. Bouttier et Yingqian Yu - Météo-France CNRM/GMME/PRECIP

AMA mars 2026

Objectif : résumer la prévision d'ensemble lors de risques de pluies extrêmes

Méthode : proposer des **seuils de décision optimaux** (P_{opt}) pour interpréter ces prévisions

Questions :

- Comment construire un **échantillon significatif** de pluies extrêmes, donc rares ?
- P_{opt} dépend-il de la région ? du type de pluie ? est-il temporellement stable ?
- Comment **interpréter physiquement** une prévision probabiliste avec ce type de produit ?

L'utilisation des prévisions d'ensemble: un problème de modélisation de la décision

modélisation **physique**:

- modèle AROME
- perturbations de la PE

membres

modélisation des **incertitudes** :

- identification de l'**enjeu**
- choix du **paramètre** impactant
- sélection de l'**échelle d'utilisation**

distributions
de probabilité
P

modélisation des **impacts** :

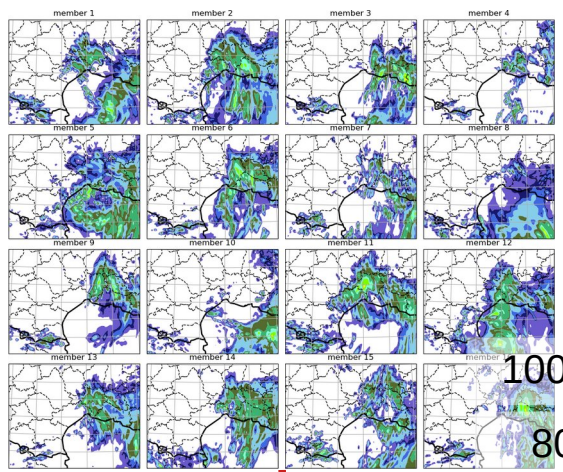
- taux de **fausses alarmes** admissible = seuil P_{opt}
- valeur pratique de cette décision = son **score**
- **confiance** dans le système = recul historique
- **explicabilité** de la décision = communication

décision: si $P > P_{opt}$
prévoir oui/non/attendre

Calcul de P_{opt} : algorithme de Bouttier et Marchal (2024)

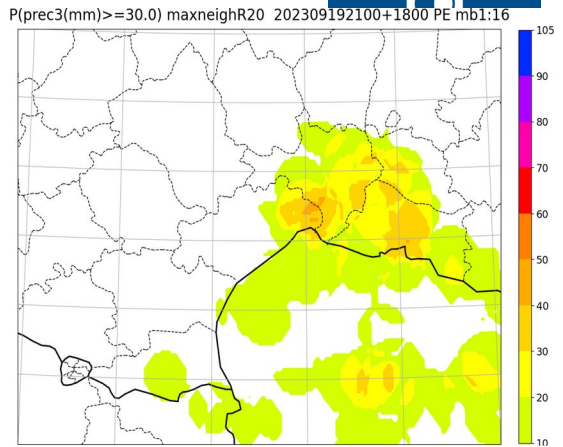
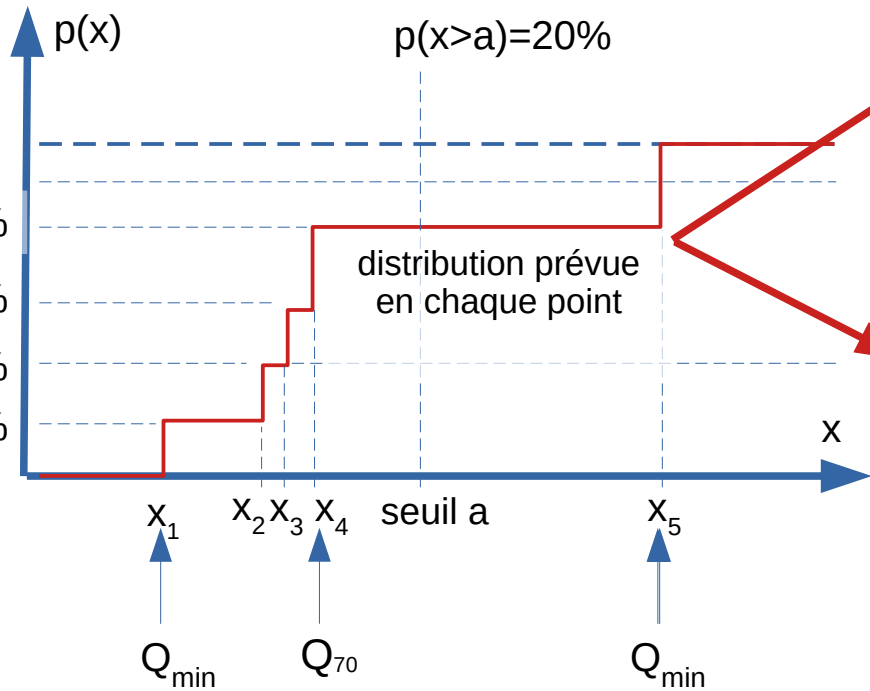


ensemble de prévisions
PE AROME
de rr3 (pluie sur 3h)

