

Prévision et assimilation d'ensemble

(méthodes Monte Carlo, quantification des incertitudes, propagation des erreurs, modélisation stochastique, logique floue,...)

Applications en météo & climat :

- prévision probabiliste, génération d'alertes
- covariances d'erreur B pour l'assimilation
- (science: études de prévisibilité, de sensibilité)

= environ 60% des ressources calcul des centres opérationnels

A cause du chaos et de la rareté des observations, les incertitudes souvent aussi grandes que les phénomènes eux-mêmes

horizon de prévisibilité : limites sur

- * l'échéance de prévision
- * son niveau de détail : spatial, temporel, intensité

Prévision et assimilation d'ensemble

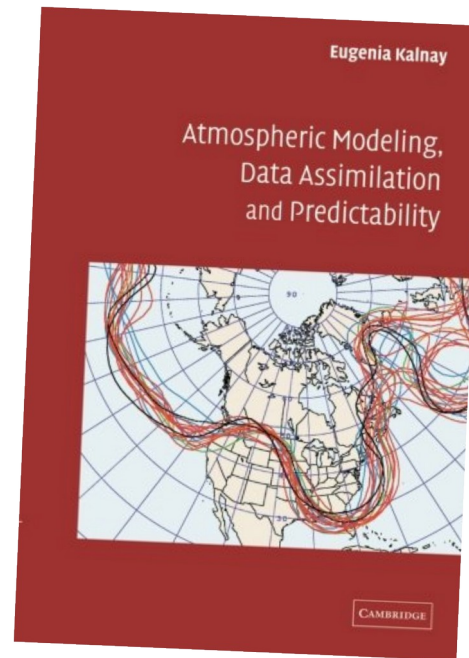
- principe de la PE
- ensembles en assimilation de données
- construction des perturbations
- métriques d'évaluation
- produits



Eugenia Kalnay

(Argentine, 1942-2024)

1ères prévisions d'ensemble en
météo (1993)



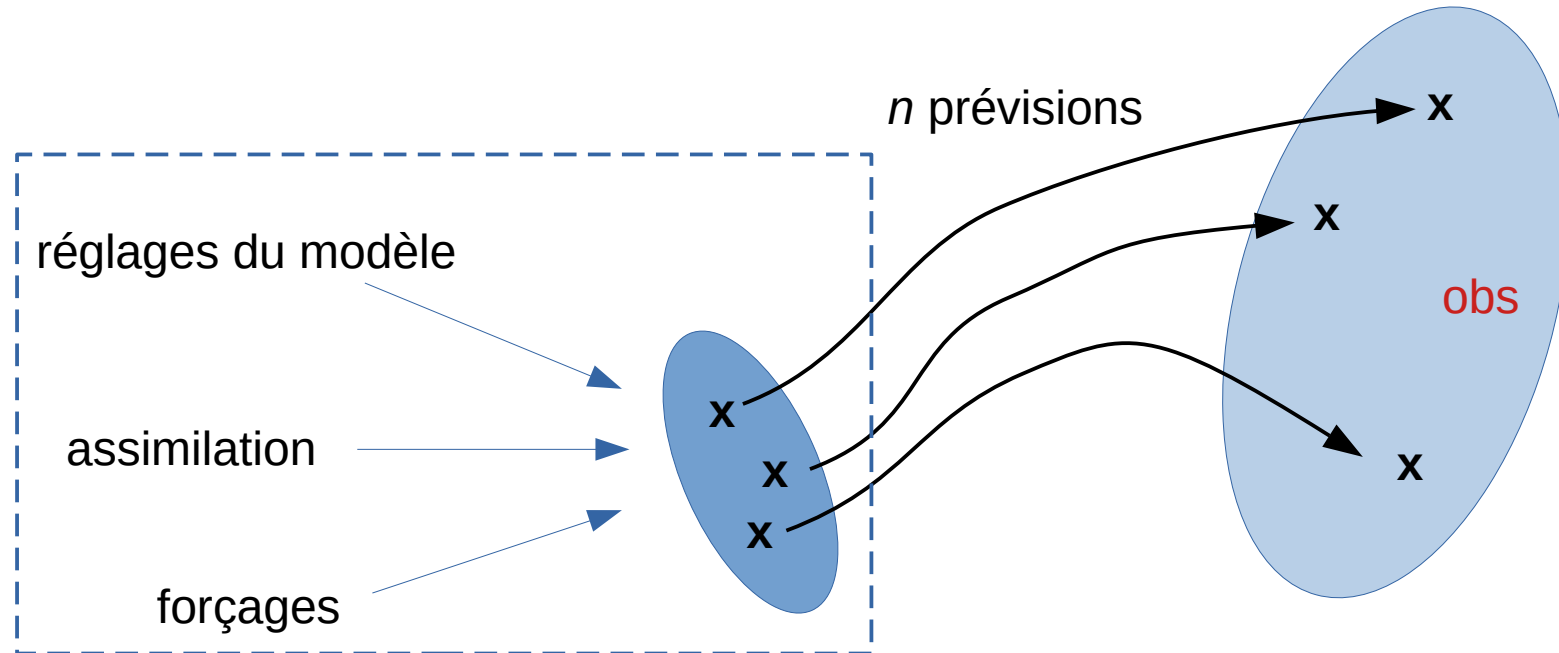
Edward Lorenz,
(MIT, 1917-2008)

théorie du chaos

déf : un **ensemble** est une *représentation discrète* des *incertitudes* en temps réel sur *notre connaissance d'un système physique*

- "discrète" = on approxime une distribution par un 'peigne' (= somme de distributions de Dirac)
- "incertitudes" = éventail des états plausibles
- "notre connaissance" = les informations dont on dispose à un instant t

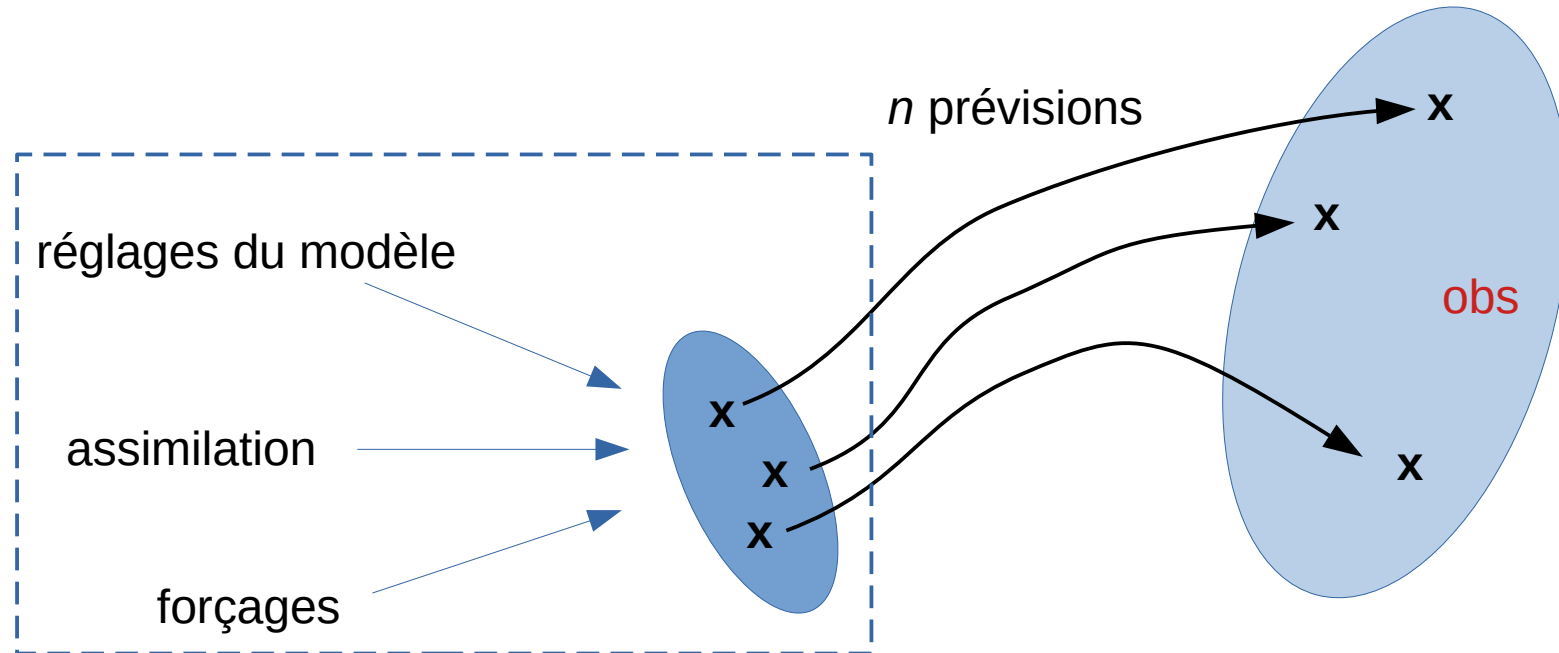
Une prévision d'ensemble caractérise un **système de prévision** (que nous avons construit) et non une réalité physique immanente



déf : un **ensemble** est une *représentation discrète* des *incertitudes* en temps réel sur *notre connaissance d'un système physique*

- "discrète" = on approxime une distribution par un 'peigne' (= somme de distributions de Dirac)
- "incertitudes" = éventail des états plausibles
- "notre connaissance" = les informations dont on dispose à un instant t

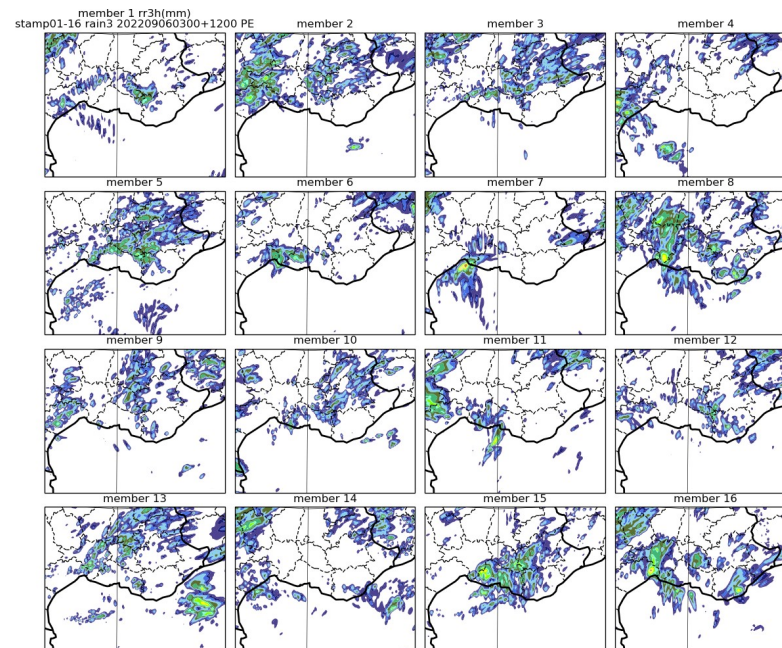
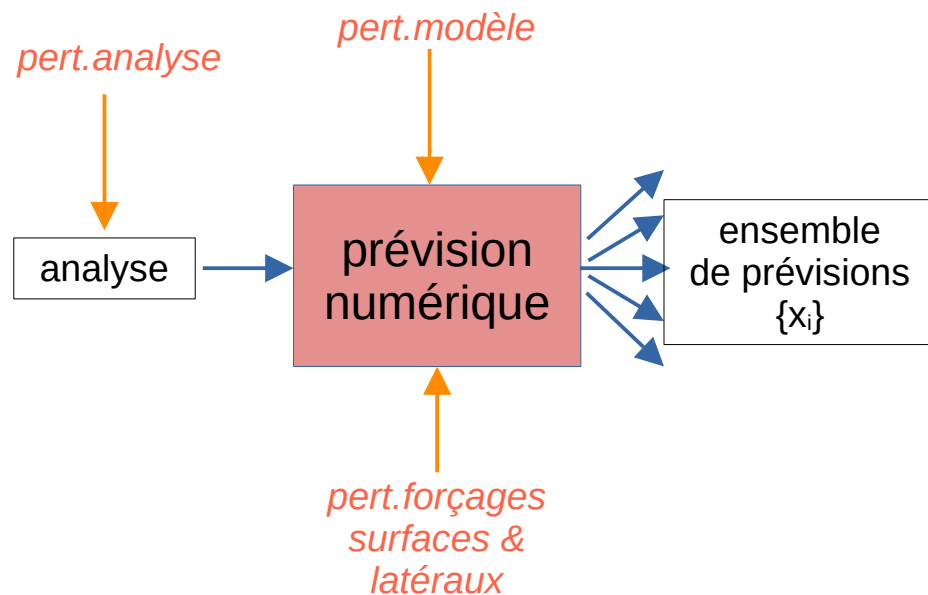
Une prévision d'ensemble caractérise un **système de prévision** (que nous avons construit) et non une réalité physique immanente



objectif : créer une dispersion des prévisions, représentative des incertitudes

principe :

- injecter des *perturbations aléatoires* dans *toutes les entrées incertaines* du système (d'analyse ou de prévision)
- laisser les perturbations mûrir : création d'*équilibres physiques*, *amplification chaotique*



Prévision d'ensemble

On lance q prévisions perturbées (**membres**) en parallèle ($q \sim 50$)

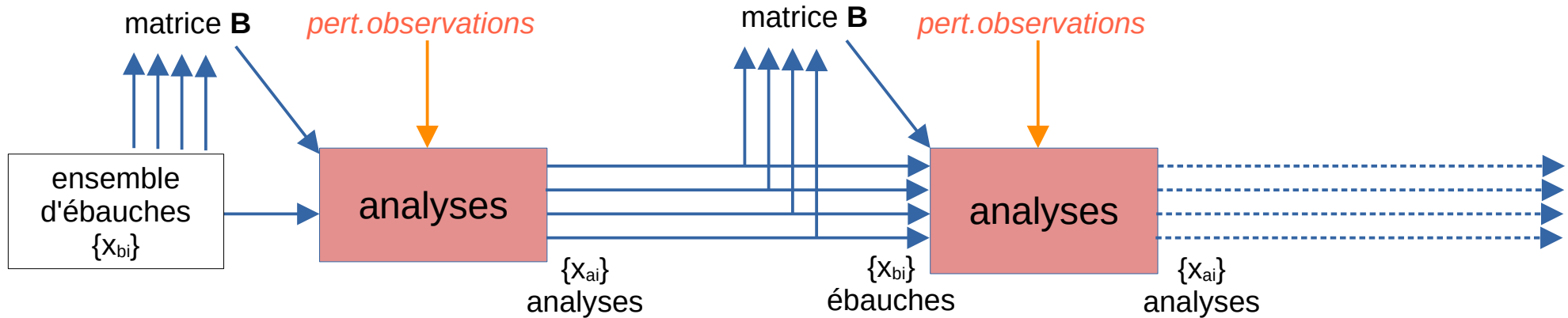
Approche naïve: en factorisant $A = L L^T$, si l'on tire aléatoirement u suivant la loi Gaussienne, alors la prévision perturbée $v = x_a + L u$ suit la loi de probabilité d'analyse impliquée par A (matrice de covariances d'erreurs d'analyse)

donc $M(v)$ suit approximativement la loi de probabilité des erreurs de prévision

Le calcul d'un ensemble $M(v_i)$, $i=1 \dots q$ coûte q fois la prévision d'un modèle.

cycles de prévision : comme la prévision d'ensemble

*cycles d'analyse : perturbation des **observations***



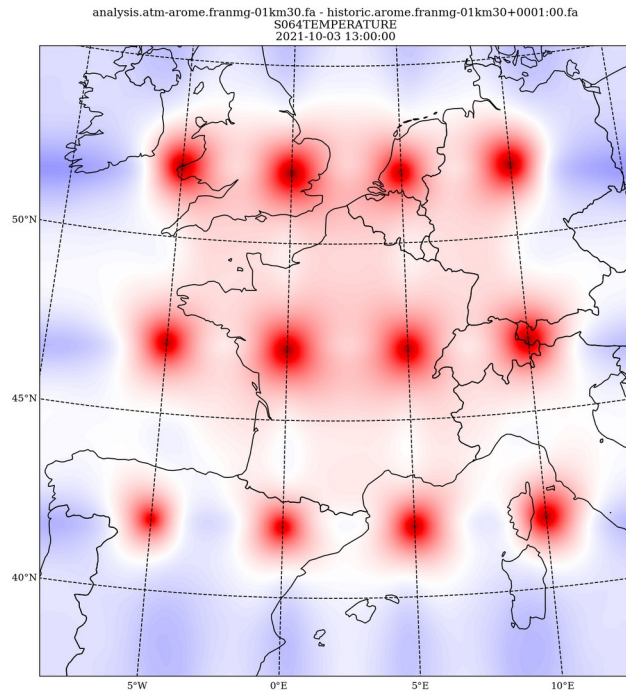
Applications des prévisions d'ensemble : relation avec l'assimilation de données

8/60

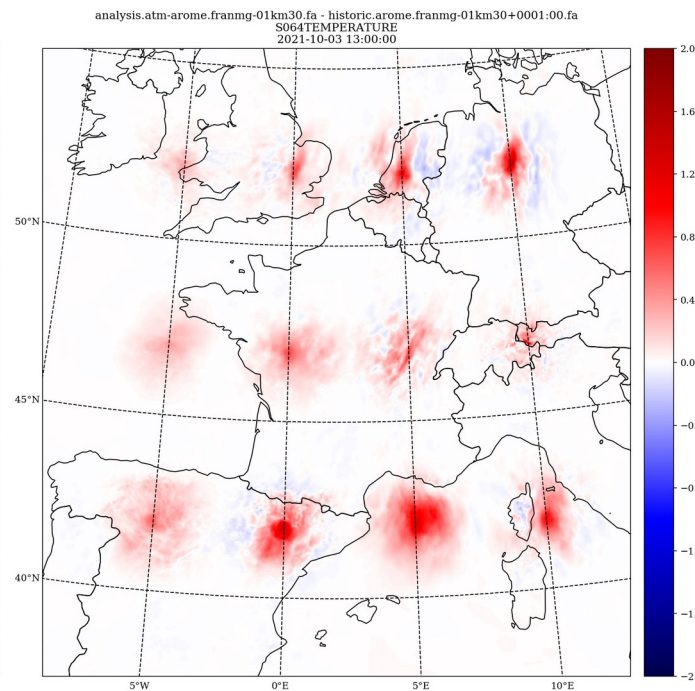
fig : 3DenVar, (P. Brousseau et V. Vogt, réunion Prévi/GMAP janvier 2022)

assim d'ensemble AE Arome 50 membres $dx=3,2km$, Jb hybride

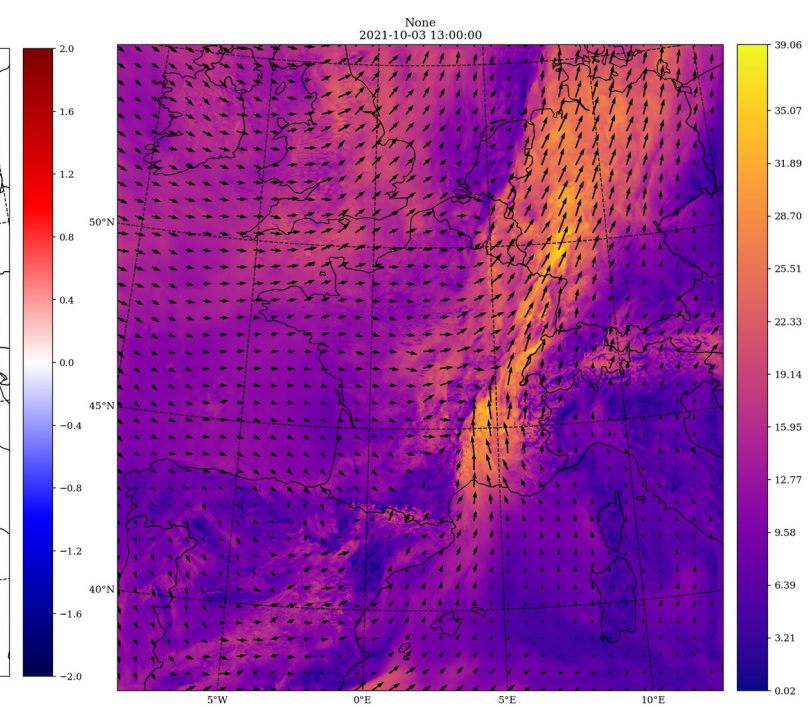
cov (T850,T850) du Jb 3DVar



cov du Jb 3DenVar



vent à 850hPa



Ensembles en assimilation de données

Deux objectifs :

- obtenir une **dispersion correcte en moyenne (fiabilité)**
- obtenir un bon signal = une dispersion qui **dépende de la situation (résolution)**

En théorie, les perturbations peuvent être définies automatiquement (filtres de Kalman, filtres particuliers, machine learning)

En pratique, on n'a jamais assez de données et de puissance calcul, donc :

- définir les perturbations avec du **sens physique**
- **optimiser des scores** sur des périodes-test
- perturber "n'importe comment" peut même produire un bon ensemble (avec les systèmes chaotiques)

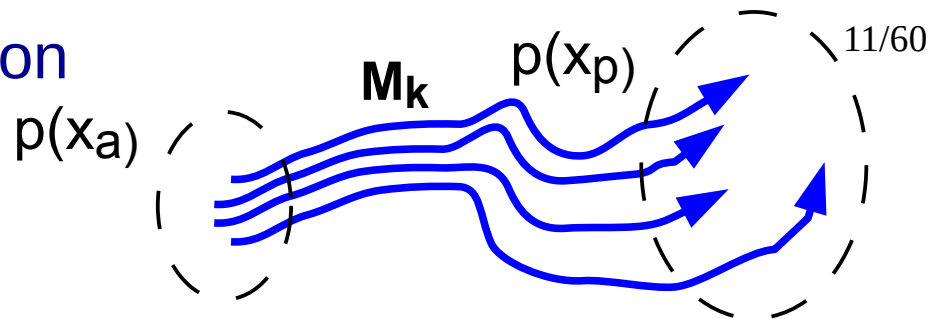
Les assimilations modernes (4DVar, 3DenVar, 4DenVar) utilisent des ensembles pour générer leurs fonctions de structure.

