Utilisation des statistiques en climat : un panorama

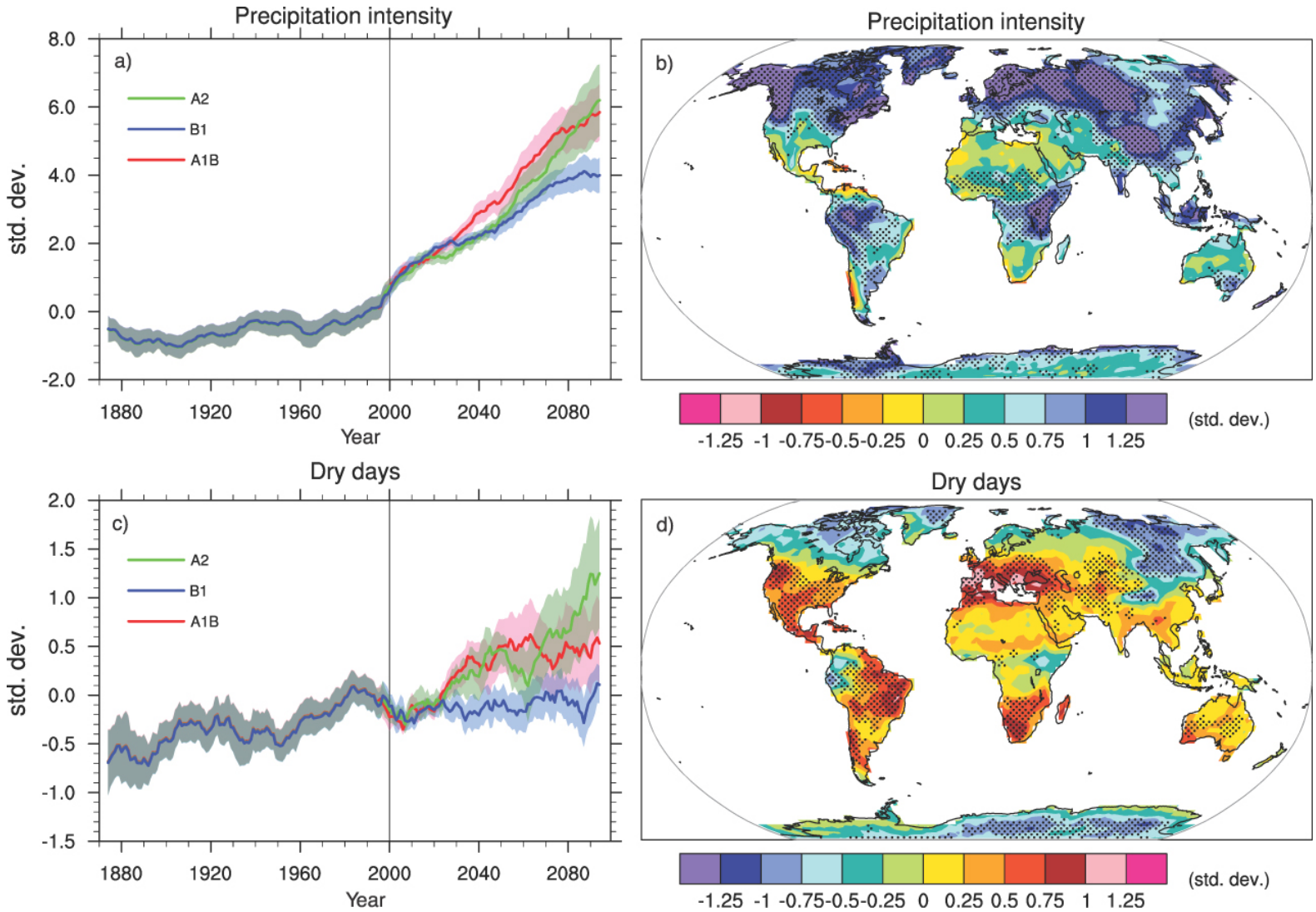
Julien Cattiaux
GAME | CNRS/Météo-France
Toulouse

ENSAI
5 Décembre 2014

Retrouver ce cours sur ma page web : http://www.cnrm-game.fr/spip.php?article629

Mail : julien.cattiaux@meteo.fr | Twitter : @julienc4ttiaux

Statistiques et climat ? Une figure du GIEC...