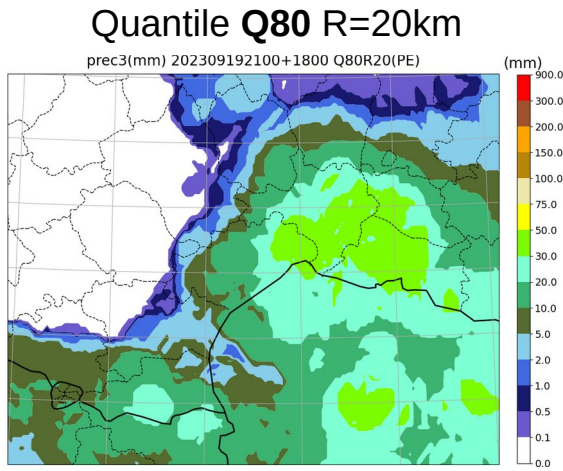


- voisinage max dans chaque membre = modélise les erreurs de position des objets extrêmes
- visu: proba ou quantile de la distribution de probabilité prévue en chaque point (~ enveloppe "raisonnable" de l'ensemble)

voisinage max
& construction de PDF
prévue



proba $P(rr3>30mm)$
rayon $R=20km$



Quantile **Q80** $R=20km$



Quel seuil de probabilité utiliser ?

→ Optimisation de P_{opt} par apprentissage statistique

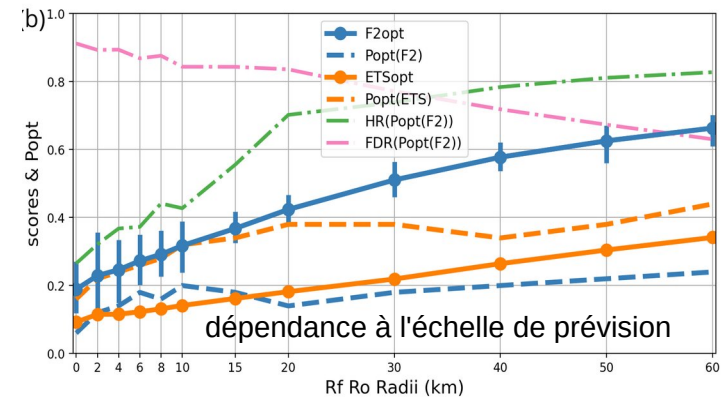
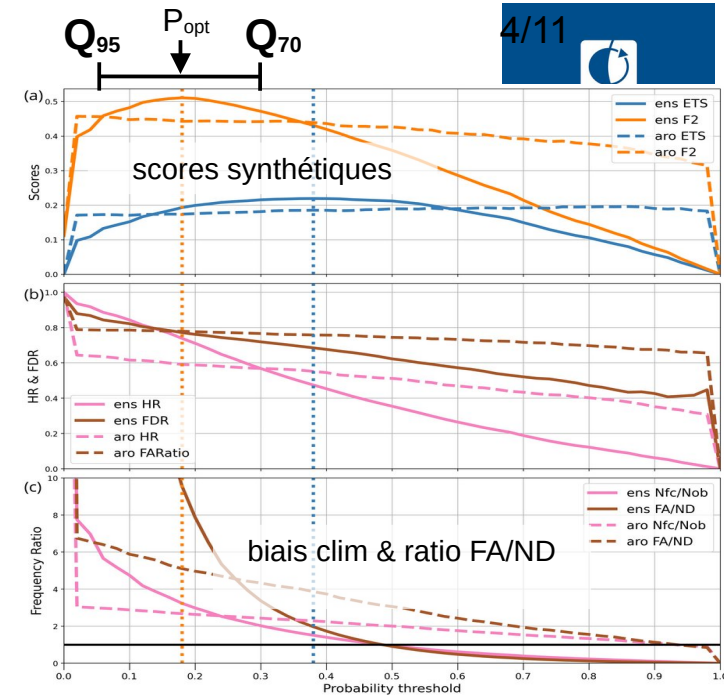
On optimise P_{opt} sur une sélection de cas passés.