"Fonction de structure" = incrément qu'une analyse va produire en présence d'une obs (différente de l'ébauche)

Les fonctions de structures combinent

- **structures "statiques"** : Jb climatologique assez trivial, avec structures quasi-homogènes, isotropes et d'assez grande échelle (~100km ou plus selon le paramètre et le niveau)
- **structures "dynamiques"** (flow-dependent) : assez fidèles à des covariances entre membres d'un ensemble (moyennant une "localisation")
- NB. pondération α entre les 2 à vérifier
- NB. en 3D/4DenVar, les structures dynamiques restent limitées par la résolution de l'ensemble

Prévision d'ensemble et assimilation



Pour estimer la matrice B:

- tirer k états x_{ak} qui suivent la loi de probabilité de x_a
- la PDF des prévisions $x_p = M(x_a)$ est approximée par k prévisions perturbées: $x_{pk} = M_k(x_{ak})$
- alors $B \sim \text{cov}(x_{bk}) \sim \text{cov}(M(x_{ak})) \rightarrow$ fournit les fonctions de structure

Pour estimer les erreurs d'analyse:

- on tire k états $x_{bk} y_k$ qui ont pour covariances B et R
- le BLUE produit un ensemble d'analyses: $x_{ak} = x_{bk} + K(y_k - H[x_{bk}])$
- alors l'erreur d'analyse $A \sim \text{cov}(x_{ak})$

Ensemble d'assimilation cyclées (EDA= ensemble data assimilation):

- on combine $x_{pk} = M_k(x_{ak})$ et $x_{ak} = x_{bk} + K(y_k - H[x_{bk}])$
- ces k systèmes d'assimilation perturbés modélisent la *propagation temporelle des erreurs*

Equations du filtre de Kalman

$$K = B H^T (H B H^T + R)^{-1}$$

$$x_a = x_b + K (y - H[x_b]) \quad \text{analyse OI}$$

$$A = (I - K H) B \quad \text{ou} \quad A = 2(J'')^{-1}$$

$$x_b(t) = M_{t-1 \rightarrow t} x_a(t-1) \quad \text{prévision}$$

$$B(t) = M'_{t-1 \rightarrow t} A(t-1) M'^T_{t-1 \rightarrow t} + Q(t)$$

avec Q = covariances de "l'erreur modèle": $x_t(t) - M_{t-1 \rightarrow t} x_{t-1}$

Le filtre de Kalman est statistiquement optimal si M est linéaire. On peut aussi l'utiliser avec un modèle non-linéaire (= filtre de Kalman étendu EKF)

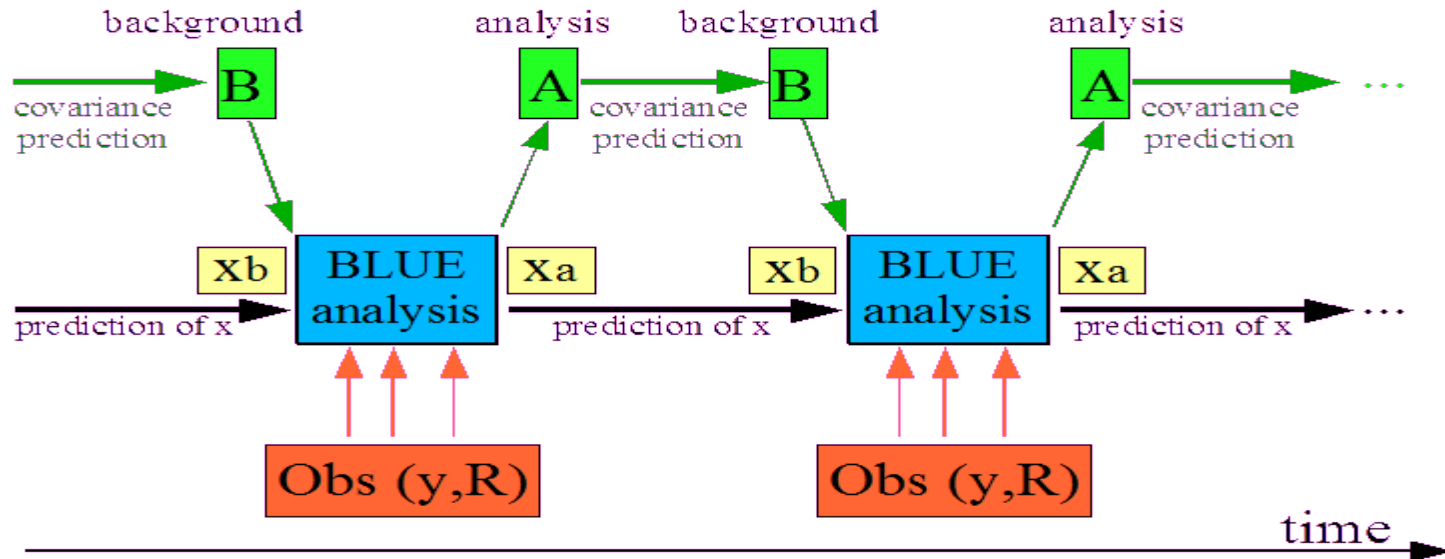
schéma du Filtre de Kalman

Le filtre de Kalman revient à compléter les équations d'analyse & de prévision de x par une récurrence sur B :

- A = covariances d'erreur d'analyse de l'OI (déjà vue)
- plus l'équation de **prévision des covariances d'erreur**:

$$B = M A M^T + Q \quad (Q \text{ erreur de modélisation})$$

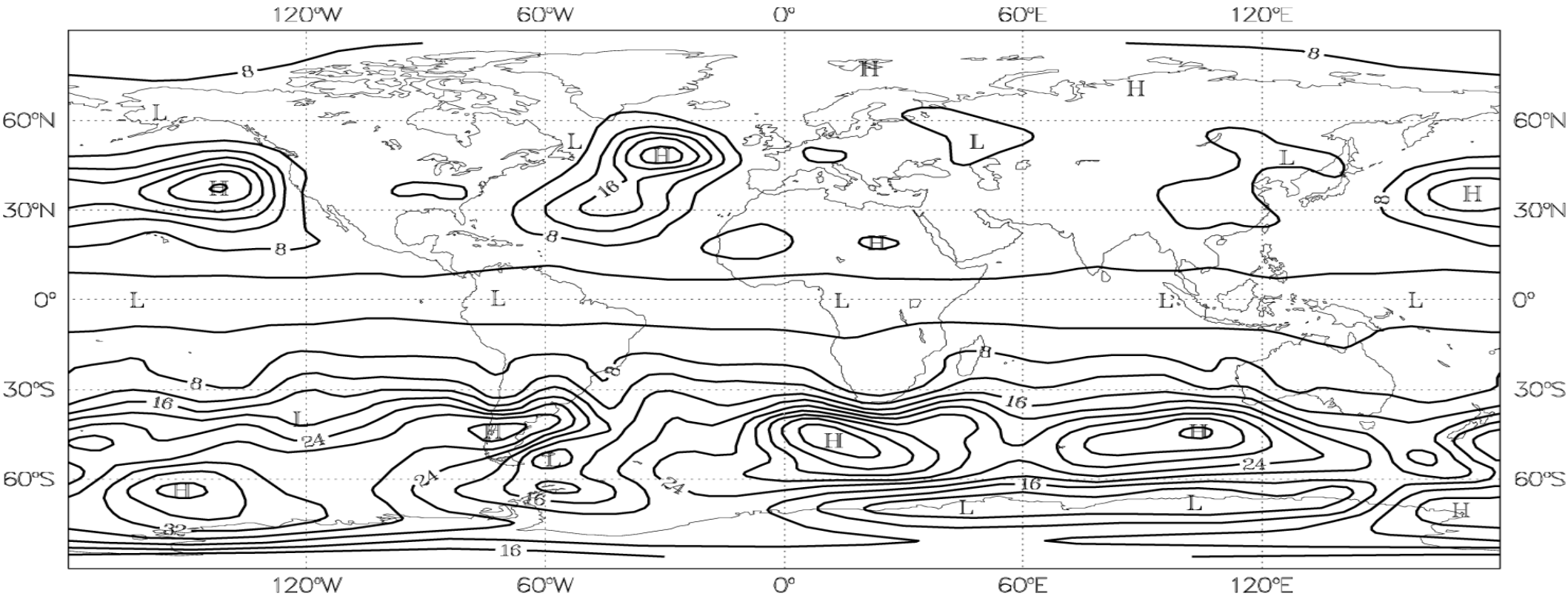
qui met à jour A avec le modèle M : **calcul physique de B**



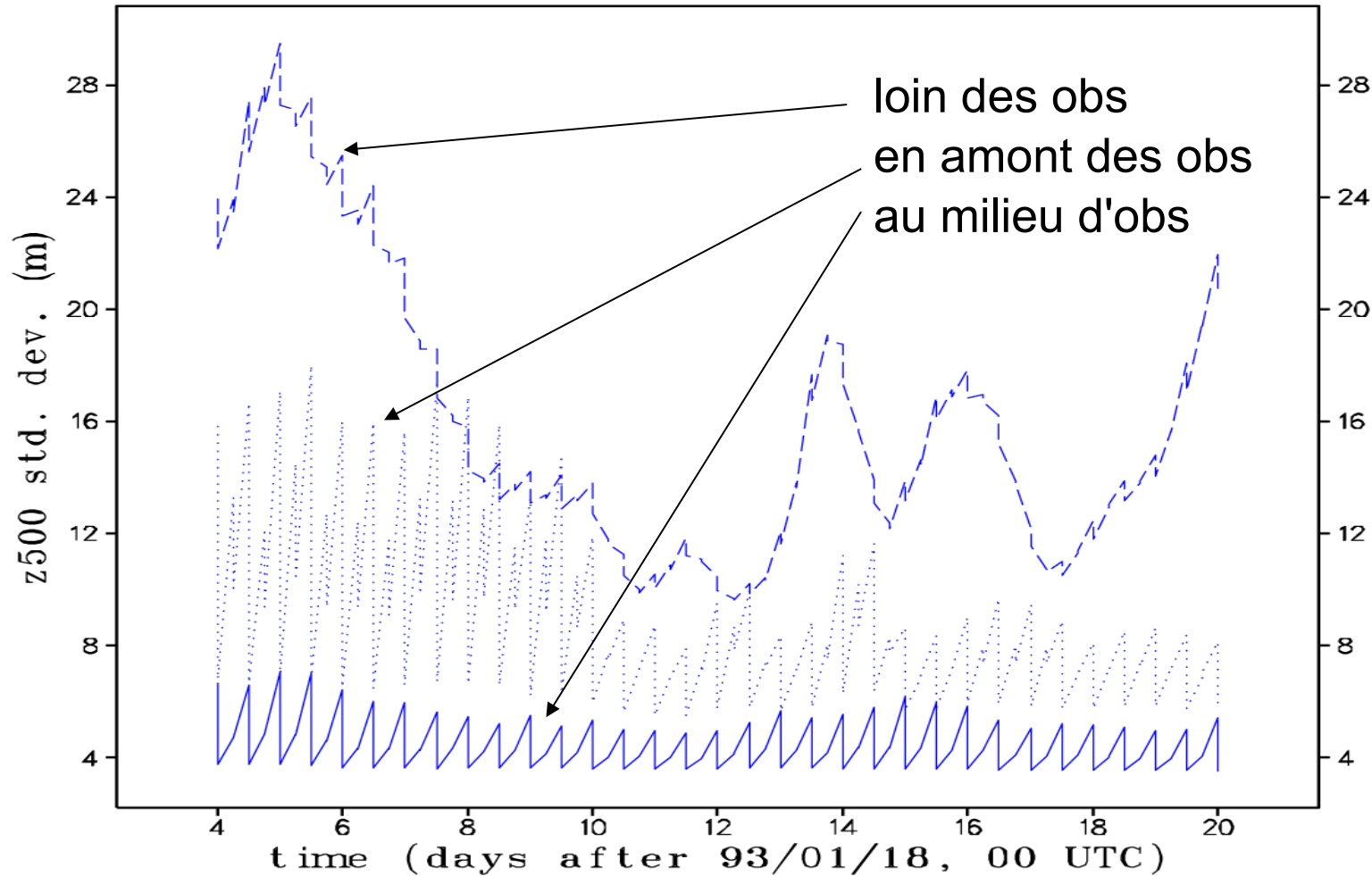
Information sur B fournie par le KF

Variances dynamiques dans B

(écart-type de z 500hPa dans l'atmosphère)



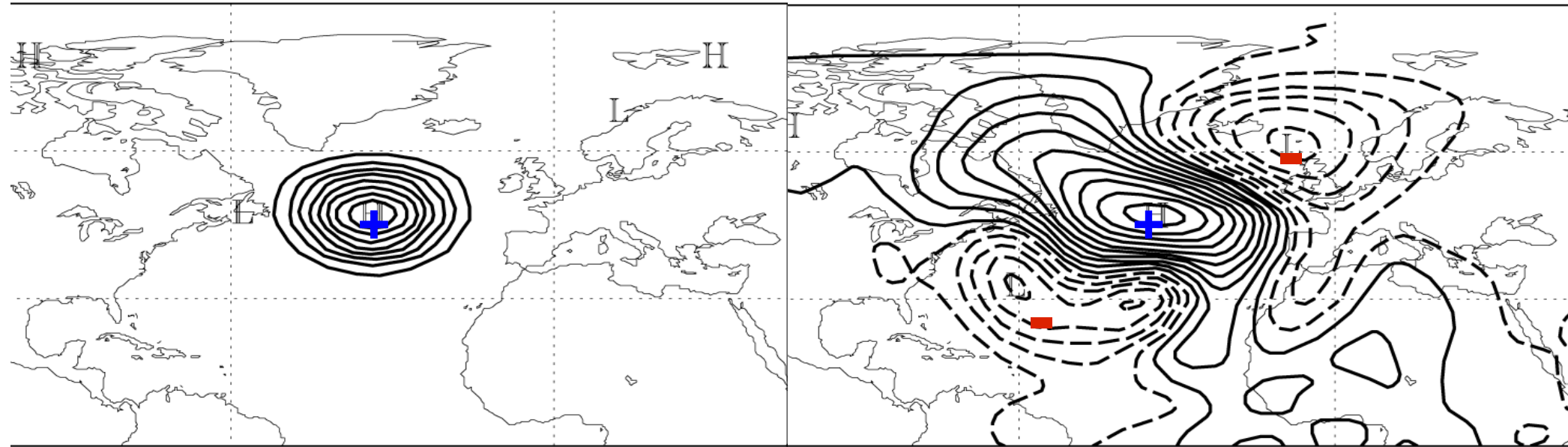
Evolution temporelle de l'incertitude dans une assimilation séquentielle en 3 points d'un KF



Exemples de fonctions de structure dynamiques (théorème: 4Dvar et KF produisent le même résultat)

3DVar: B statique modélisé avec
corrélations gaussiennes,
structure= BH^T

filtre de Kalman et 4DVar:
B produit par le modèle
structure= BM^TH^T



Evolution dynamique des covariances d'erreur en filtre de Kalman

$$\mathbf{B} = \mathbf{M} \mathbf{A} \mathbf{M}^T + \mathbf{Q}$$

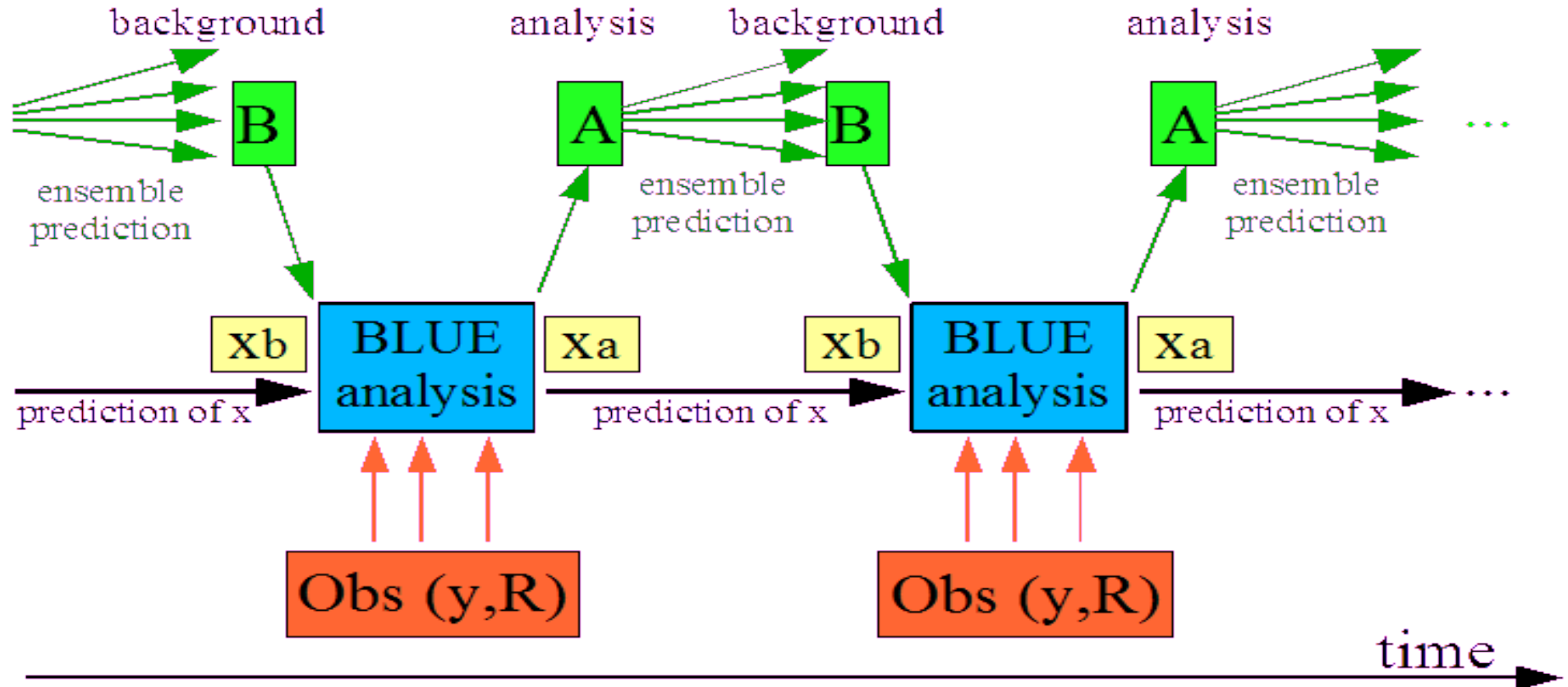
- **amplification** des variances par le modèle dans les zones dynamiquement instables (=peu prévisibles)
- **réduction des variances par l'analyse** au voisinage des observations
- **transport** des erreurs le long du flux

'Divergence du filtre de Kalman': si les variances deviennent trop petites, l'analyse n'utilise plus les obs.