Source : IPCC AR4 (2007) Fig. 10.18.

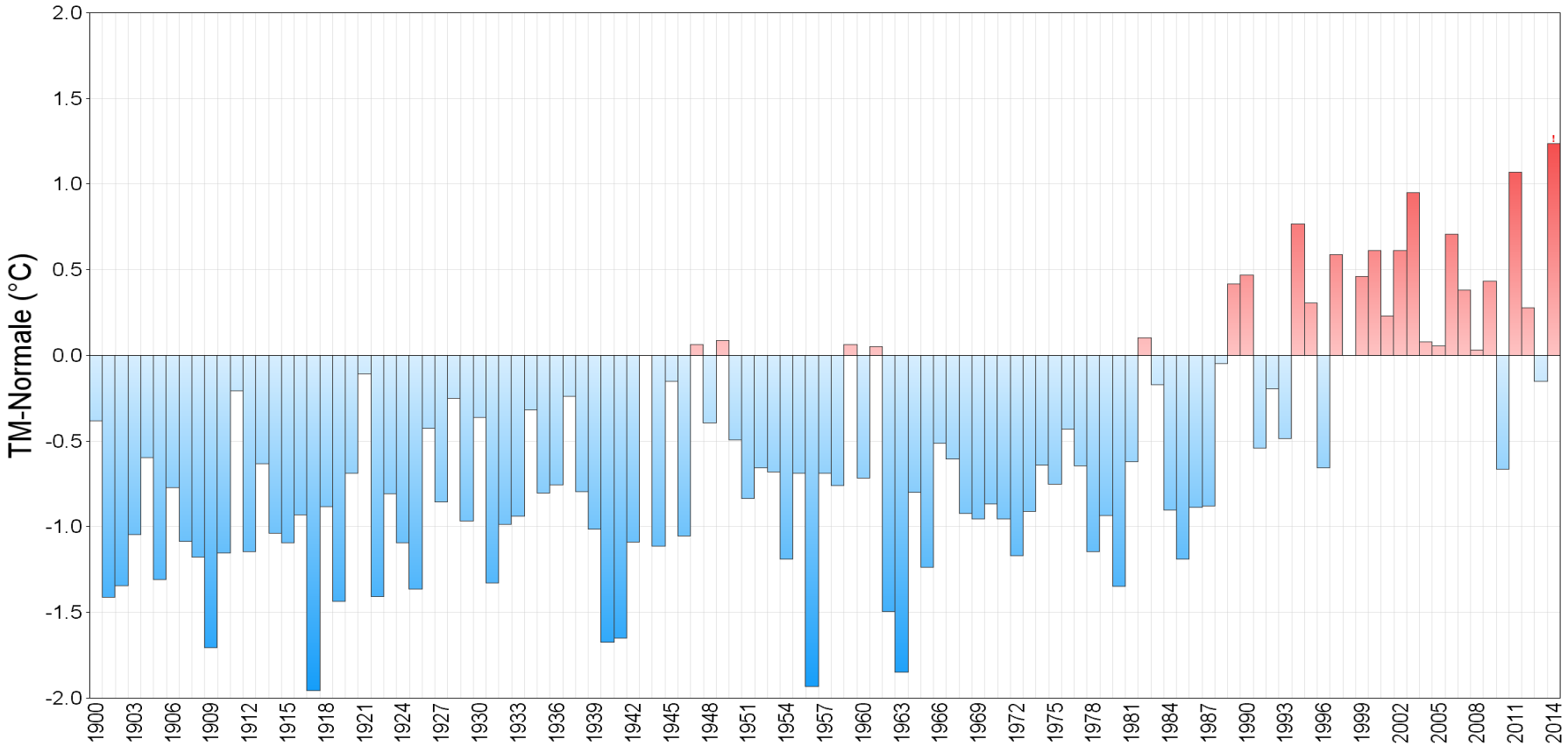
Intro empruntée au cours de Pascal Yiou (LSCE).