Equivalent à un **quantile** $Q_{opt} = 1 - P_{opt}$

Ex: pluies sur 6h avec un échantillon de 1 an :

- $Q_{opt} = Q_{82}$ pour les **fortes pluies (rr6~30mm) à résolution 30km** (si on vise 4x plus de fausses alarmes FA que de non-détections ND)
- alors le taux de détection **TD~80%**, **FA~80%**, biais fréquentiel fc/ob = 3
- **tous les quantiles entre Q70 et Q95** sont aussi quasi-optimaux
- résultat **sensible aux échelles d'utilisation** spatiale et temporelle : **prévisions très dégradées sous 20km et 3h**
- P_{opt} dépend un peu de l'**intensité** au-delà de 45mm/6h.
- très peu de l'échéance si <36h, de la saison

*Problème : avec 1 an de données on ne peut traiter que des intensités $rr6 \leq 30mm$ environ -> **besoin d'un échantillon plus long***



Construction d'un échantillon de pluies extrêmes avec toute l'archive opérationnelle PE AROME

Indice de sévérité : $Idx(rr6) = Max_{quotidien} (Q_{80ensemble} (Q_{99spatial} (rr6)))$
 Sélection des jours avec le plus grand Idx :

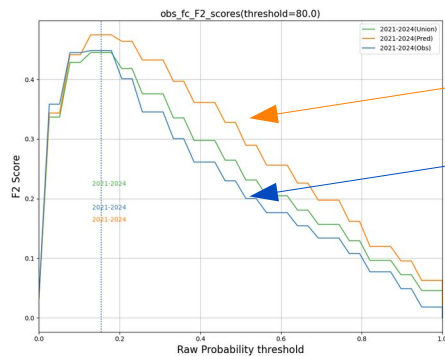
- **observé**, $Q_{99}=74\text{mm}$
 - **ou prévu** par PE Arome, $Q_{99}=66\text{mm}$
- **83 jours sur 5 ans ("orange" ou "rouge")**

Résultat :

NB les scores sont **meilleurs sur les jours à pluies prévues qu'à pluie observée** (= biais de sélection)

mais l'impact de ce biais est négligeable sur P_{opt}

-> prépare l'utilisation de la réanalyse ARRA sur 60 ans



score sur cas prévus

score sur cas observés

PEARO
2020

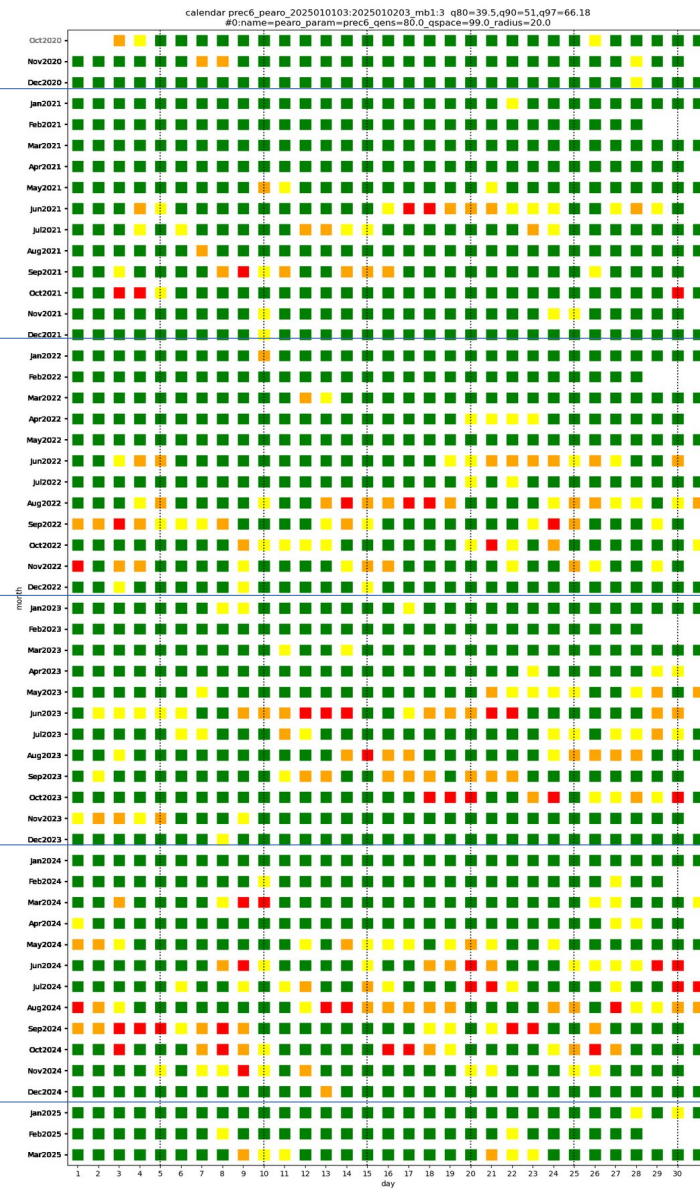
2021

2022

2023

2024

2025



Question : faut-il un modèle de décision spécifique aux événements méditerranéens ?