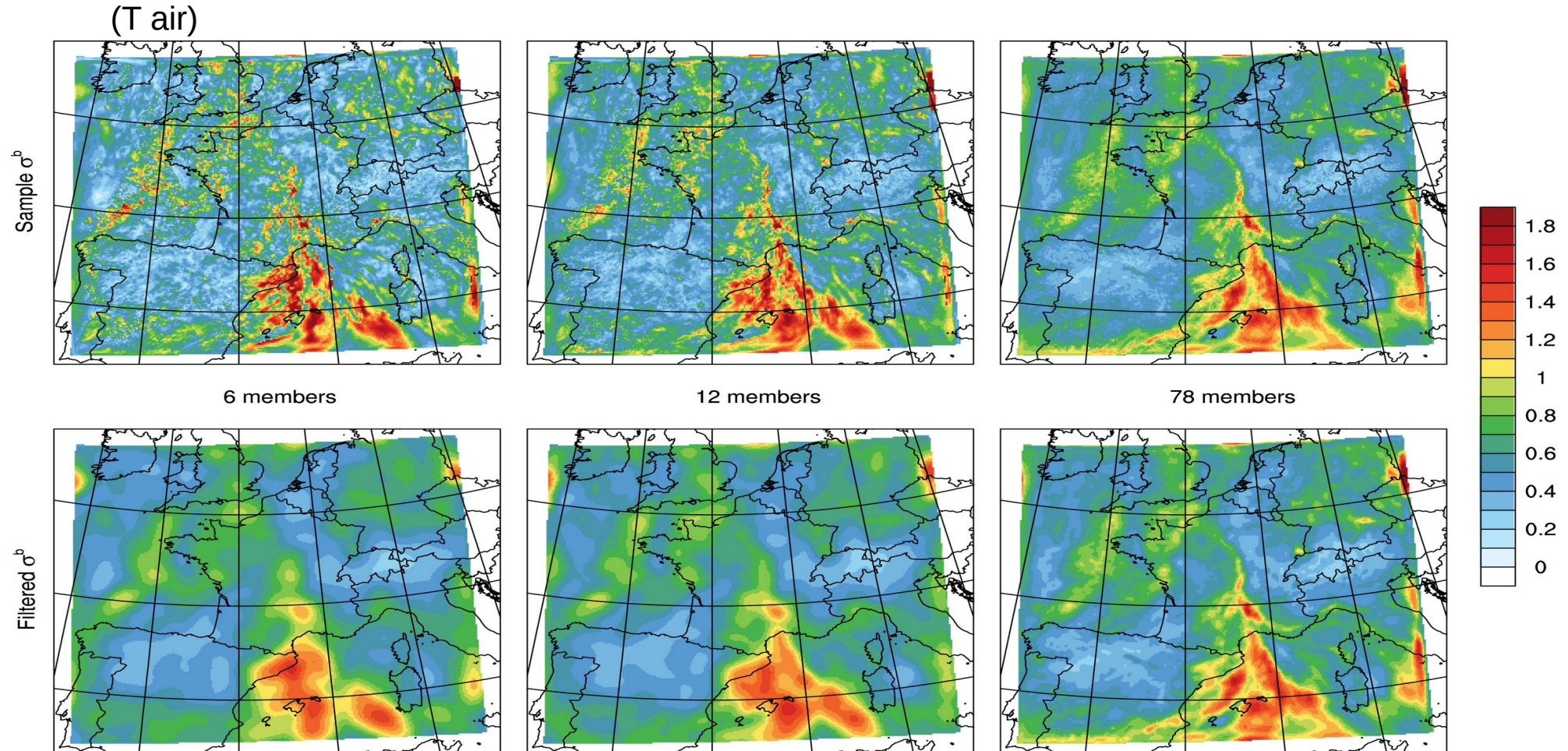
- dispersion des erreurs = étalement des corrélations
- activité ondulatoire = corrélations négatives

Filtre de Kalman d'ensemble:

remplacer l'équation $B = MAM^T + Q$ qui est très chère,
par des prévisions d'ensembles



Exemples de variances d'ébauche ensemblistes



Trois types de Filtre de Kalman d'ensemble:

EnSKF 'square root ensemble filter':

- ensemble de $x_b \rightarrow B \rightarrow \text{BLUE} \rightarrow A \rightarrow$ ensemble d'analyses x_a
- on perturbe aléatoirement les observations avec R

ETKF 'ensemble transform Kalman filter':

- comme EnSKF, plus une *restriction du calcul du gain K au sous-espace engendré par les x_{bk}*
- on ne perturbe pas les obs (=filtre 'déterministe')
- peut être combiné avec une stratégie de sélection locale des observations ('LETKF' = local ETKF)
- plus compliqué mais moins coûteux que le EnSKF

EDA 'ensemble data assimilation':

- comme EnSKF, mais *on fait autant d'analyses que de membres*
- évite de passer par A
- plus coûteux que EnSKF, mais plus simple, et performant

le futur: **EnVar** : hybridation d'EDA et de 4DVar.

En résumé sur l'assimilation d'ensemble:

Le filtre de Kalman est une méthode pour produire 1 assimilation

- avec une équation d'évolution de B
- on peut utiliser des méthodes stochastiques (=des prévisions ou des assimilations d'ensemble) pour produire B
- le but reste de produire un seul x_a optimal à chaque instant

L'assimilation d'ensemble est une méthode pour produire des PDFs de x_a et de x_b

- on calcule en parallèle k instances perturbées d'une assimilation
- cette assimilation peut être un filtre de Kalman ou une autre méthode
- elle peut utiliser ou non les x_{bk} pour contruire son B

Ces 2 approches peuvent être combinées ou non.



Techniques de perturbation des ensembles

Deux objectifs :

- obtenir une **dispersion correcte en moyenne (fiabilité)**
- obtenir un bon signal = une dispersion qui **dépende de la situation (résolution)**

En théorie, les perturbations peuvent être définies automatiquement (filtres de Kalman, filtres particuliers, machine learning)

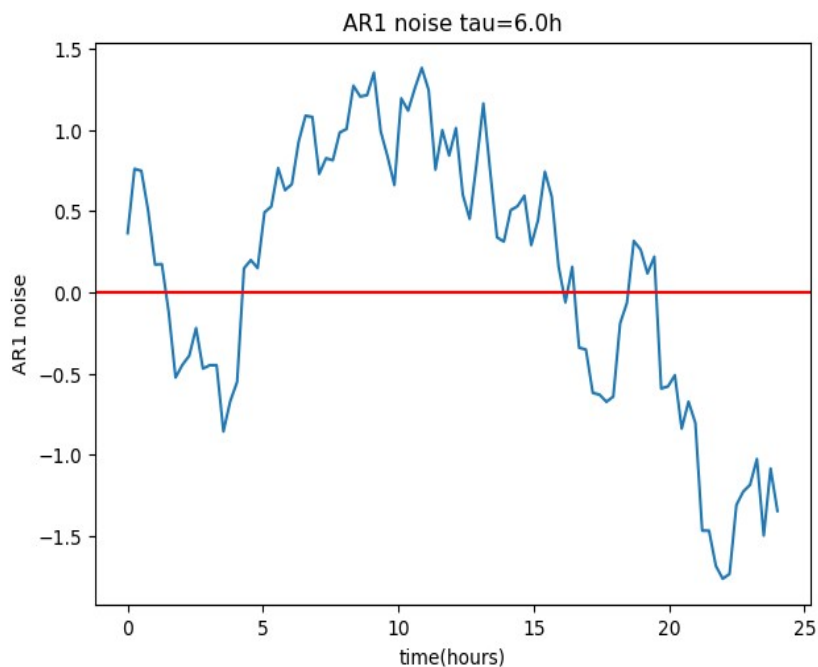
En pratique, on n'a jamais assez de données et de puissance calcul, donc :

- définir les perturbations avec du **sens physique**
- **optimiser des scores** sur des périodes-test
- perturber "n'importe comment" peut produire un bon ensemble (si système chaotique)

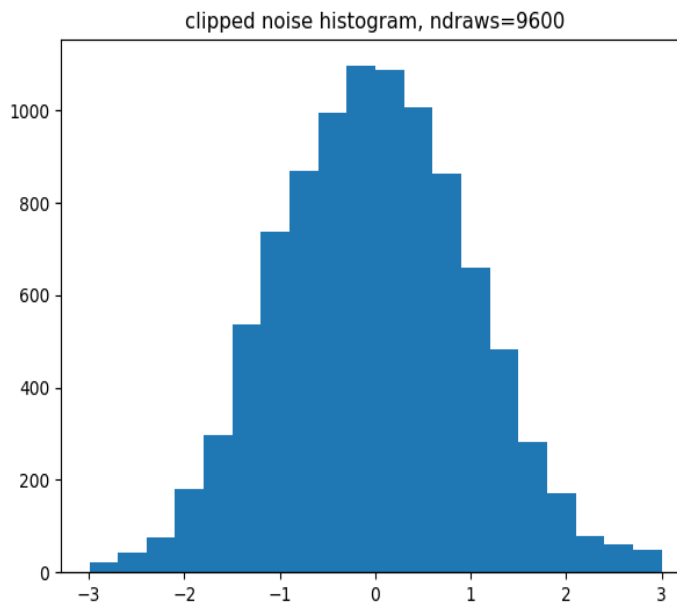
Techniques de perturbation : générateur de bruit autocorrélé pour respecter la régularité spatio-temporelle des variables physiques

bruit temporel avec processus aléatoire AR1 :

série temporelle

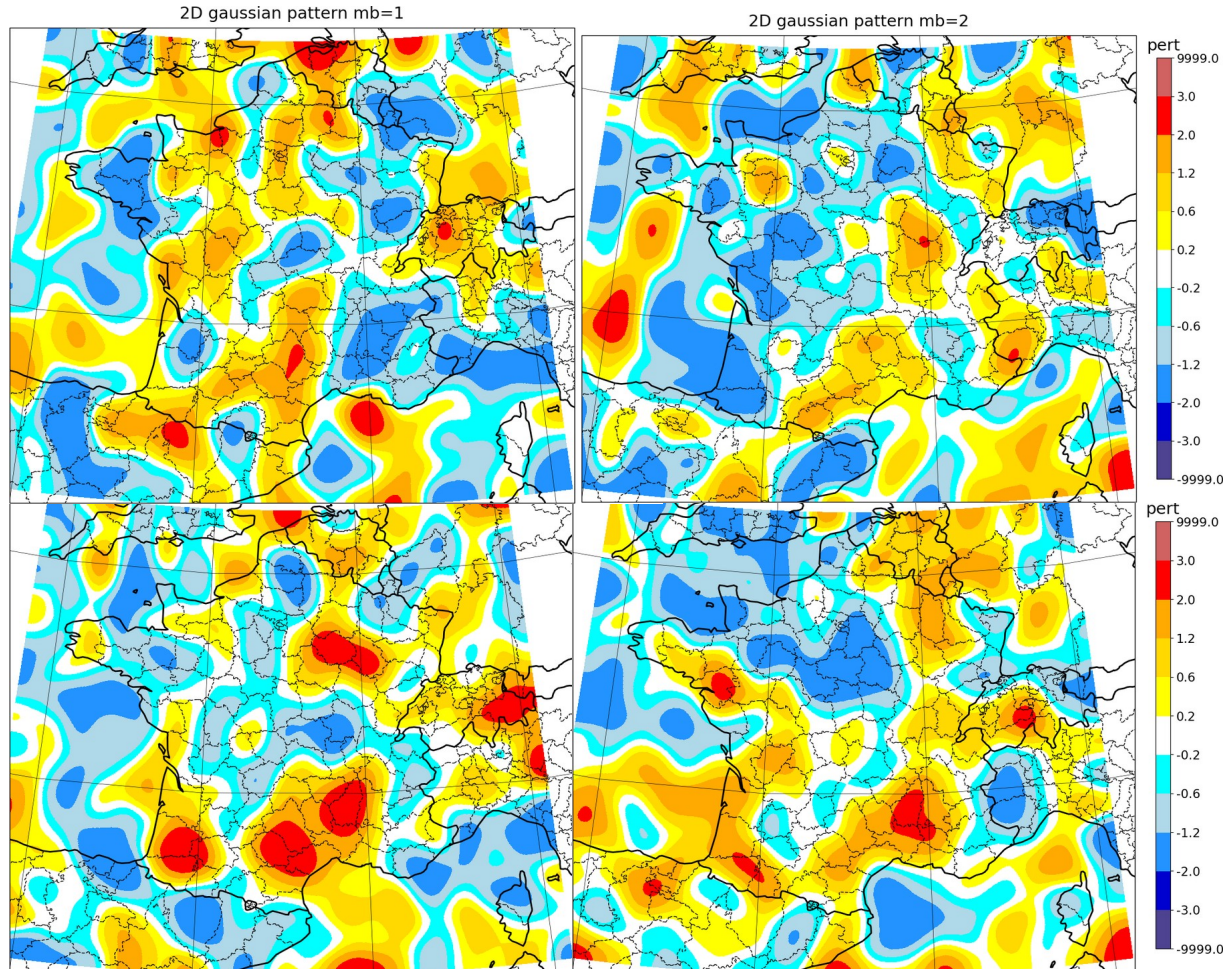


distribution sur une longue période



champ de bruit 2D : par lissage spatial d'un bruit blanc (corrélation résultante quasi-gaussienne)

4 réalisations



On perturbe en priorité des paramètres impactants -> propagation des perturbations dans le reste du modèle

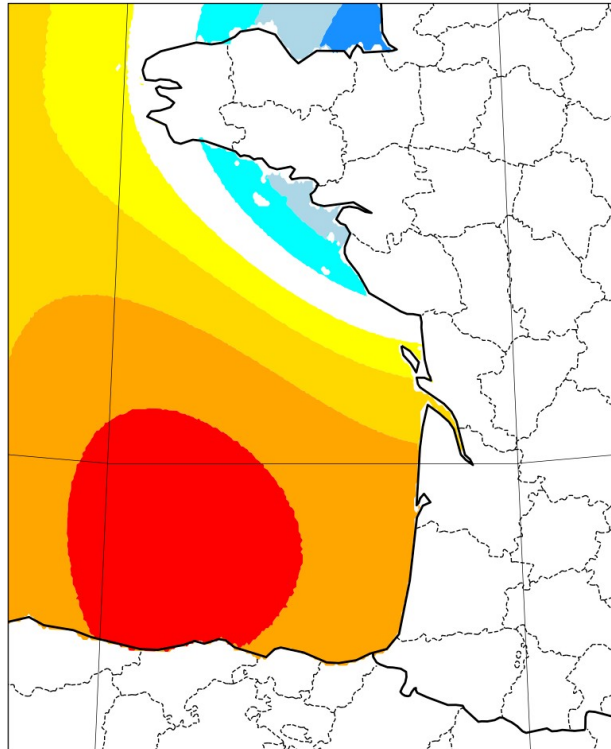
ex de champ impactant : température de surface de la mer

perturbation de Tmer

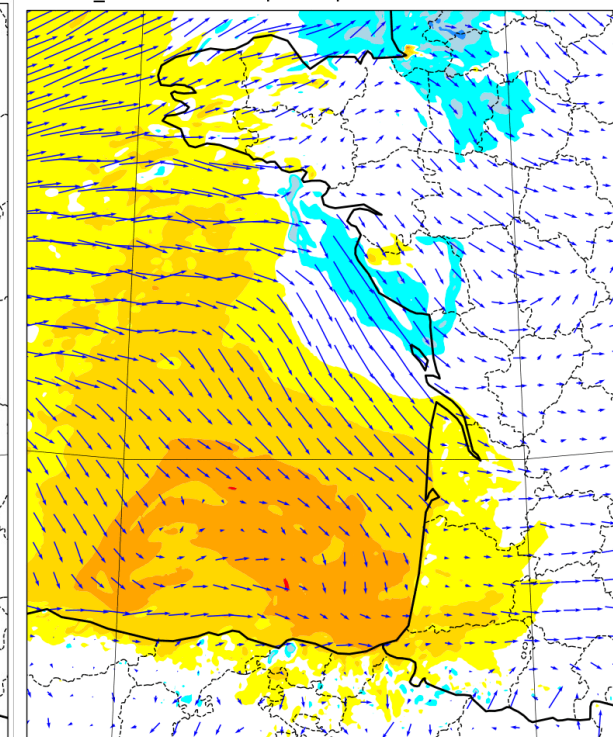
réponse de T10m après 10h

impact sur l'écart-type de T10m
après 10h (16 membres)

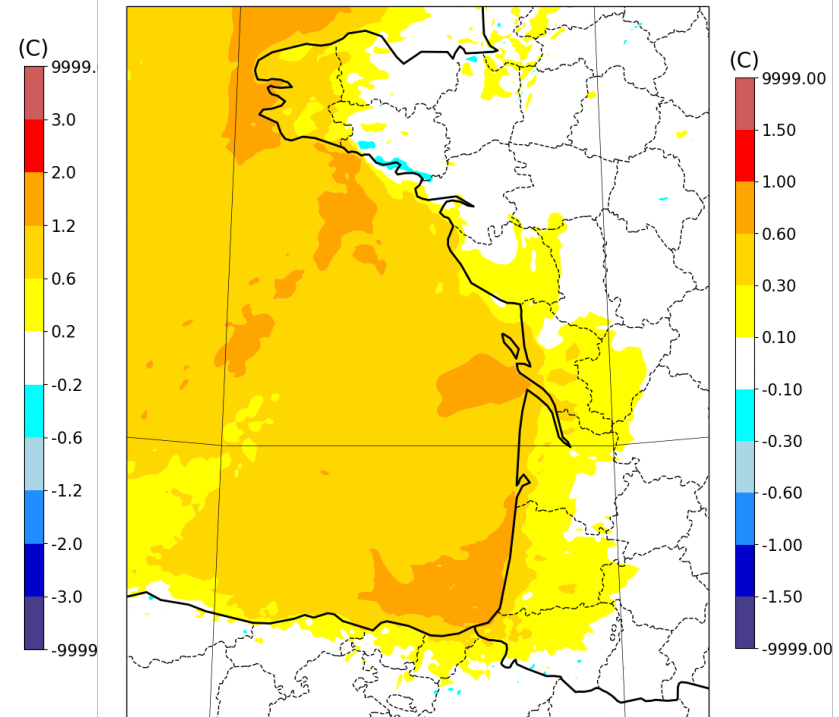
mb01 Ts diff(PEaroper-expGBRD) 202108142100+10



mb01 T_10m diff(PEaroper-expGBRD) 202108142100+10



T_10m(C) stdev diff(PEaroper-expGBRD) 202108142100+10



Techniques de perturbation

- perturber "n'importe comment" peut produire un ensemble utile :
exemple du modèle chaotique de Lorenz

(fig. Palmer, 2016)

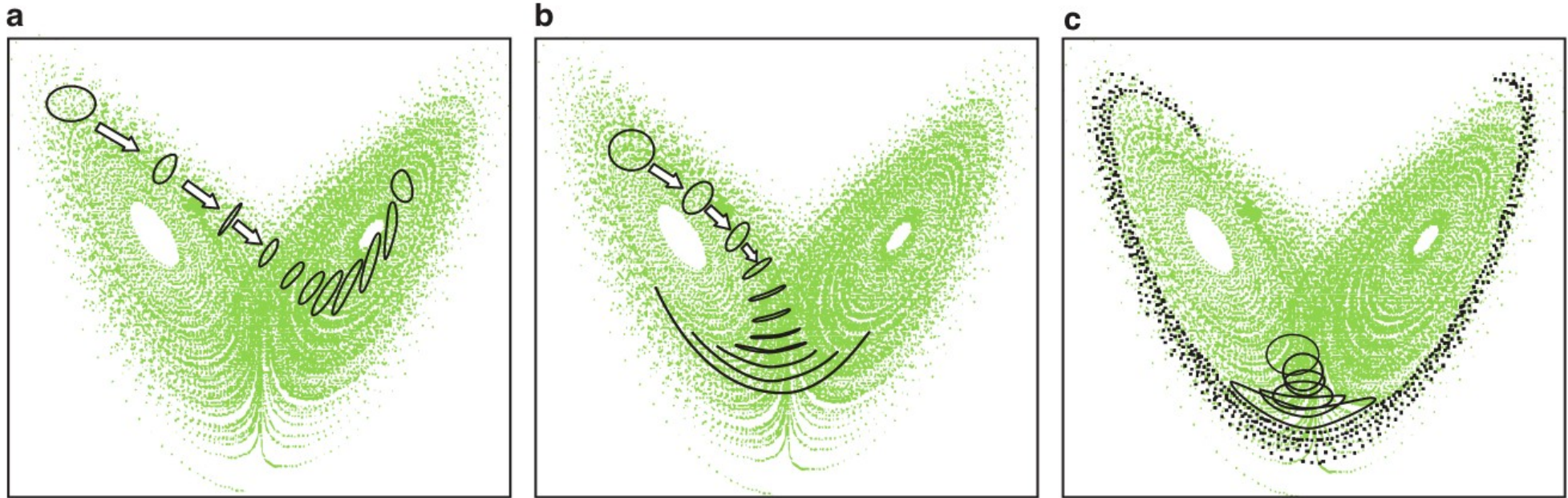
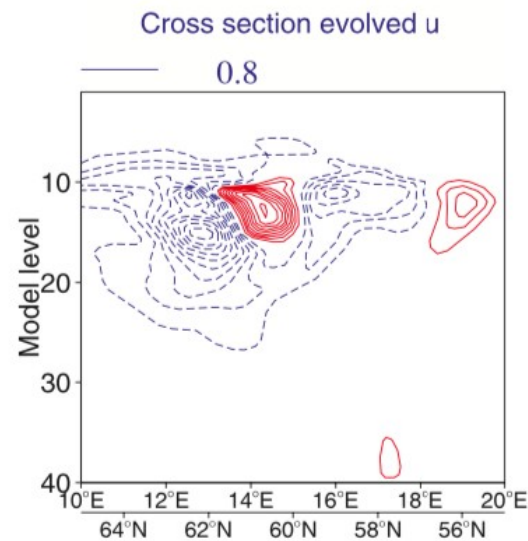
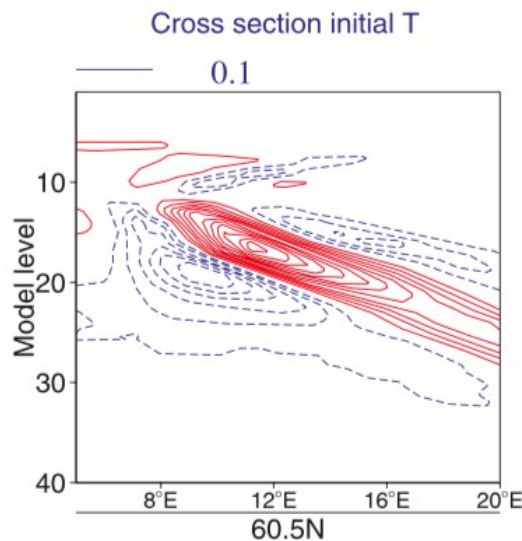
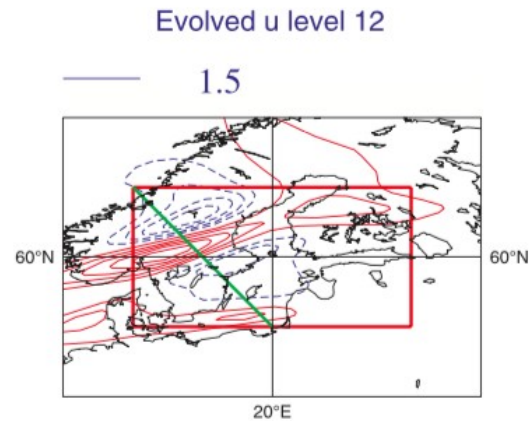
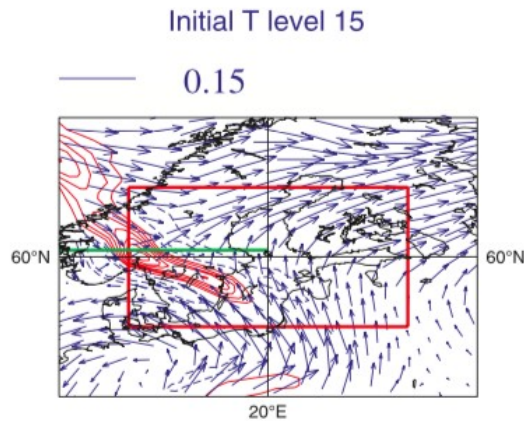