Statistiques et climat ?...et sa légende

Figure 10.18. Changes in **extremes** based on multi-model simulations from nine global coupled climate models, adapted from Tebaldi et al. (2006). (a) Globally **averaged** changes in precipitation intensity (defined as the annual total precipitation divided by the number of wet days) for a low (SRES B1), middle (SRES A1B) and high (SRES A2) scenario. (b) Changes in **spatial patterns** of simulated precipitation intensity between two **20-year means** (2080-2099 minus 1980-1999) for the A1B scenario. (c) Globally averaged changes in dry days (defined as the annual **maximum** number of consecutive dry days). (d) Changes in spatial patterns of simulated dry days between two 20-year means (2080-2099 minus 1980-1999) for the A1B scenario. Solid lines in (a) and (c) are the 10-year **smoothed** multi-model **ensemble means**; the envelope indicates the ensemble mean **standard deviation**. Stippling in (b) and (d) denotes areas where at least five of the nine models concur in determining that the change is **statistically significant**. Extreme indices are calculated only over land following Frich et al. (2002). Each model's **time series** was **centred** on its 1980 to 1999 average and **normalised** (rescaled) by its standard deviation computed (after **detrending**) over the period 1960 to 2099. The models were then **aggregated** into an ensemble average, both at the global and at the grid-box level. Thus, changes are given in units of standard deviations.

Une série climatique...

Température annuelle moyenne France centrée sur 1981–2010



Données : Météo-France.

. . . et un tas de questions

- Qualité des données ?

$\longrightarrow$ Homogénéisation.

- Propriétés statistiques et dépendance spatio-temporelle ?

$\longrightarrow$ Analyse spectrale, analyse en composantes principales, classification automatique.

- Tendances ? Liens avec d'autres variables ?

$\longrightarrow$ Tests d'hypothèses.

- Les tendances traduisent-elles un changement climatique ? Si oui, quelles causes ?

$\longrightarrow$ Détection et attribution.

- Comment anticiper la suite ?

$\longrightarrow$ Prévisions et projections.

- Comment décrire les événements extrêmes ?

$\longrightarrow$ Théorie des records, théorie des extrêmes.