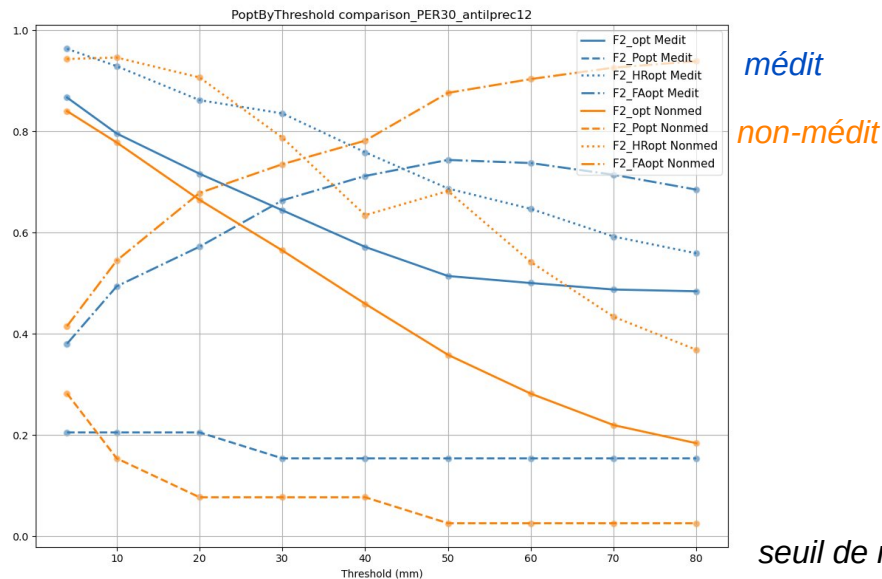
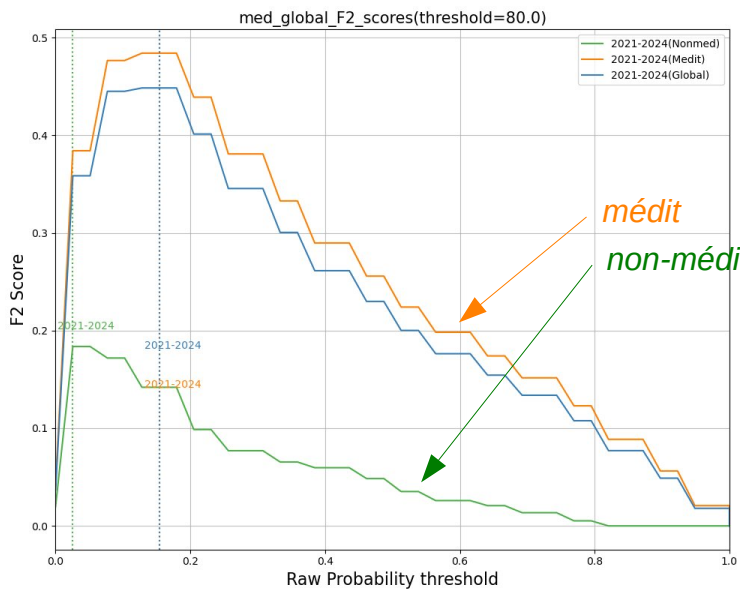
Méthode : identification subjective des cas "médit". Ils constituent la majorité des cas de HPE.

Résultats : sur les cas médit, par rapport aux non-médit,

- P_{opt} est plus élevé de ~10%
- les scores sont meilleurs

hyp: orages orographiques plus prévisibles que les orages de plaine ?