Figure 1 Scientific basis for ensemble forecasting. In a nonlinear system the growth of initial uncertainty is flow dependent – here illustrated with the Lorenz (1963) model. The set of initial conditions (black circle) is located in different regions of the attractor in (a), (b) and (c).

perturbations "vecteurs singuliers"
à croissance maximale sur la
situation météo du jour

maximise
 $\| M(x+dx) - M(x) \| / \| dx \|$
par méthode de Lanczos

coupes verticales
de vecteur singulier



Techniques de perturbation

Etat initial:

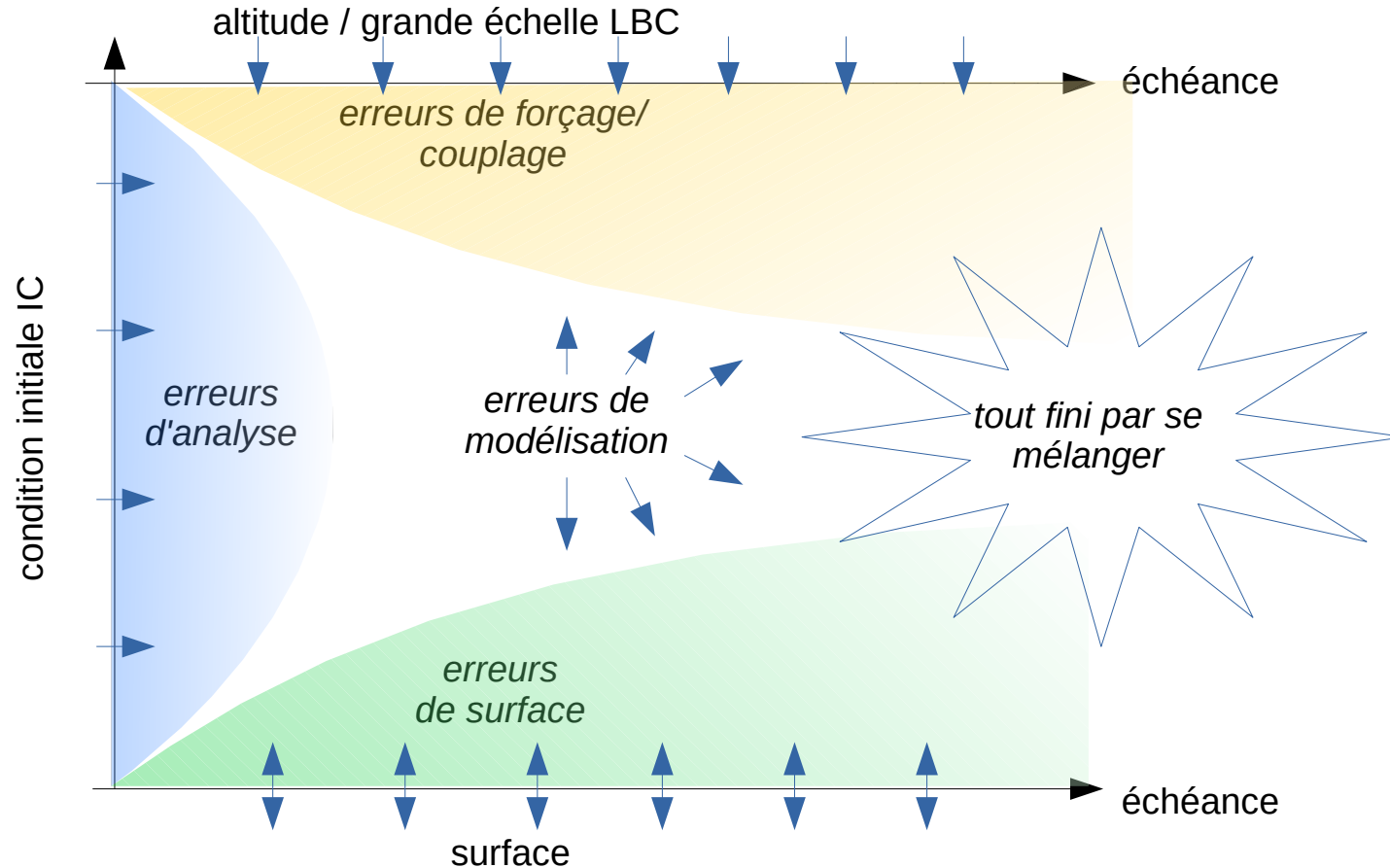
- 'breeding': utiliser les différences entre prévisions produites hier
- vecteur singuliers (singular vectors): calculer les n perturbations de l'analyse dont l'énergie va croître le plus vite (ce sont les plus grands vecteurs propres de MAM^T , calculables par méthode de Lanczos)
- assimilation d'ensemble
- filtre de Kalman d'ensemble

Modèle:

- perturber les paramètres du modèle
- utiliser plusieurs physiques
- injecter du bruit dans l'équation d'évolution du modèle (physique stochastique)

Techniques de perturbation : effet comparé des perturbations initiales (IC) et latérales (LBC) et modèle

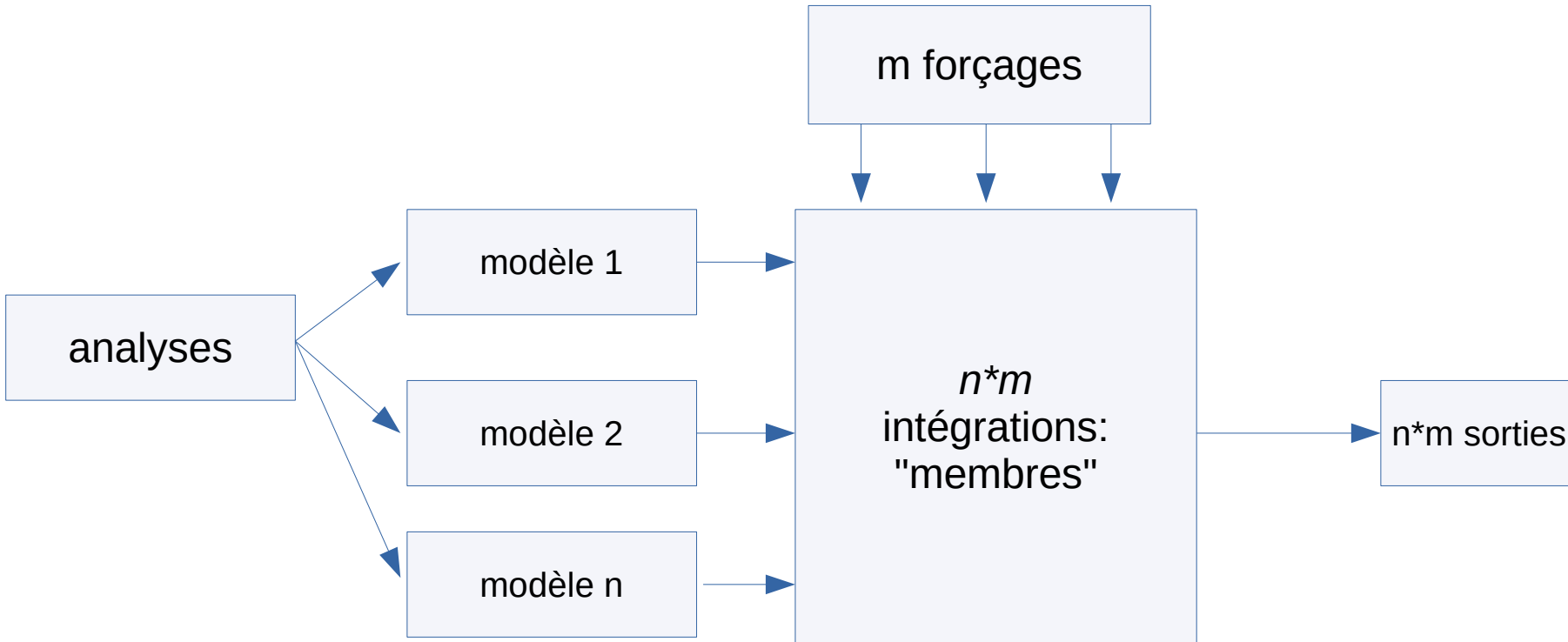
dans un modèle à aire limitée (LAM) :



flèches bleues
= sources d'erreurs
= perturbations
ensemblistes

En panachant des modèles, analyses, forçages, etc., souvent déterministes

On ne sait pas trop de quel système cela calcule l'incertitude, mais c'est moins cher...

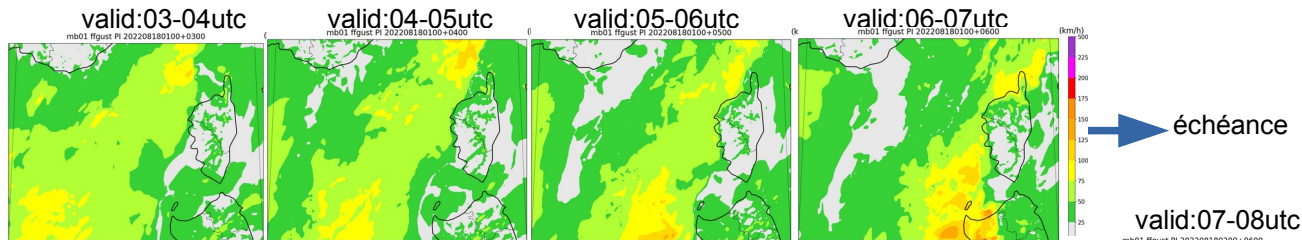


ensembles décalés : 1 modèle au fil de ses mises à jour

31/60

chaque rangée =
1 prévision par heure
(cartes de rafales)

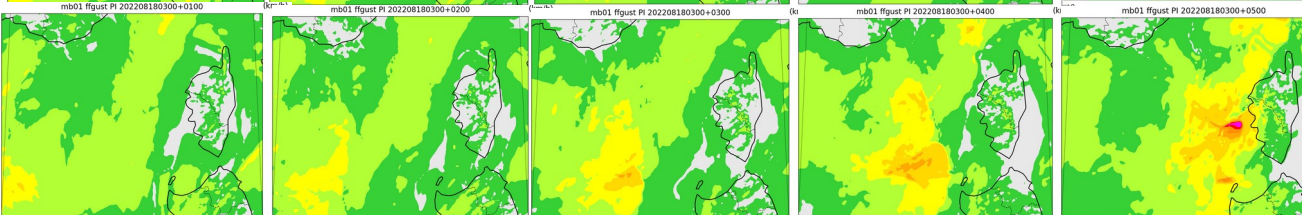
01utc



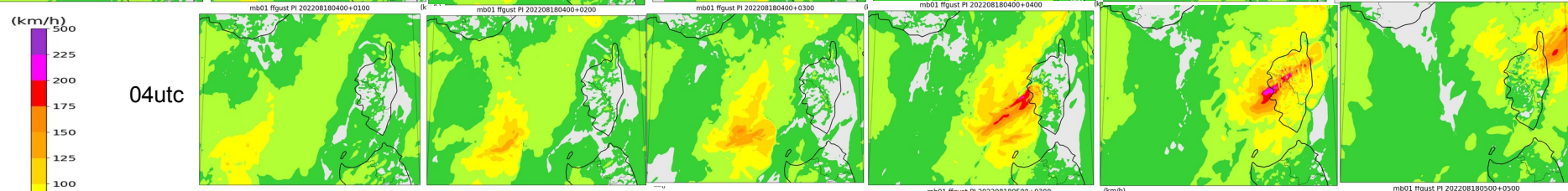
02utc



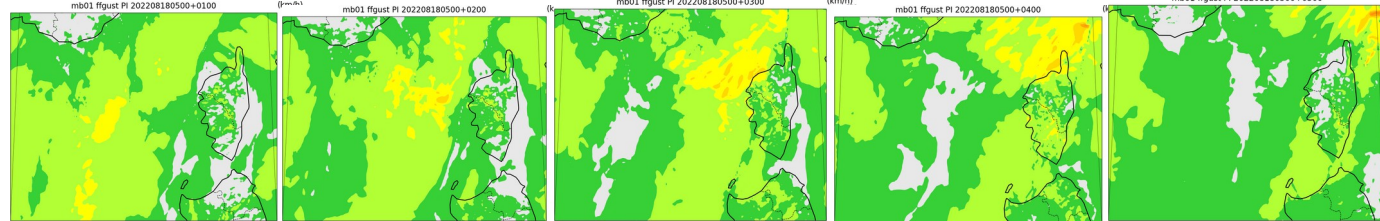
03utc



04utc



05utc

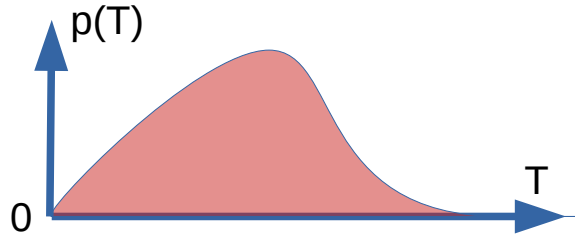


heure de
lancement

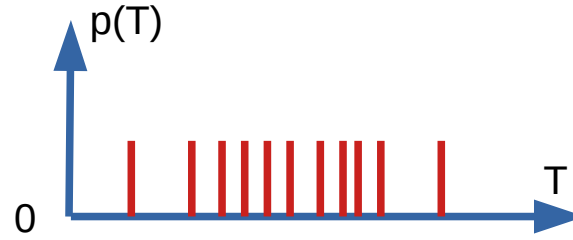
Métriques d'évaluation des ensembles

- principe de la PE
- construction des perturbations
- métriques d'évaluation
- produits

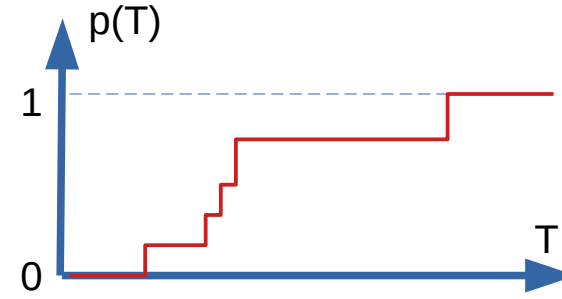
Exemples : incertitude d'un **paramètre réel** (ex : T en 1 point) :



une fonction de distribution continue
(PDF : probability density function)



une PDF discrète



une PDF discrète
représentée par son
intégrale (CDF :
cumulative density
function)

1 marche par membre

Quelques "statistiques" càd des manières de résumer une PDF :

- moyenne
- min, max
- quantile Q_x : valeur de T telle que x % de la distribution est en-dessous
- médiane Q_{50} : valeur de T telle que 50 % est en-dessous
- probabilité de dépasser un seuil t : $P(T > t) = \text{intégrale (surface) de la PDF à droite de } t$

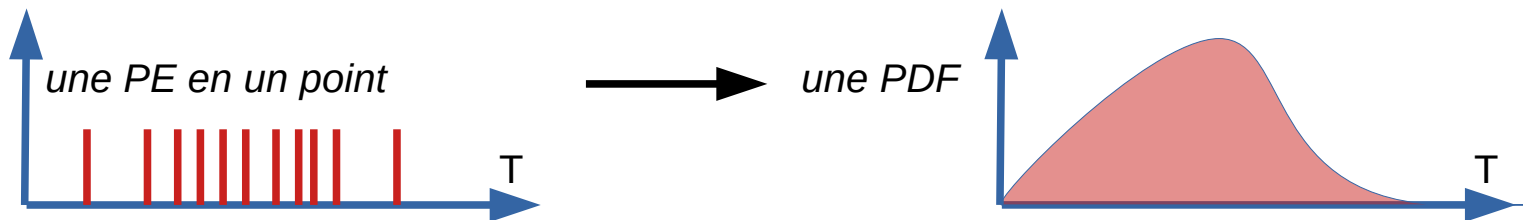
La proba $P(T > Q_x)$ de dépasser un quantile Q_x est donc $1-x$: 10 % de chances d'être au-dessus du Q_{90} .

Scores et évaluation des prévisions d'ensemble

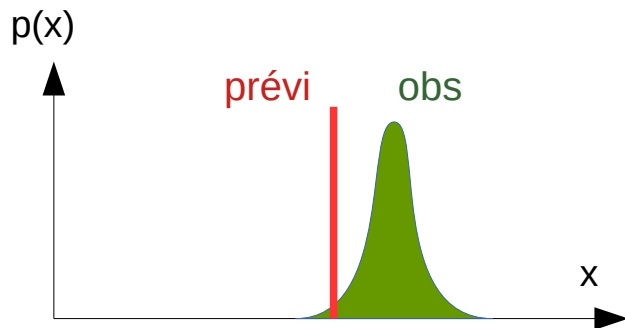
Un très vaste sujet, difficile à comprendre et à utiliser proprement !

Difficultés :

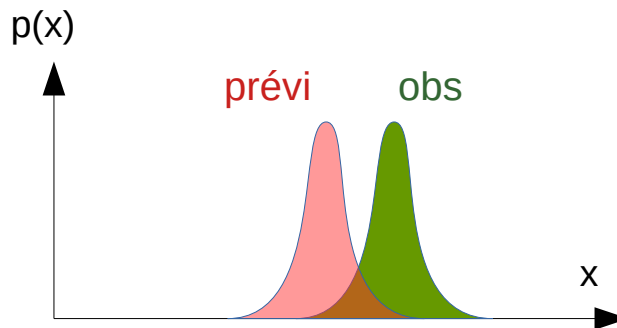
- la PE mesure une incertitude donc **on ne peut pas la juger sur un seul cas**
- la PE ne fournit pas 1 valeur mais une fonction (une distribution) : infinité de manières de l'utiliser → **rester proche de l'utilisation que l'on va en faire** (ne pas espérer un "score universel")
- besoin de gros échantillons pour faire des statistiques stables
- les erreurs d'obs ont souvent un impact significatif
- on peut évaluer non pas la PE mais une PDF que l'on en déduit (et ce n'est pas la même chose!)
- profusion de scores redondants et/ou trompeurs
- les scores les plus simples à expliquer ne sont pas les plus utiles (et vice versa)



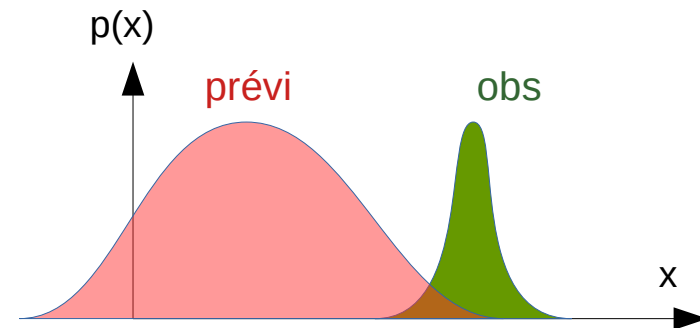
- est-ce la **précision des prévisions** du système déterministe ?
- ou la **fidélité de la distribution de la PE aux erreurs de prévision** ?