- P_{opt} s'extrapole mieux aux intensités extrêmes



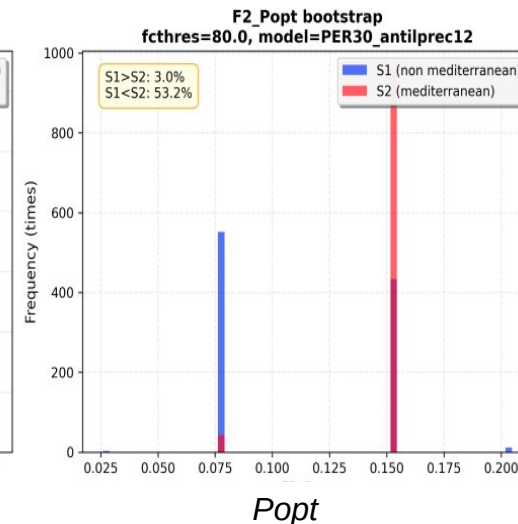
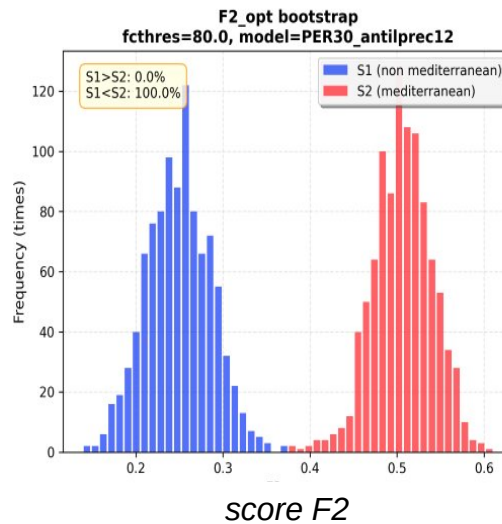
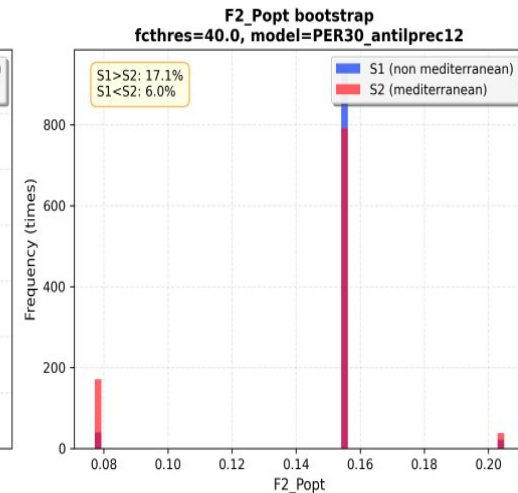
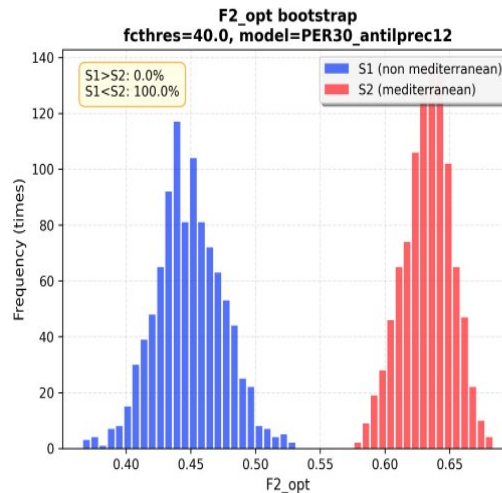
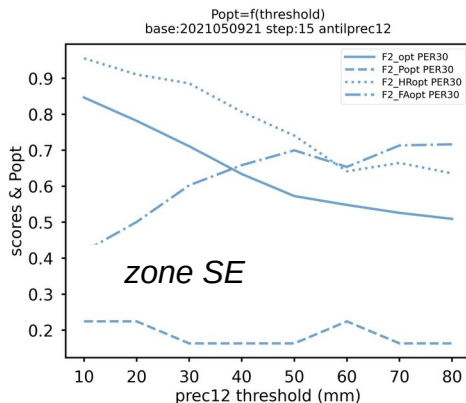
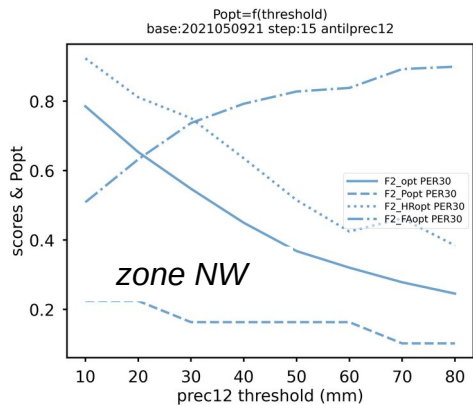
Spécificités des pluies méditerranéennes

Question : les différences médit/non-médit se résument-elles à des différences géographiques SE/NW ?

Méthode : on coupe la France en 2 en **isolant le coin SE** (2E,45N) et on compare les Popt de ces 2 parties sur tout l'échantillon (jours médit + non-médit)

Résultats :

- **les prévisions restent meilleures dans le SE** que le NW (différences robustes de score F2)
- **P_{opt} reste plus élevé dans le SE**, mais cette différence de P_{opt} est peu robuste (le test du bootstrap échoue)