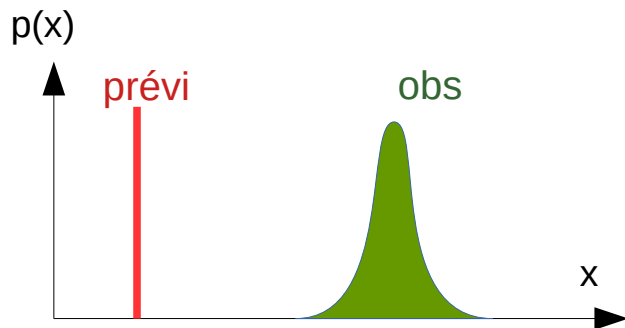
"bonne" prévi déterministe



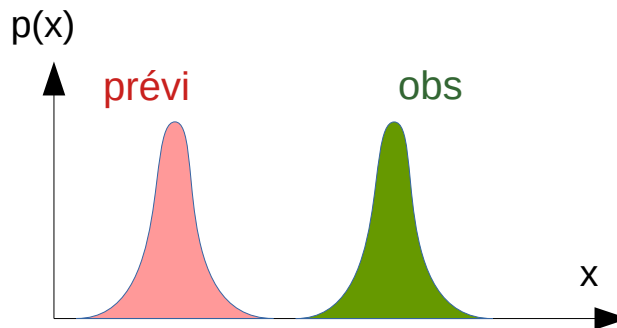
"bonne" prévi probabiliste



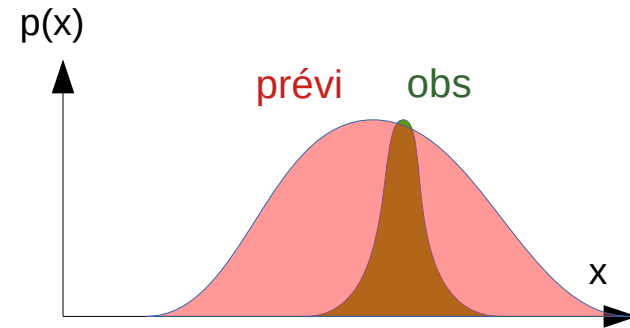
forte incertitude, mais bonne prévi



"mauvaise" prévision déterministe



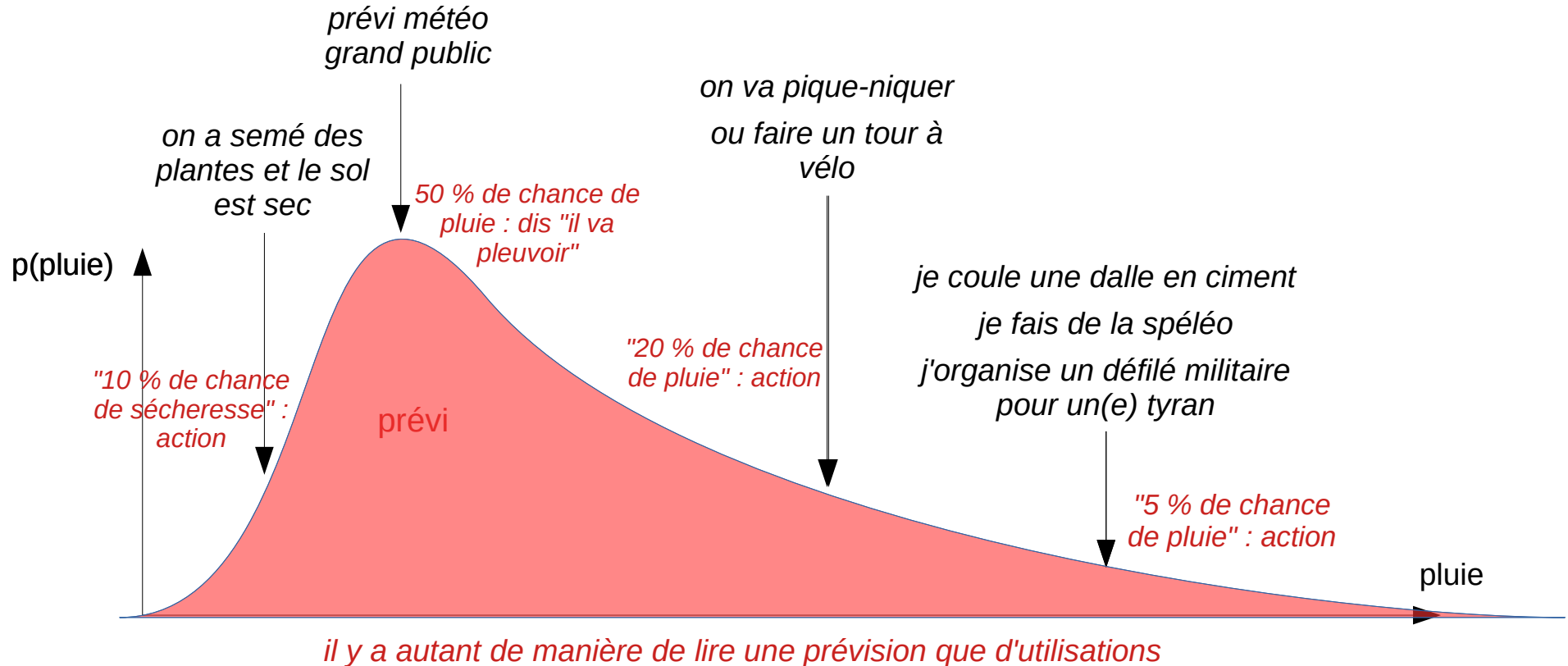
"mauvaise" prévision probabiliste



??? indécidable sur 1 seul cas

-> la notion de "**qualité d'une PE**" est ambiguë :

- est-ce la **précision des prévisions** du système déterministe ?
- ou la **fidélité de la distribution de la PE aux erreurs de prévision** ?
- ou son **"utilité"** ?



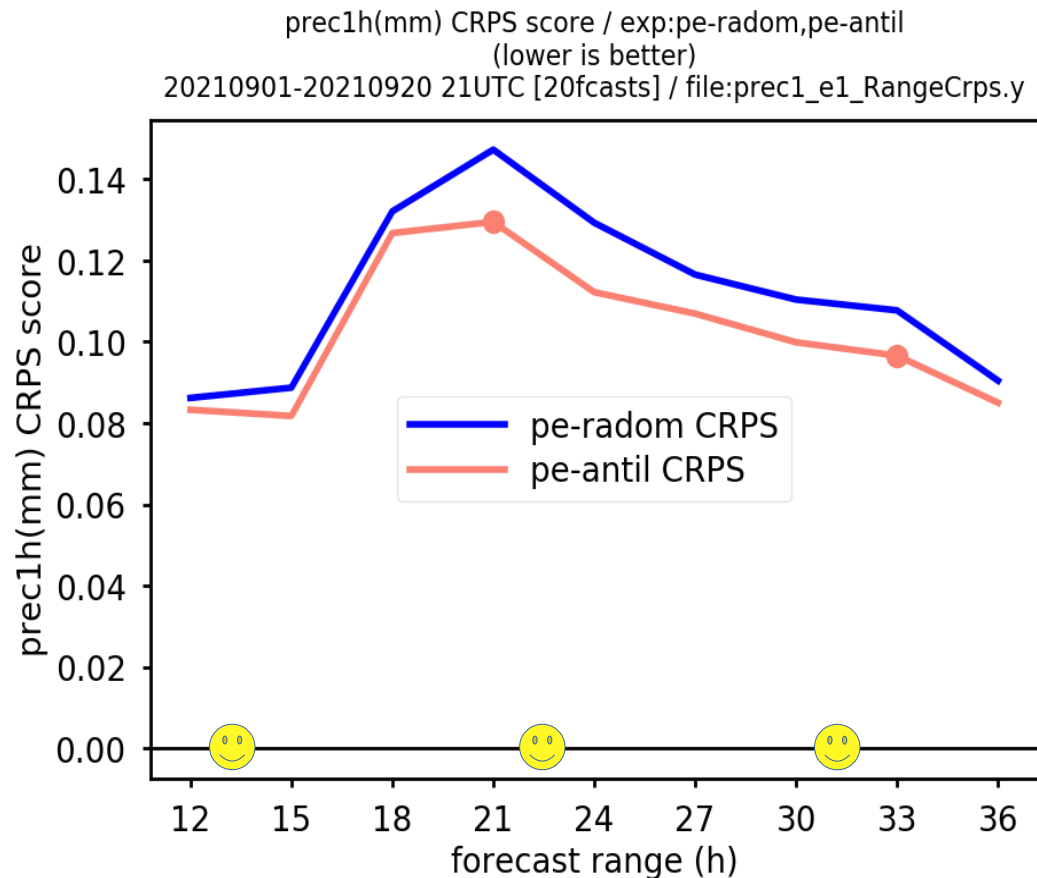
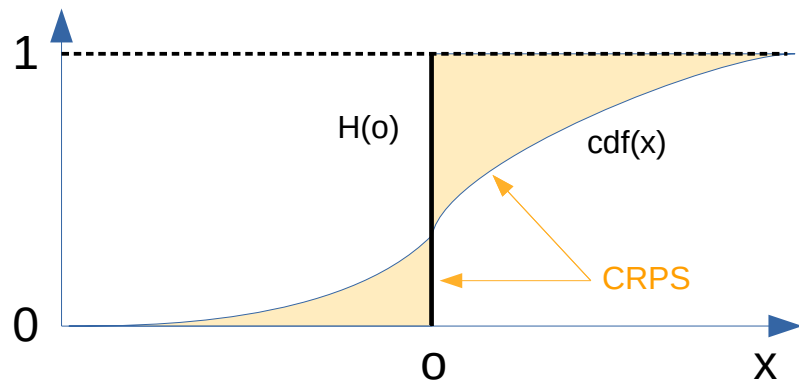
CRPS : un score intégral pour les prévisions réelles "Continuous Ranked Probability Score"

Déf : moy de CRPS = $\int [cdf(x_i) - H(obs)]^2 dx$ avec

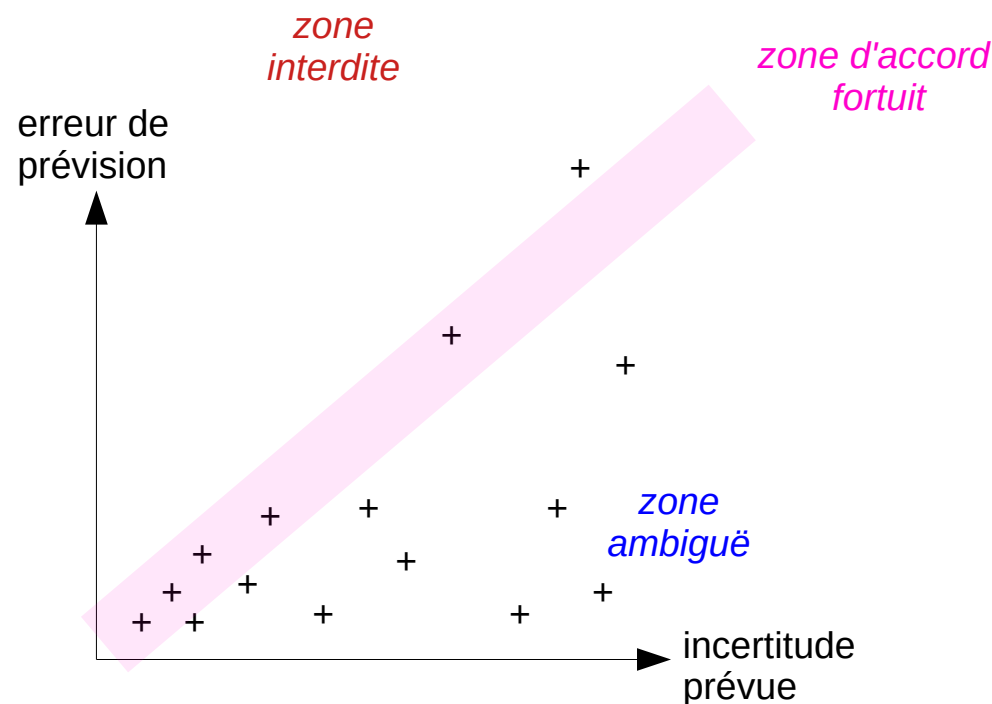
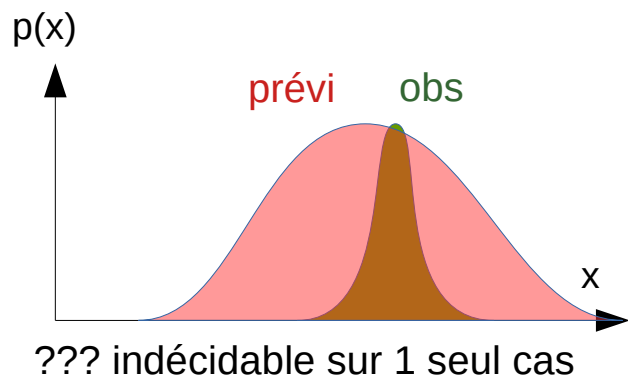
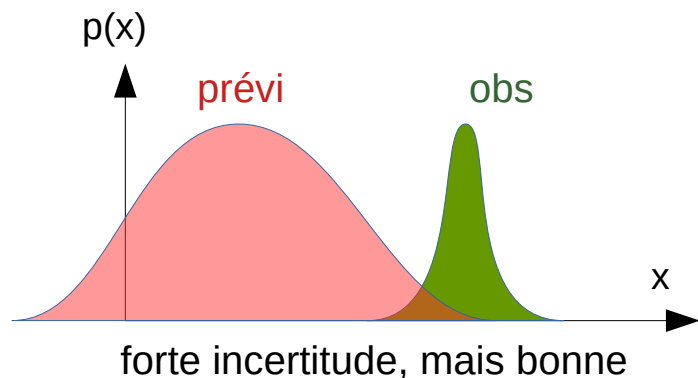
- $cdf(x_i)$ la distribution prévue par l'ensemble x_i
- $H(o)$ la fonction de Heaviside de l'obs o

Remarques :

- CRPS = $\int (\text{Brier sur tous les seuils } y) dy$
- pour une prévi déterministe, CRPS(x) = $|x - o|$ (c'est la MAE : mean absolute error)



- "erreur": distance entre prévision et observation de la réalité
- "dispersion": incertitude estimée a priori sur la prévision
- notion complexe car ce sont 2 quantités de natures différentes



Principe : lorsque l'ensemble prévoit un événement binaire avec x % de probabilité, est-il observé x % du temps ?

C'est un **calcul de biais des probabilités** qui **stratifie les prévisions par classe de probabilités prévues** :

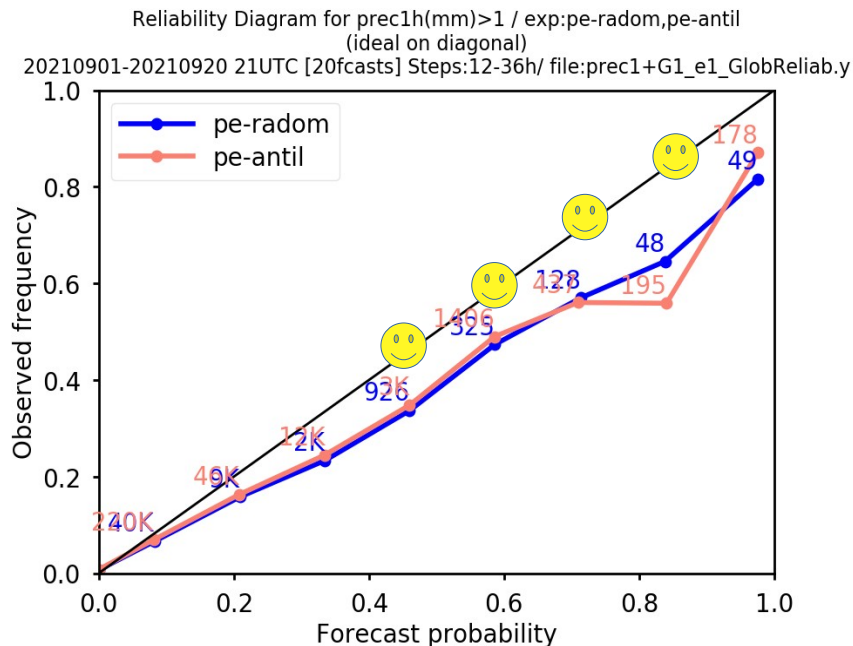
1) on fait un histogramme des probabilités prévues

2) on regarde dans chaque classe x à quelle fréquence y l'événement est observé : idéalement $y=x$

3) courbe trop plate = ensemble trop confiant (probas prévues trop proches de 1 ou de 0)

pièges : variations des populations de classes, et présence de classes peu peuplées

C'est un "sanity check" visuel bon pour la com, mais pas une mesure de qualité (utiliser plutôt le score de Brier)



niveau 4: **les scores de décision** : (~ "résolution") : la PE aide-t-elle à prendre de bonnes décisions ?

Dépend de **ce que veut** l'utilisateur, de la **qualité du système** de prévision, et de la **qualité des perturbations** :

- avantage : ils reflètent assez précisément l'utilité pratique d'un ensemble
- inconvénient : parfois difficiles à comprendre, et à relier aux ingrédients de l'ensemble
- nécessitent de bien modéliser l'utilité pour l'utilisateur

Exemples : pour un **évènement binaire y** :

- le **diagramme ROC** (relative operating characteristic) et d'innombrables scores qui s'en déduisent (CSI, Jaccard, PSS, HSS, AUC, etc)
- le diagramme de **valeur économique potentielle (PEV)**
- les **scores d'IA** (precision-recall, F1, F2, etc)

Tous utilisent la **technique de la table de contingence** :

- on résume les erreurs de prévision de y par un quadruplet (a,b,c,d)
- les scores ROCA, CSI, HSS, PEV, F1, etc sont des fonctions de (a,b,c,d)

	obs=ooui	obs=non
prévi=ooui	a	b (FA)
prévi=non	c (ND)	d

Scores et évaluation des prévisions d'ensemble : la table de contingence, base des scores "de décision"

41/60

Première étape : définir l'évènement que l'on score :

- résume les erreurs de prévision de probabilité d'un évènement binaire y : $p(y)$
- typiquement, pour scorer un paramètre réel x , on regarde les erreurs de prévision du **dépassement d'un seuil y par x - il y aura un score pour chaque combinaison (x,y) : $p(x>y)$**

Deuxième étape : construction de la table de contingence (a,b,c,d) pour les $p(x)$:

- on **définit** que la prévision est "oui" si la PE dit que $p(x)>z$ **z est un seuil de probabilité** (y est un seuil de valeur)
en général "on prévoit oui si z % de membres ont $x>z$ "
- sur un grand nombre n de paires $[p(x), z]$, on compte les 4 combinaisons :
 - a : détections (événement observé et prévu)**
 - b : fausses alarmes (prévu et pas observé)**
 - c : non-détections (observé et pas prévu)**
 - d : non-événement (ni observé ni prévu)**
- on calcule le score= $f(a,b,c,d)$ avec (a,b,c,d) fonction des 2 seuils (x,y)

	obs=oui	obs=non
prévi=oui	a	b (FA)
prévi=non	c (ND)	d

Remarques :

- (a,d) =bonnes prévis, (b,c) =mauvaises prévis.
- $a+b+c+d = \text{Nobs}$ et $a+c = \text{NobsYes}$ donc espace de dimension 2

Brier : un score intégral pour des prévisions binaires (donc % un seuil de valeur)

déf : moy de $B = \sum [p(x) - o]^2$
avec o fonction indicatrice de l'obs :

- $o=1$ si Obs=Oui
- $o=0$ sinon

prévision parfaite si $B=0$:

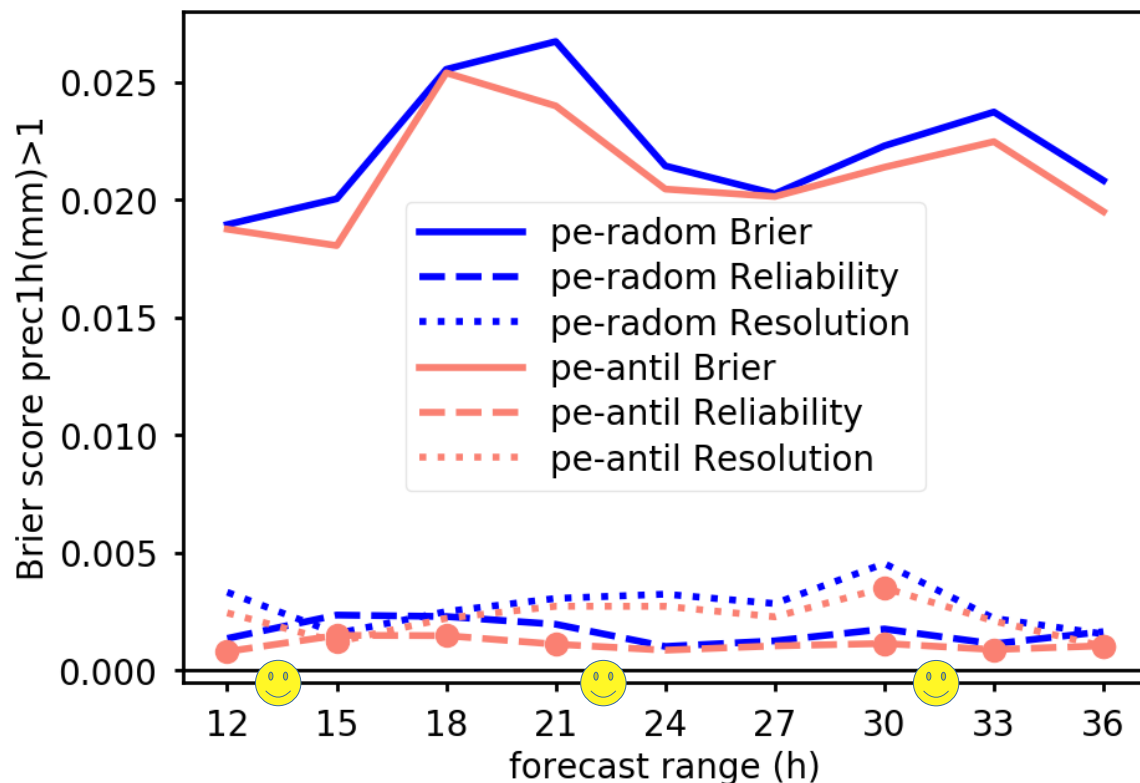
- il faut ne prévoir que des 1 et des 0
- et en même temps que les obs

Variante "Brier Skill Score" :

$$BSS = (B - B1) / (B2 - B1)$$

où B1, B2 sont les scores de prévis
de référence (ex : la clim)

Brier score prec1h(mm)>1 / exp:pe-radam,pe-antil
(lower is better)
20210901-20210920 21UTC [20fcasts] / file:prec1+G1_e1_RangeBrier.y



Défini pour un événement binaire :

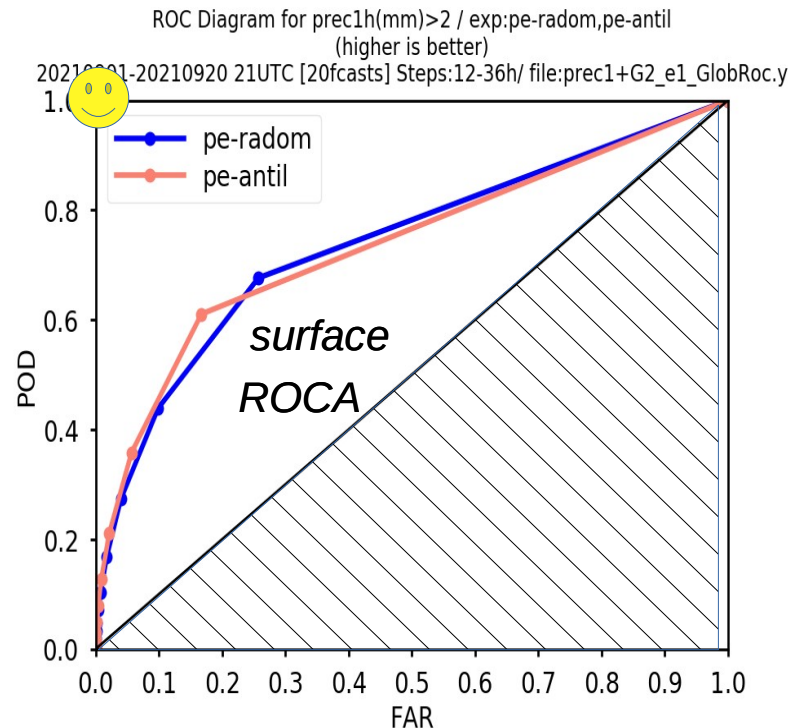
- on fixe le seuil de valeur y
- on fait varier le seuil de proba $z \rightarrow$ table de contingence paramétrée : $(a,b,c,d)=f(z)$ que l'on résume par les points
 - **abscisse** $FAR(z) = b/(b+d)$ **false alarm rate** (variante : certains utilisent $b/(b+a)$)
 - **ordonnée** $POD(z) = a/(a+c)$ **probability of detection**
- le ROC est la **courbe paramétrique** $[FAR(z), POD(z)]$ pour z dans $[0,1]$

Lecture de la courbe ROC :

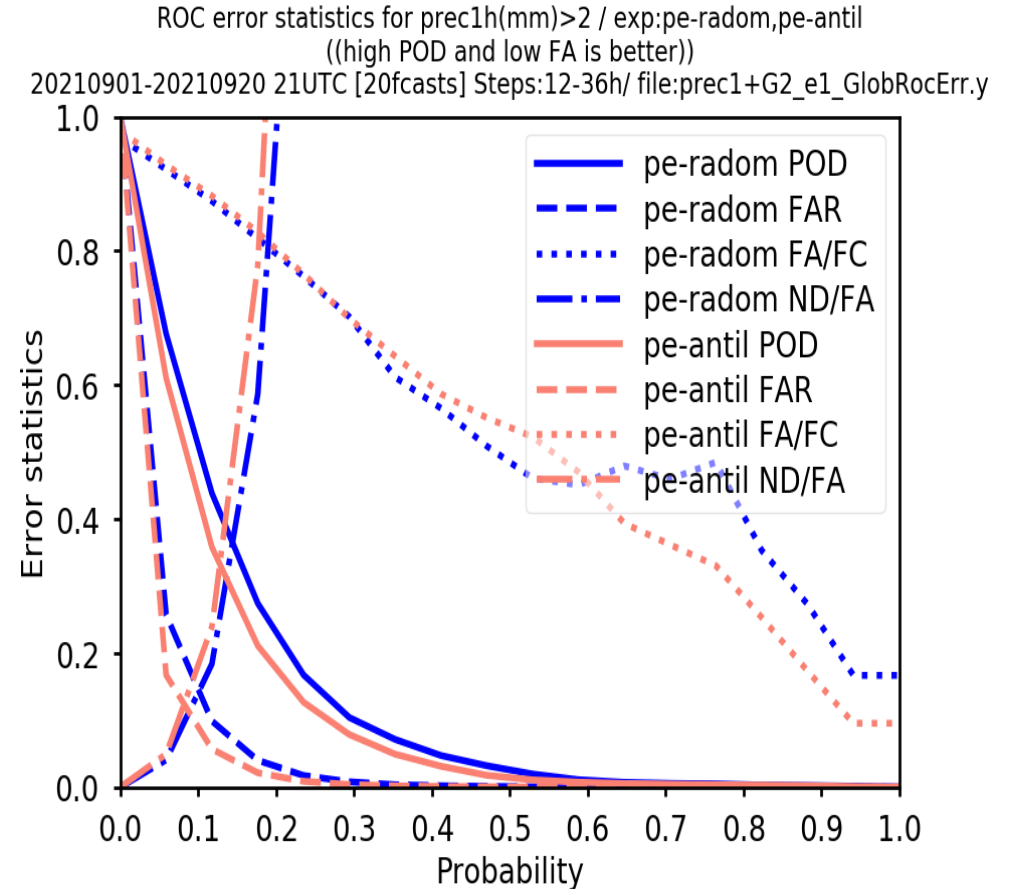
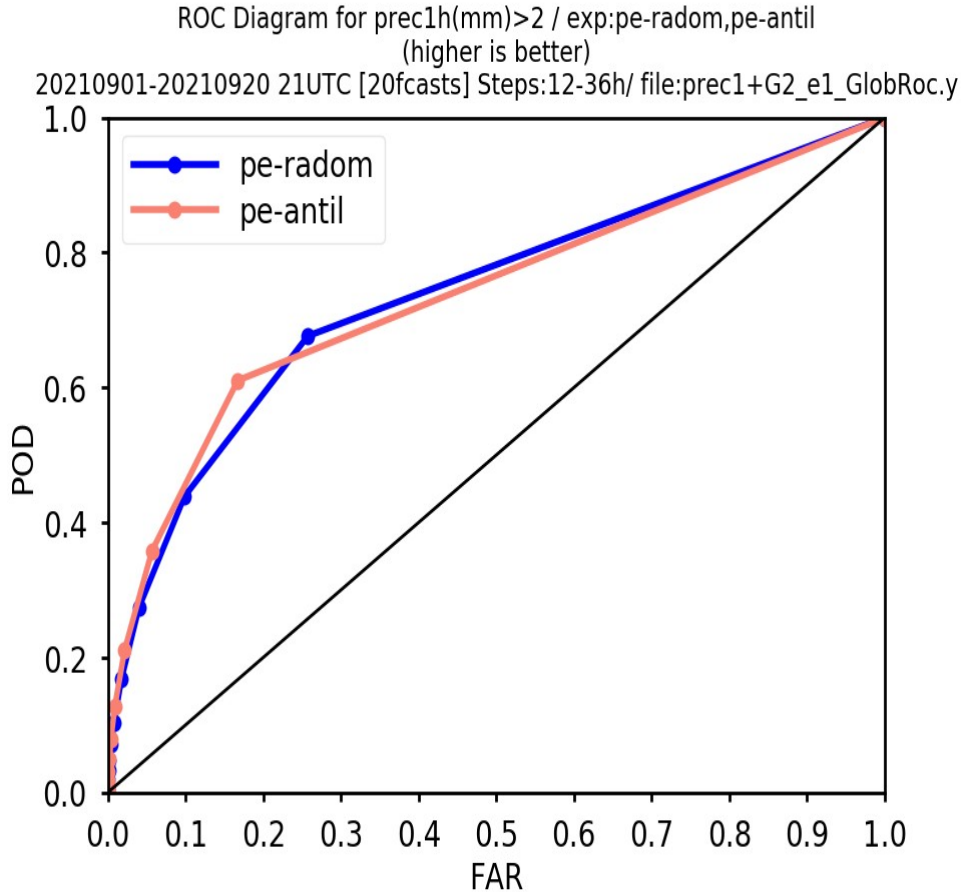
- chaque point correspond à un niveau de probabilité z (pas toujours montré!)
- si $z \sim 0$, on prévoit presque toujours $\rightarrow$ POD et FAR sont grands : point (1,1)
- si $z \sim 1$, on ne prévoit qu'à coup quasi-sûr $\rightarrow$ POD et FAR sont petits : point (0,0)
- prévision aléatoire : sur la diagonale ($FAR=POD$)
- prévision parfaite : si $(FAR, POD)=(0,1) \rightarrow$ **PE meilleure si plus près de (0,1)**

Interprétation :

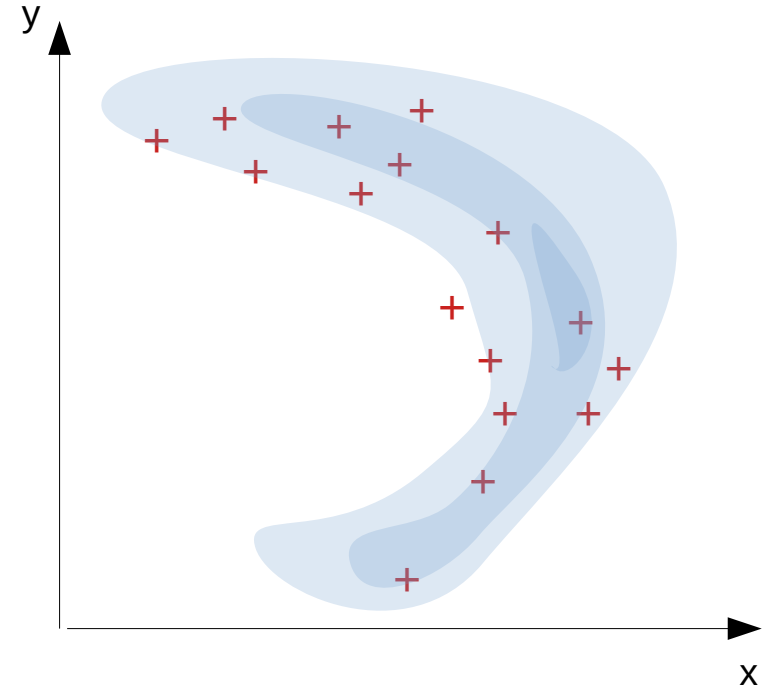
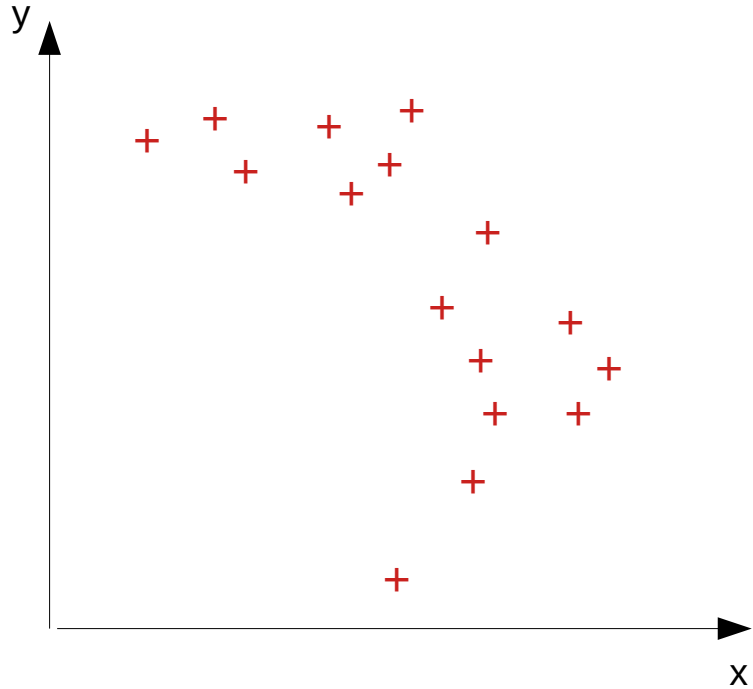
- mesure un **contenu en information** mais cache les pbs de fiabilité.
- ambigu : plusieurs manières de définir "plus près de (0,1)", par ex score ROCA
- mélange qualité des PDFs et qualité du modèle de prévi
- difficile à utiliser pour les événements rares



on peut "déplier" la courbe ROC en fonction des probabilités prévues (-> cf diagramme de fiabilité pour voir si elles sont bien calées)



2) The curse of dimensionality : cette distribution est-elle bimodale ? quel est l'état le plus probable ?



Produits des ensembles

- objets
- variables scalaires
- champs (cartes, scénarios)

Enjeux :

- résumer l'information
- éviter la surcharge cognitive
- corriger les biais de modélisation
- aide à la décision, ex: alertes (= risque + vulnérabilités + confiance)

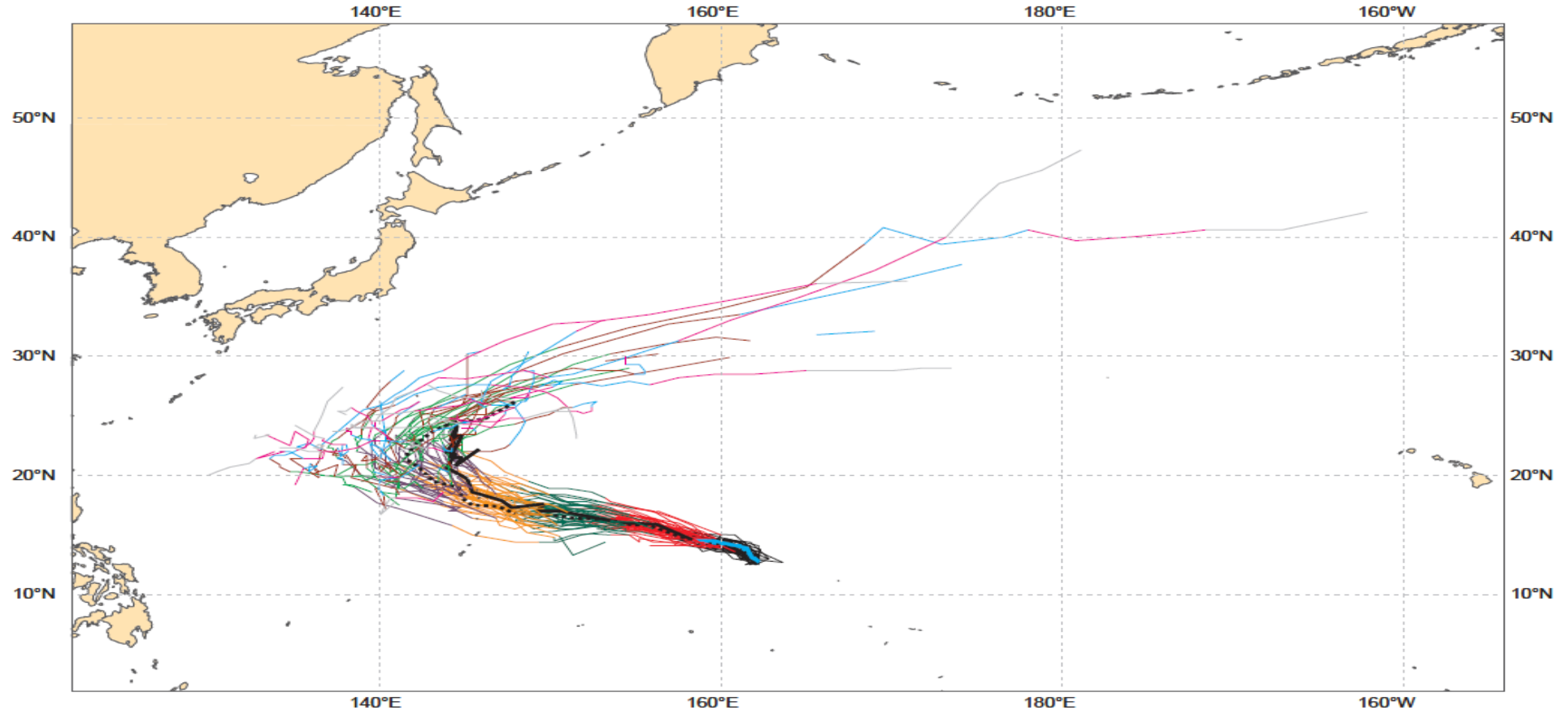
prévision cyclonique par un ensemble

Date 20151013 00 UTC @ECMWF

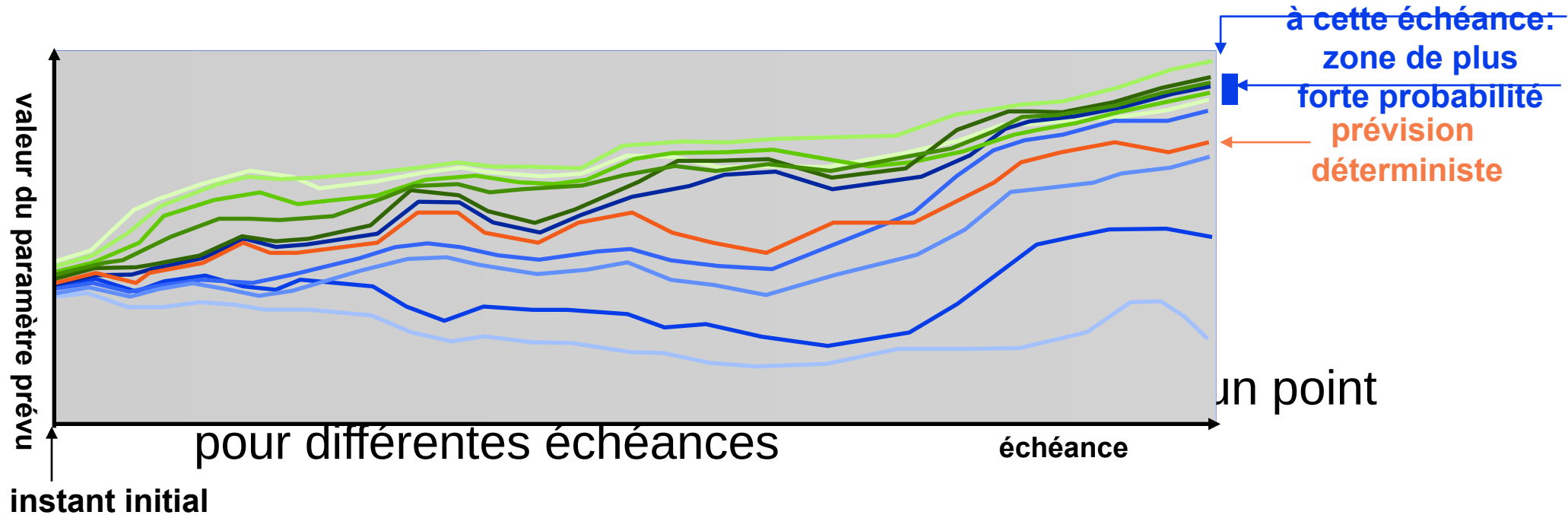
Individual trajectories for **25W** during the next **240** hours

tracks: **thick solid**=HRES; **thick dot**=CTRL; **thin solid**=EPS members [coloured]

0-24h 24-48h 48-72h 72-96h 96-120h 120-144h 144-168h 168-192h 192-216h 216-240h



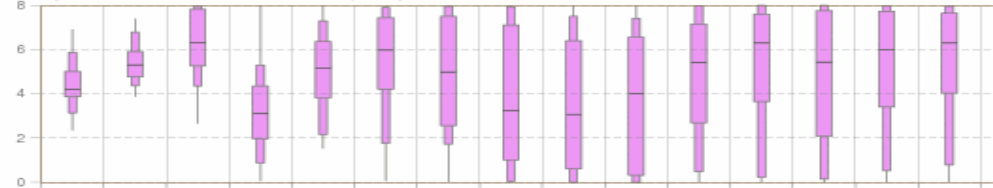
exemple de produit de prévision d'ensemble



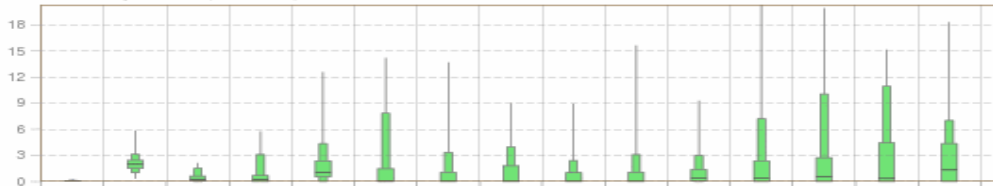
exemple de produit de prévision d'ensemble