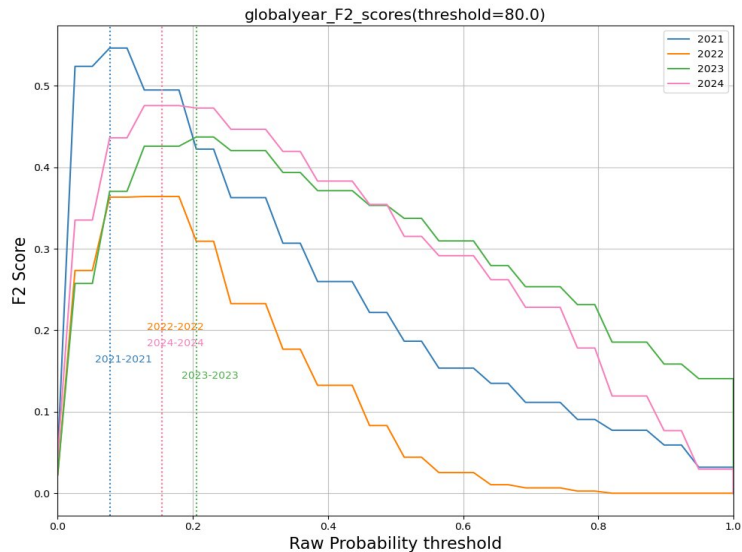
Question: les P_{opt} sont-ils stables d'une année sur l'autre ?

But : comprendre si les statistiques de PE peuvent être extrapolées hors de l'échantillon d'apprentissage

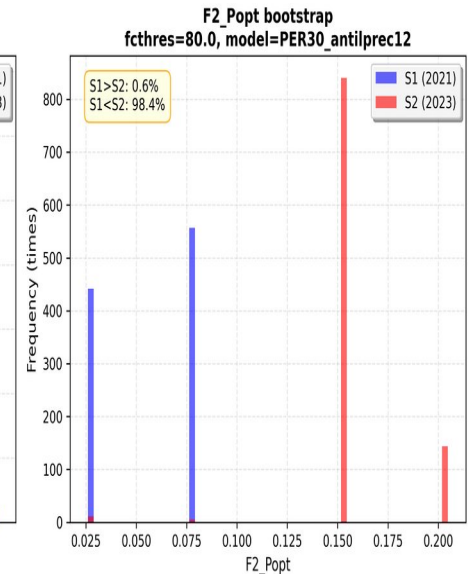
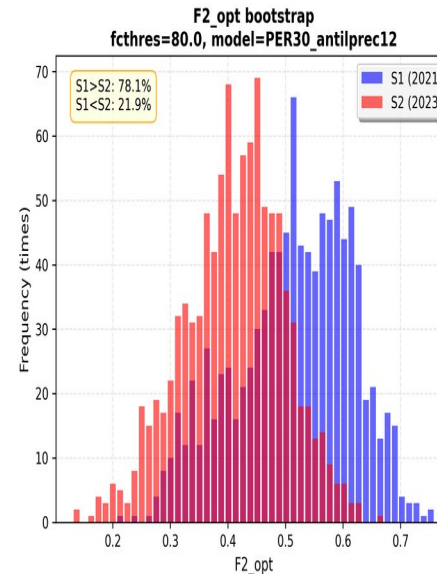
Méthode : calcul de P_{opt} sur les HPE de 4 années

Résultats sur la variabilité interannuelle :

- P_{opt} varie beaucoup : Q_{opt} entre Q78 et Q90 selon l'année
 - P_{opt} était beaucoup plus faible en 2021 qu'en 2022, 2023, 2024
- hyp : 2021 avait plus d'orages de plaine et moins de HPE orographiques (à confirmer !)



seuil
 $rr12 > 80mm$



Questions: les décisions fondées sur un P_{opt} moyen sont elles acceptables au quotidien ?