EPS Meteogram
Toulouse 44.09°N 1.5°E (EPS land point) 140 m
Extended Range Forecast based on EPS Distribution Wednesday 25 January 2012 00 UTC

Daily mean of Total Cloud Cover (okta)



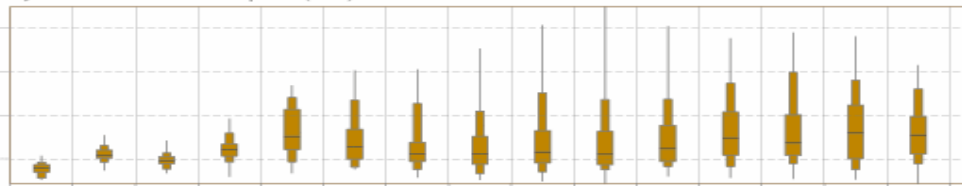
Total Precipitation (mm/24h)



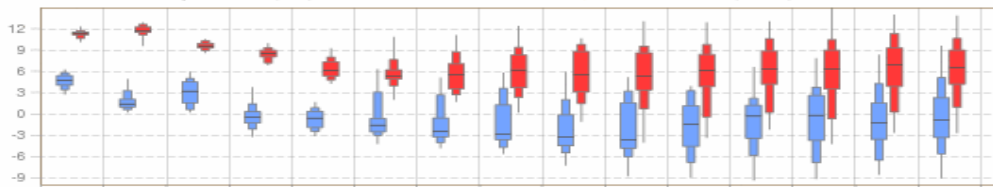
Daily distribution of 10m Wind Direction



Daily mean of 10m Wind Speed (m/s)



2m min/max temperature (°C) reduced to the station height from 230m (T319)



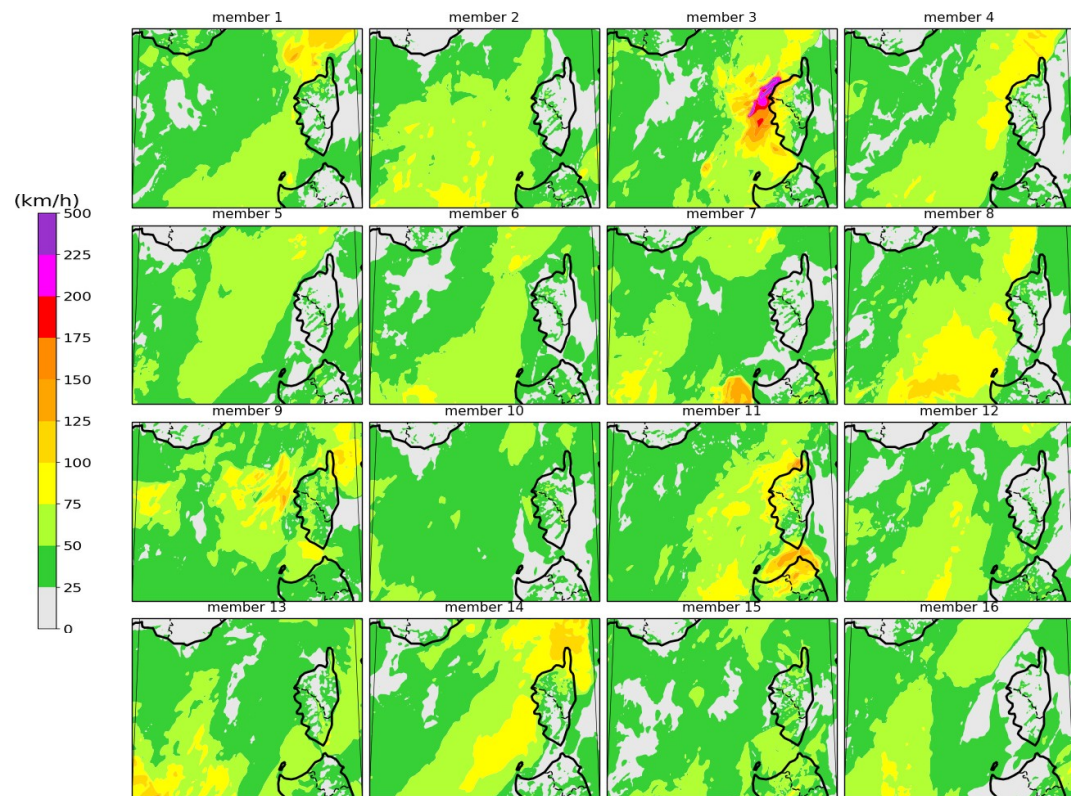
Wed 25 Thu 26 Fri 27 Sat 28 **Sun 29** Mon 30 Tue 31 Wed 1 Thu 2 Fri 3 Sat 4 **Sun 5** Mon 6 Tue 7 Wed 8
January 2012 February 2012



Magics++ 2.5.1

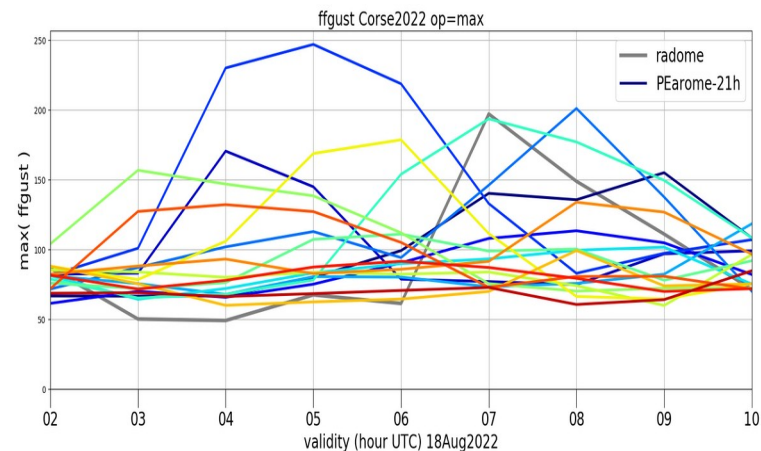
représentations d'une prévision d'ensemble

"vignettes" : tous les membres d'une PE
pour 1 échéance, 1 paramètre

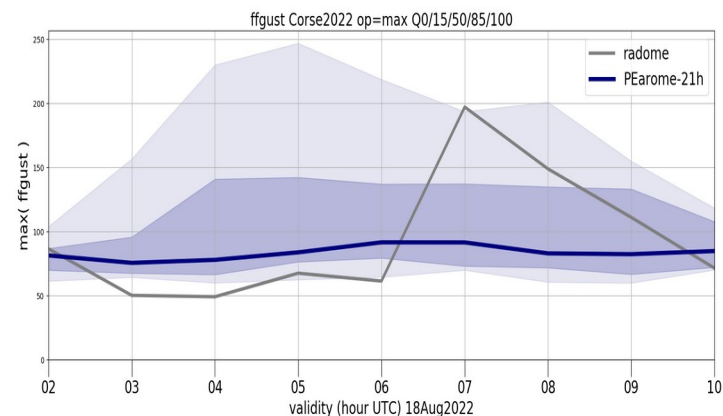


50/60

"spaghetti" : 1 courbe par membre (+ obs en gris)

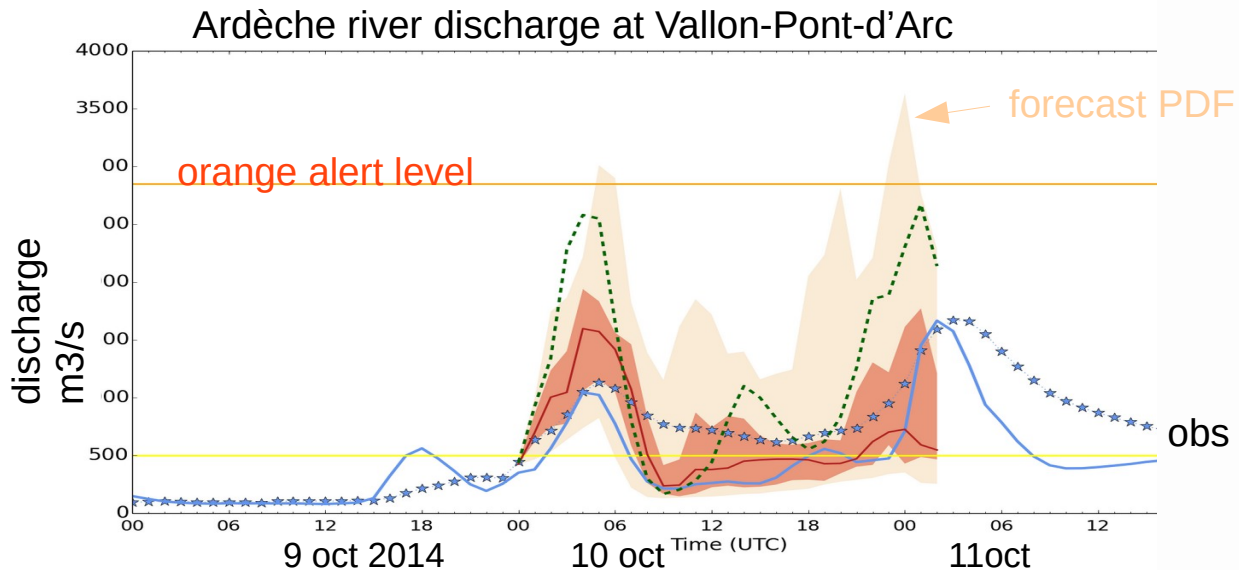
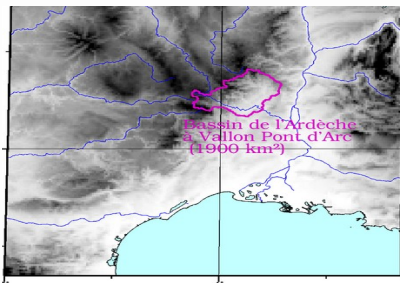
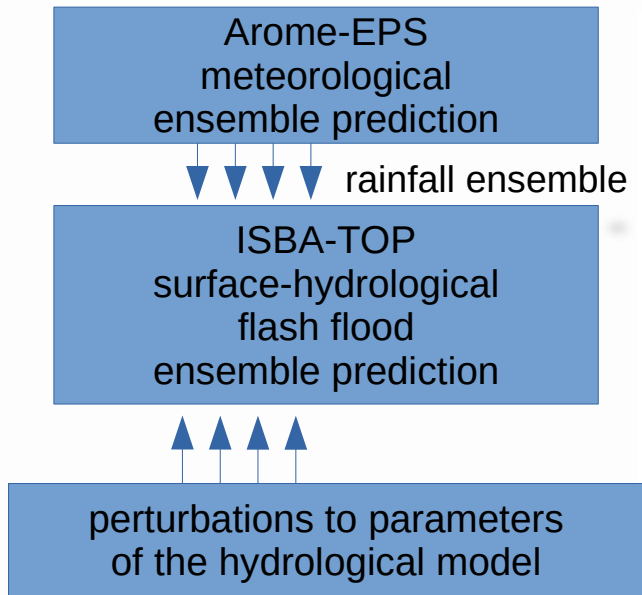


probas ou quantiles prévus



Uncertainty in flash-flood prediction

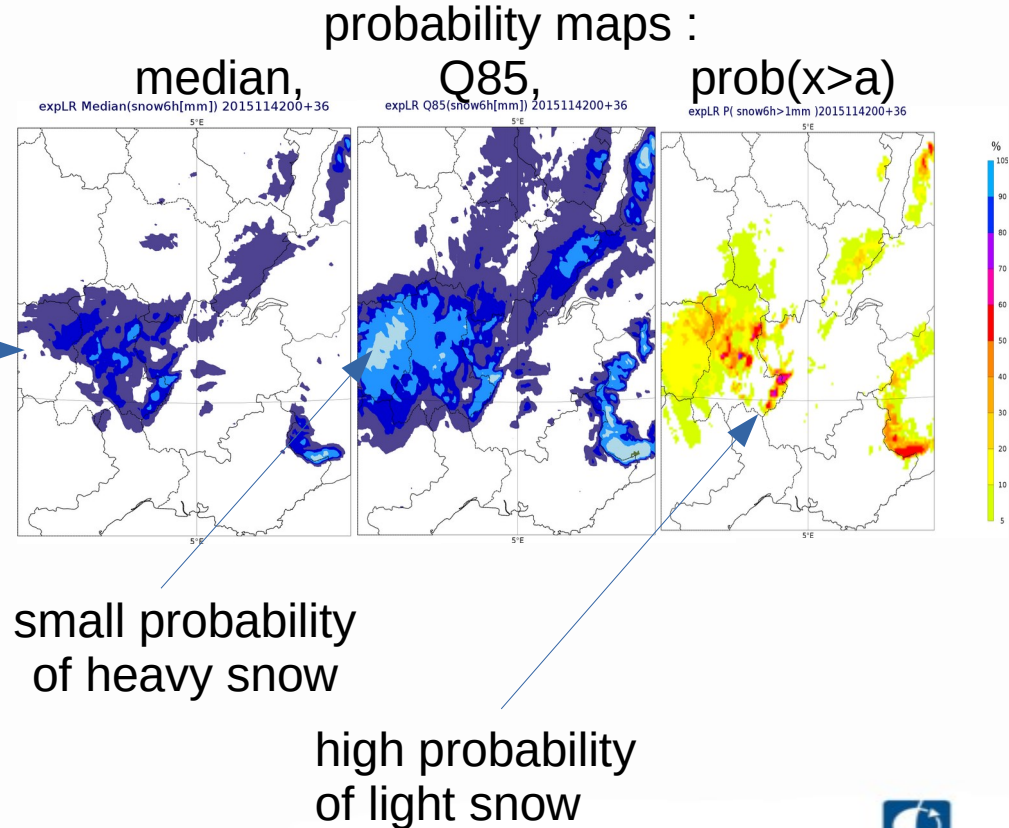
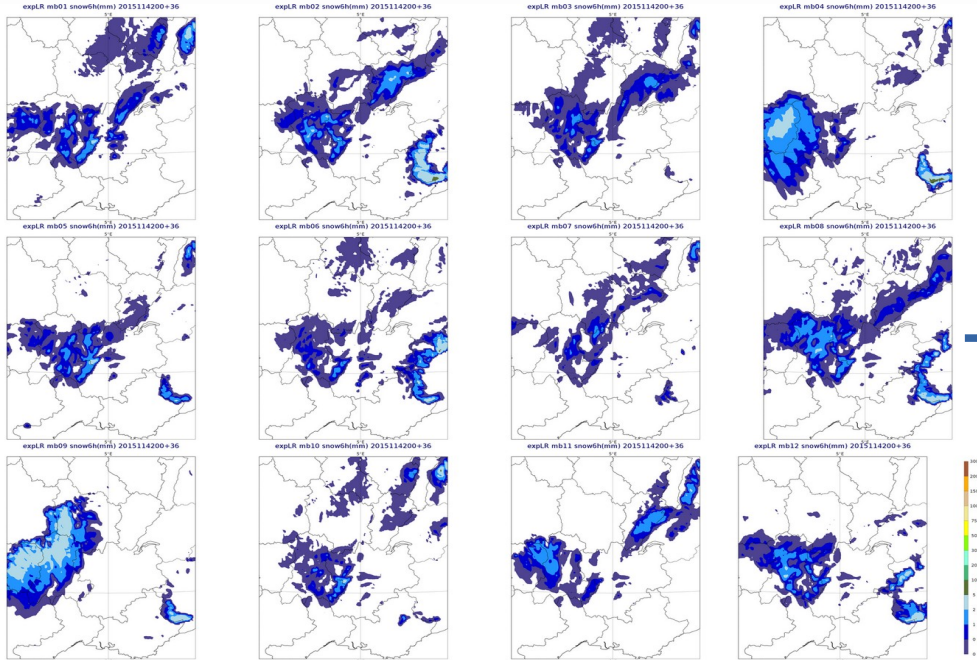
- various hydrological models, forced by atmospheric ensemble



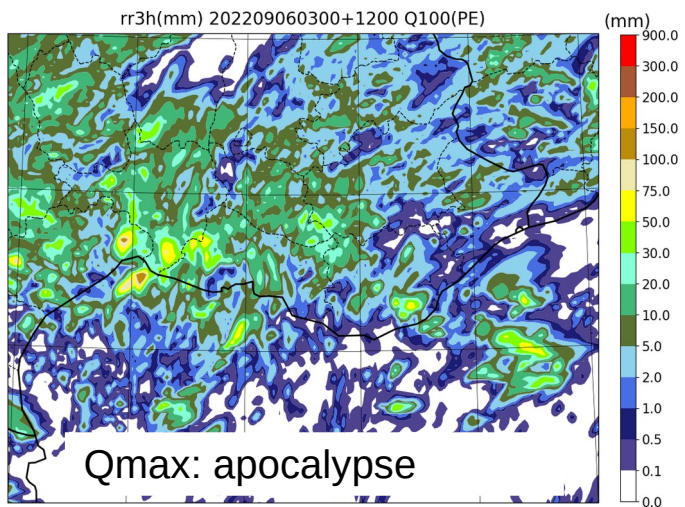
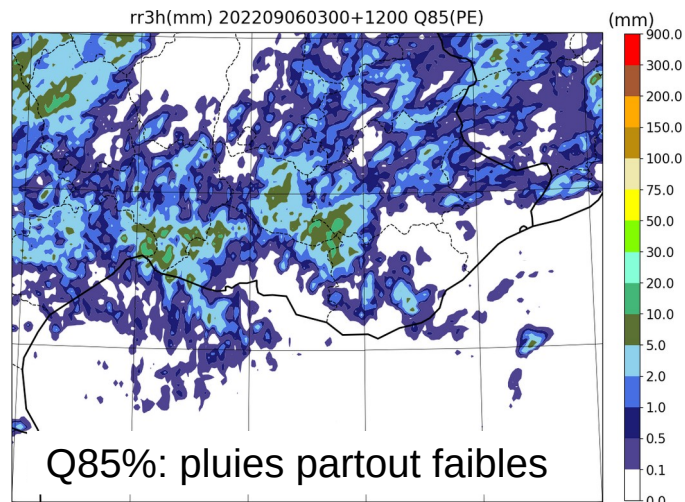
- ☆ measured discharge
- deterministic ISBA-TOP simulation driven by radar
- deterministic ISBA-TOP driven by Arome-France
- ISBA-TOP ensemble median, driven by Arome ensemble
- ISBA-TOP ensemble 25-75 % quantiles
- ISBA-TOP ensemble min-max

Arome-EPS : sample snow forecast

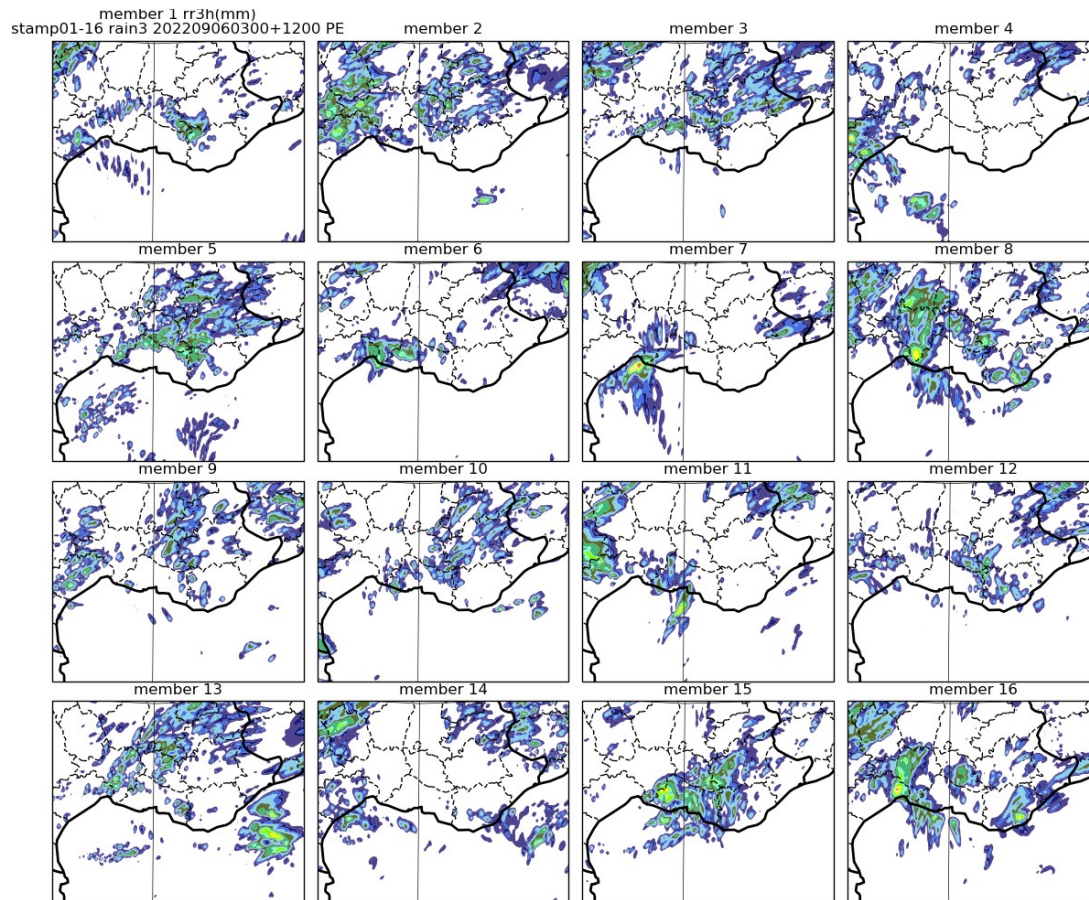
raw ensemble: 12 members



atténuation des structures physiques dans les quantiles



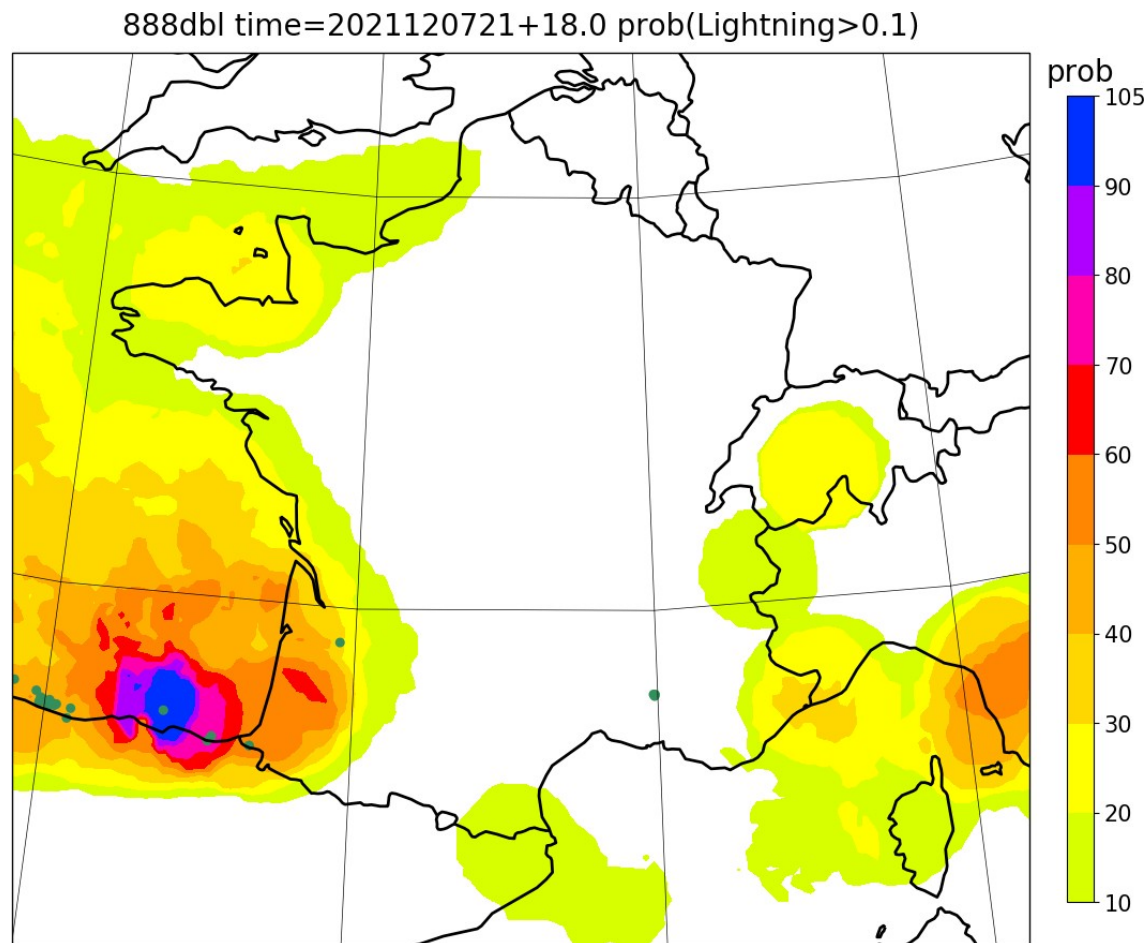
les membres : pluies intenses à positions incertaines



couleurs : probabilité d'orage
prévue en chaque point

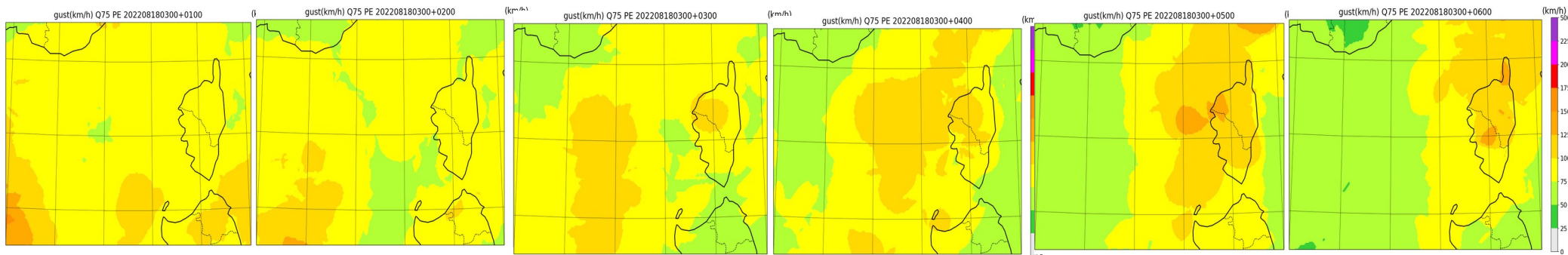
points verts : orages observés

...est-ce une bonne ou une
mauvaise prévision ?

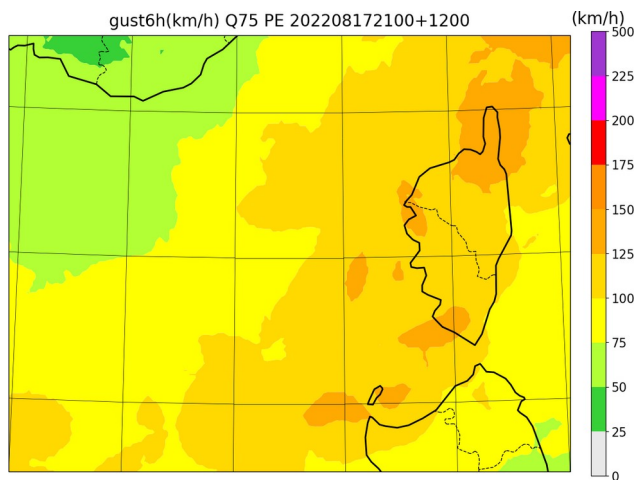


3) La non-additivité des probabilités et des quantiles :

le Q75 de vent horaire prévu, durant 6h consécutives :



le Q75 du vent sur ces 6h :



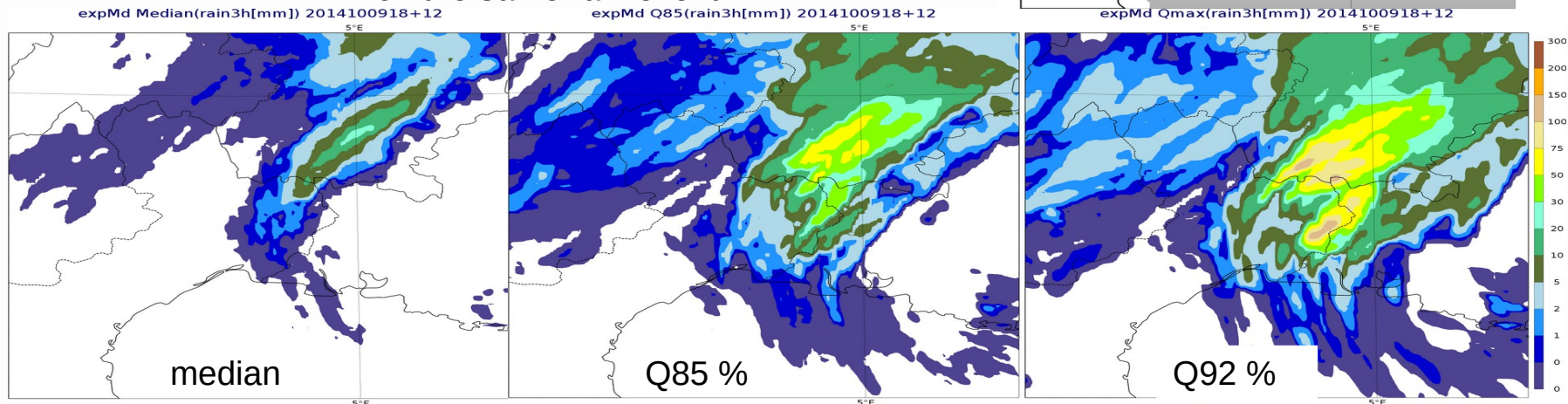
C'est plus que le max des Q75 horaires : il est très probable qu'il y aura un orage quelque part durant la période, mais on ne sait pas quand

warnings of dangerous small-scale events

case study :

- small-scale intense raining cells with high spread
- location/timing uncertainties can hide probabilities
- decision-making needs to choose a cost/loss ratio (~false alarm tolerance)
- the most likely forecast is not the most useful !

*Three points of view
on the same rain event :*

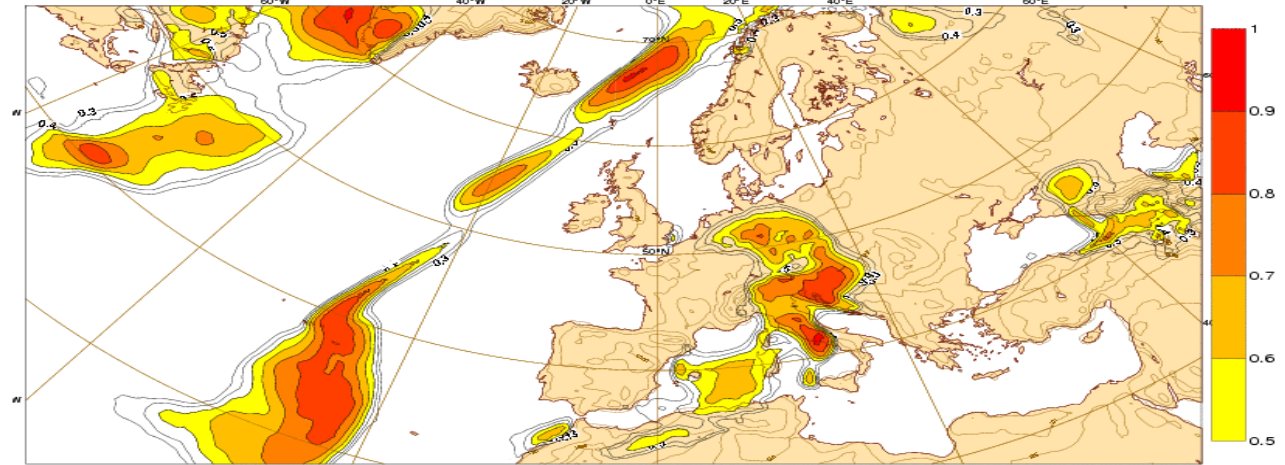


prévision d'extrêmes climatiques par des ensembles

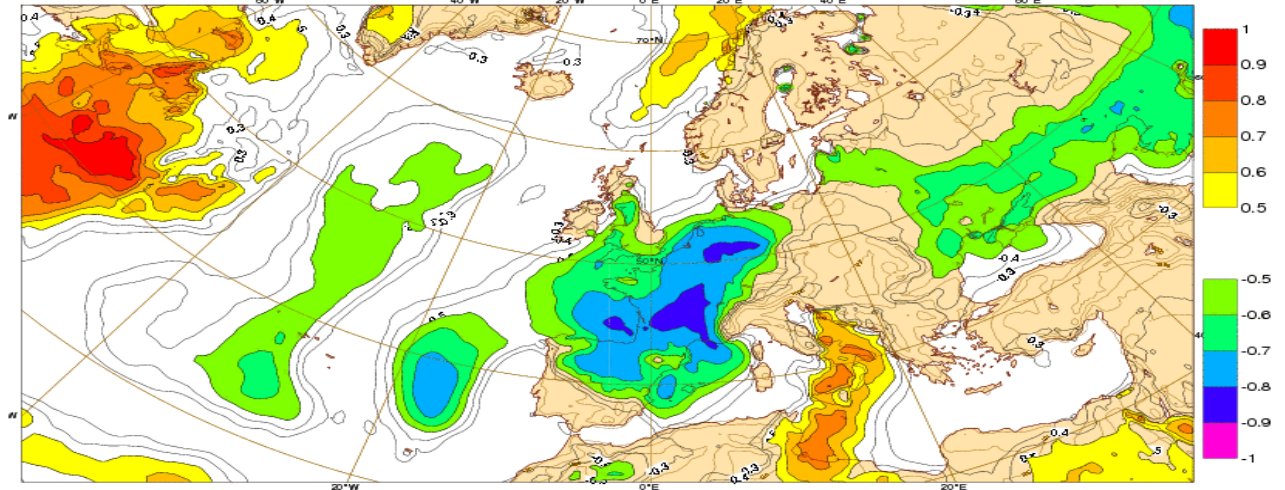
probabilité que la prévision soit dans un quantile extrême du climat en chaque point

Tuesday 13 October 2015 00UTC @ECMWF Extreme forecast index t+024-048 VT: Wednesday 14 October 2015 00UTC - Thursday 15 October 2015 00UTC
Surface: Total precipitation index

pluies



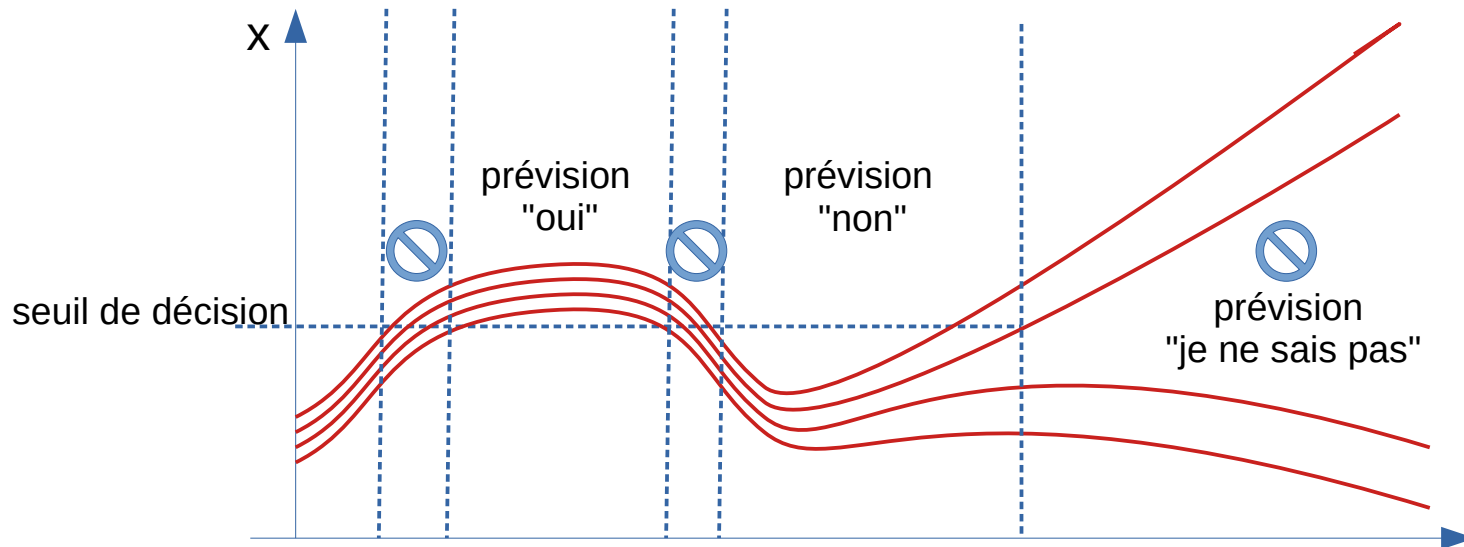
température



le problème d'horizon de prévisibilité : un événement peut être

- prévu "oui" si sa proba $>$ seuil ET le taux de fausses alarmes attendu est assez faible
- prévu "non" si sa proba $<$ seuil ET le taux de non-détection attendu est assez faible
- non-prévisible si (FA est trop fort) ou (TD est trop faible)... ou les deux :

Explique que les PE soient actuellement peu utilisées (aversion au risque)



Annexe: autres applications des prévisions d'ense "sensibilité ensembliste"

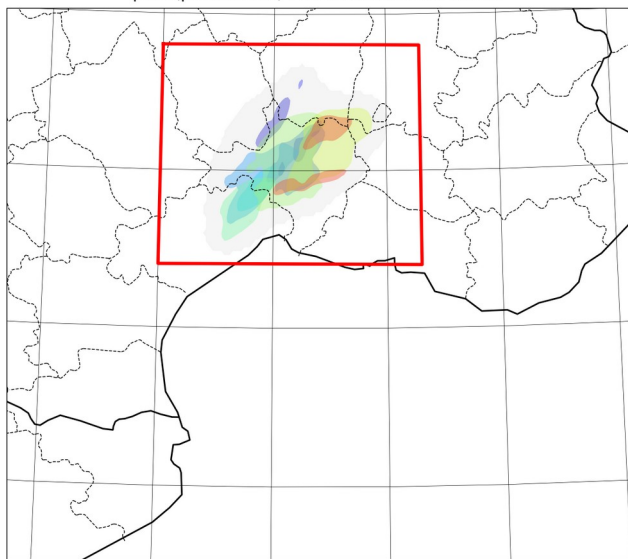
sensibilité des pluies à la vitesse du vent 5h avant

Target Object = highest rain fcasts

couleurs : membres rr3>30mm prévus par la PE
en gris : "objet rr3" par floutage des membres
en rouge : cadrage a priori

"paintball" plot

paint(prec3>30.0) PEarome 2021091321+13

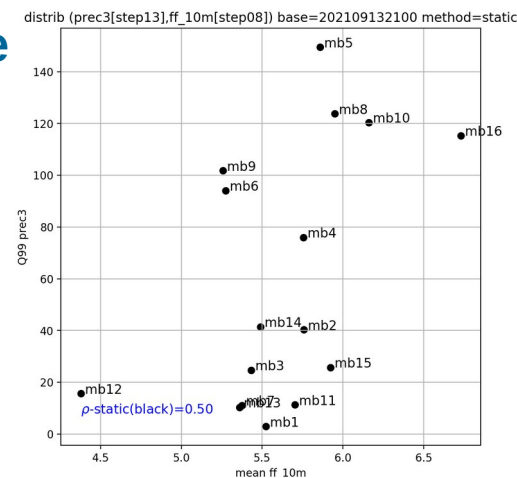
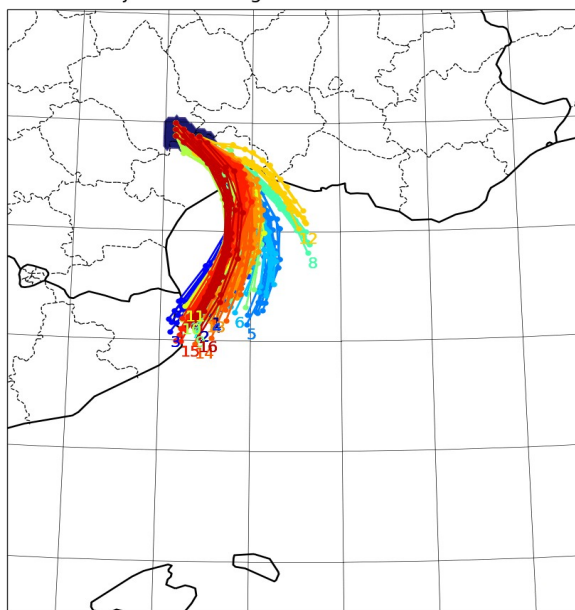


Source object = ensemble backtrajectories

bleu : cible = "objet rr3" prévu par la PE
grisé : origines des rétrotrajectoires U500m sur 5h
lignes : trajectoires colorisées par membre

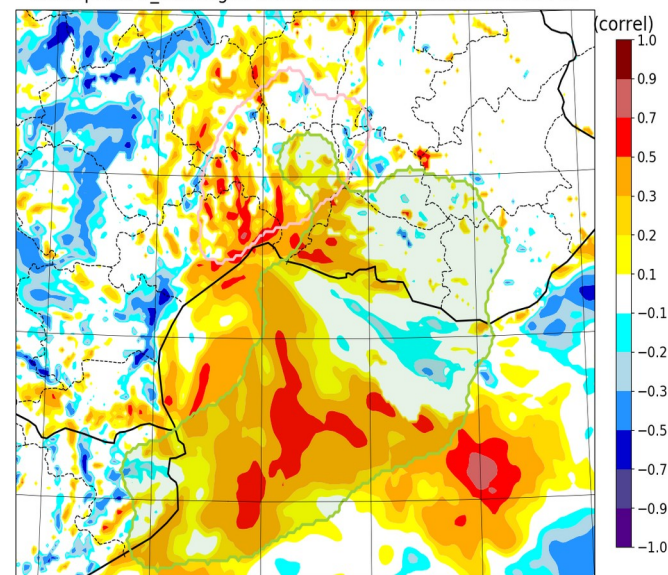
"jellyfish" plot

backtraj-mb01-16-lag5h PEarome 2021091321+6



sensitivity map / cross-correlation

correl prec3-ff_10m lag5h 16mb sourctime:202109132100+08



FIN