Méthode : comparer le P_{opt} du jour au P_{opt} annuel pour les fortes pluies

Résultats : sur 18 jours en 2022, selon le score F2,

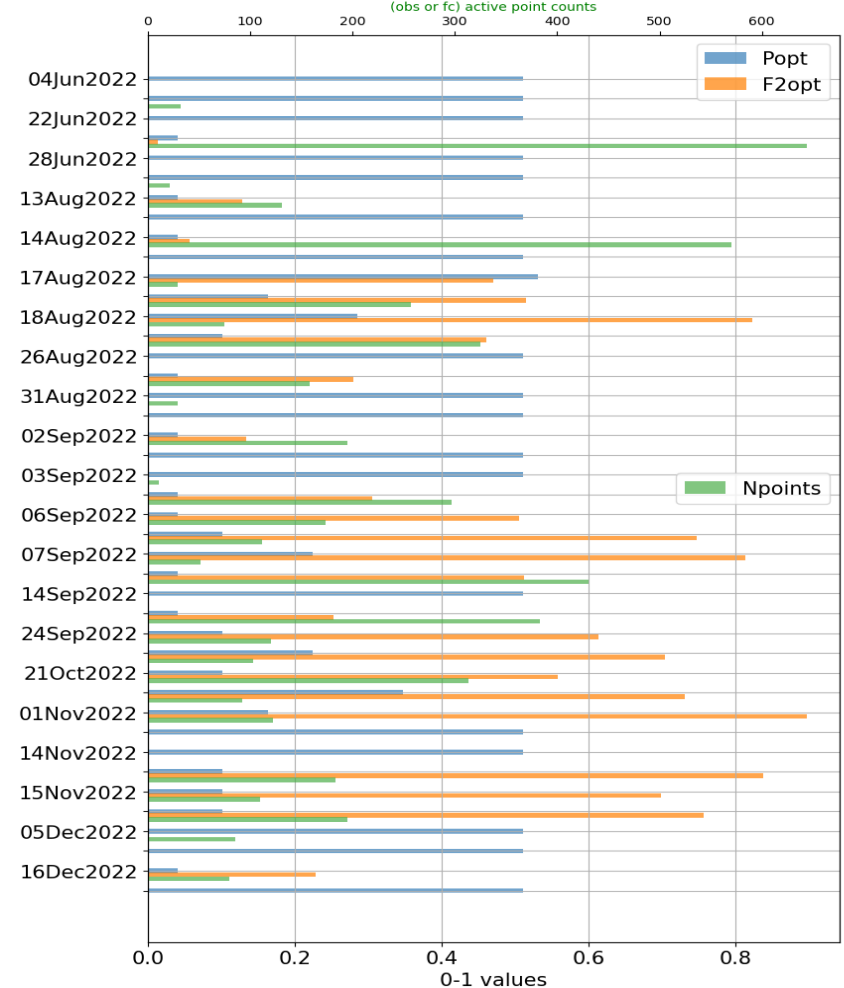
- 7/36 cas de "prévisions ratées"
- 17/36 "prévisions acceptables" (surtout des cas Médit)
- 7/36 "prévis optimales grâce à Popt"

et aussi :

- 11/36 de cas sont optimaux avec $Q75 < Q < Q_{max}$
- les prévis avec $Q < Q75$ ne sont presque jamais acceptables

en bref : sur 50% des cas temps, le Popt annuel moyen aurait permis de "bonnes" prévisions.

Popt timeseries prec12 thres=80.0(mm) steps(h)=(15, 27, 12) Popt=0.163
(obs or fc) active point counts



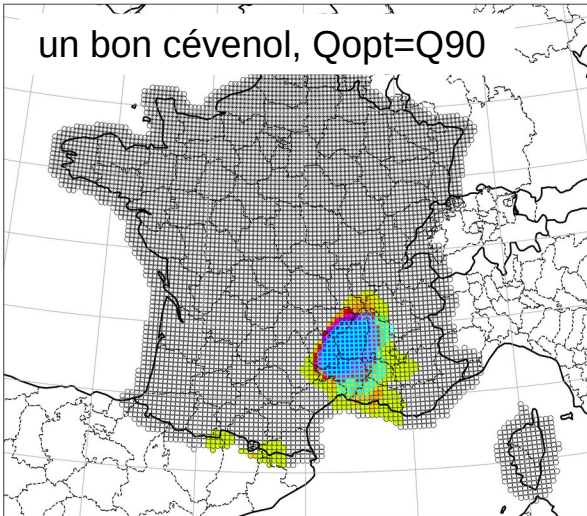
Variabilité quotidienne des P_{opt} - exemples

cartes de probas prévues
& obs de $rr12 > 80mm$



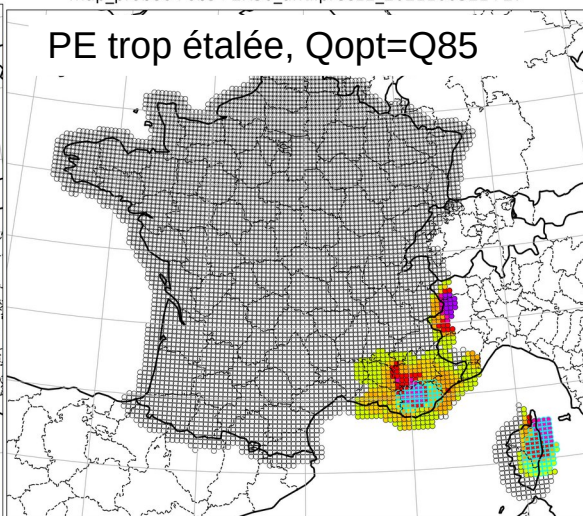
map_prob80+obs PER30_antilprec12_2021100221+27

un bon cévenol, $Q_{opt}=Q90$



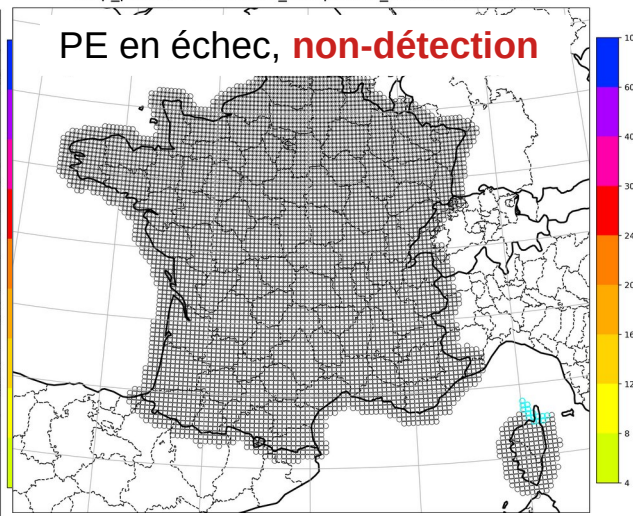
map_prob80+obs PER30_antilprec12_2021100321+27

PE trop étalée, $Q_{opt}=Q85$



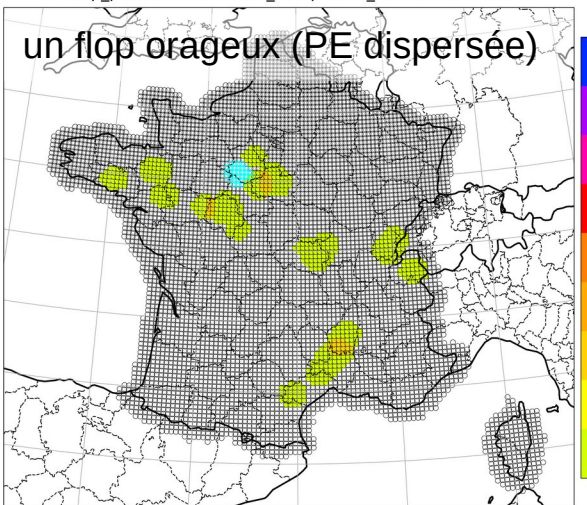
map_prob80+obs PER30_antilprec12_2021111321+15

PE en échec, **non-détection**



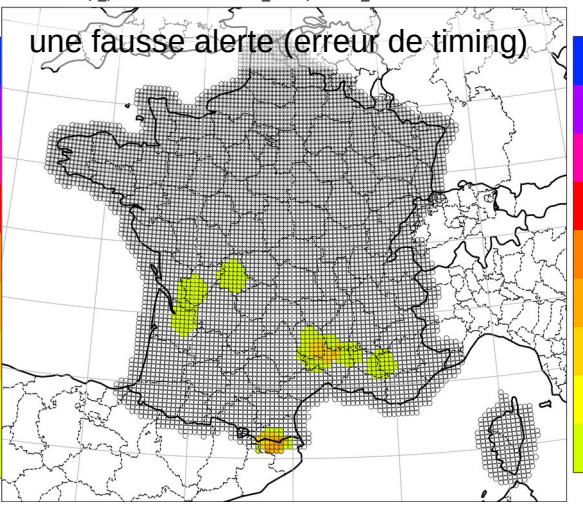
map_prob80+obs PER30_antilprec12_2022062121+27

un flop orageux (PE dispersée)

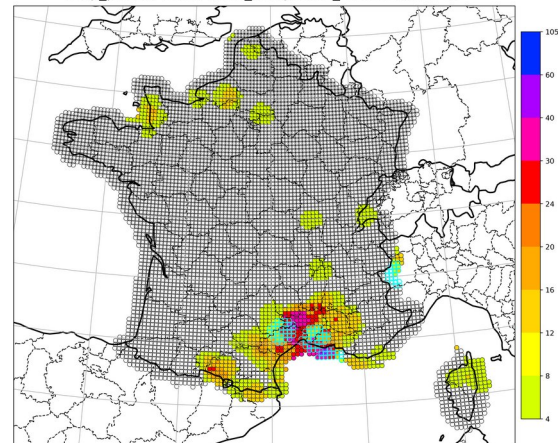


map_prob80+obs PER30_antilprec12_2022081221+27

une fausse alerte (erreur de timing)



map_prob80+obs PER30_antilprec12_2022081621+27



Conclusions :



Sur la **construction de l'échantillon** d'apprentissage :

- 5 ans d'archive conviennent pour des critères de type vigilance orange
- il faut s'appuyer sur **prévisions et observations** pour éviter les biais de sélection

Sur la **variabilité géographique** de P_{opt} :

- la règle de décision optimale diffère entre les cas méditerranéens et les autres

Sur la **variabilité temporelle** de P_{opt} :

- variabilité significative d'une année sur l'autre
 - grande variabilité quotidienne
- besoin de mieux modéliser l'influence du régime météo dans la prédiction de P_{opt}

Perspectives (projet IRICLIM) :

- étendre l'échantillon avec la réanalyse ARRA, pour enrichir l'algorithme avec plus de physique
- agrandir l'ensemble (multimodèle, voisinage temporel)
- associer des scénarios physiques à P_{opt} , pour l'interprétation humaine