Plan

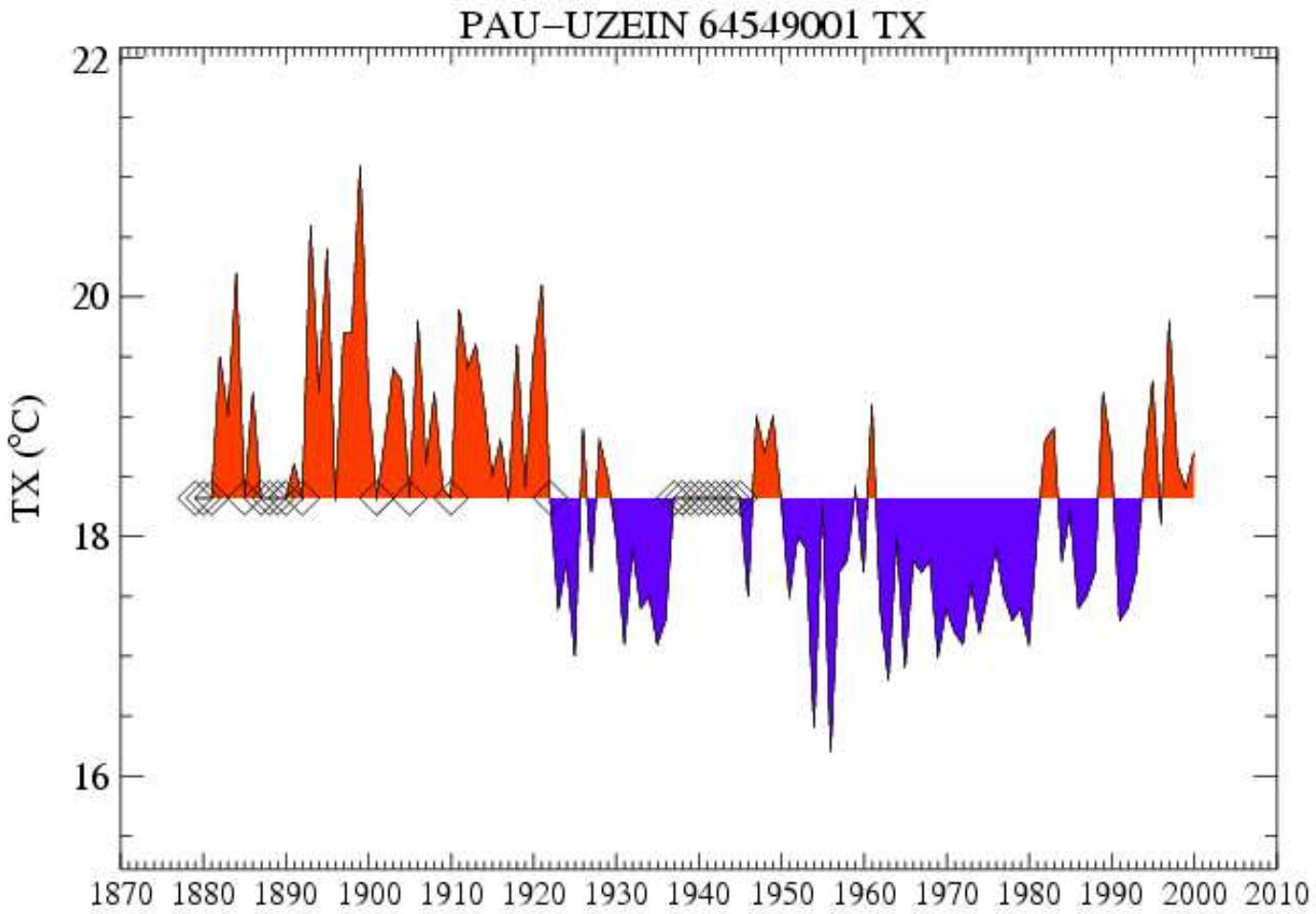
Plan

Le problème 1/2

- **1912 : Ecole Normale de Lescar**

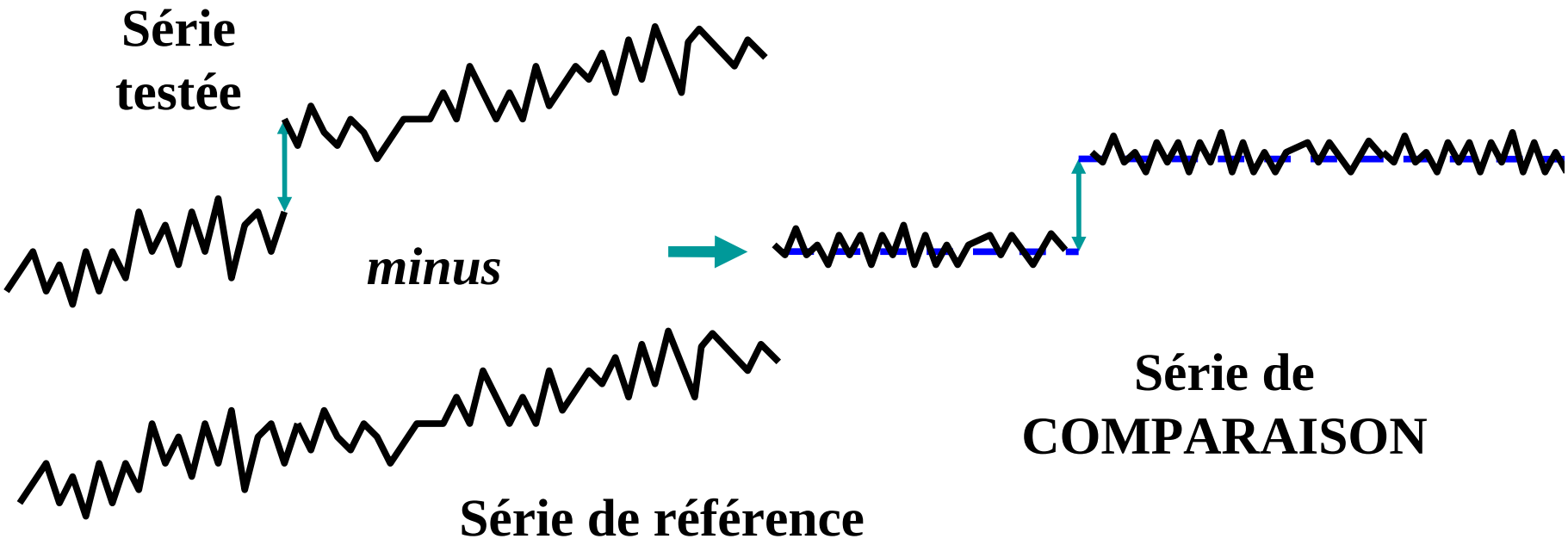
- **2006 : Aérodrome de Pau-Uzein**

Le problème 2/2



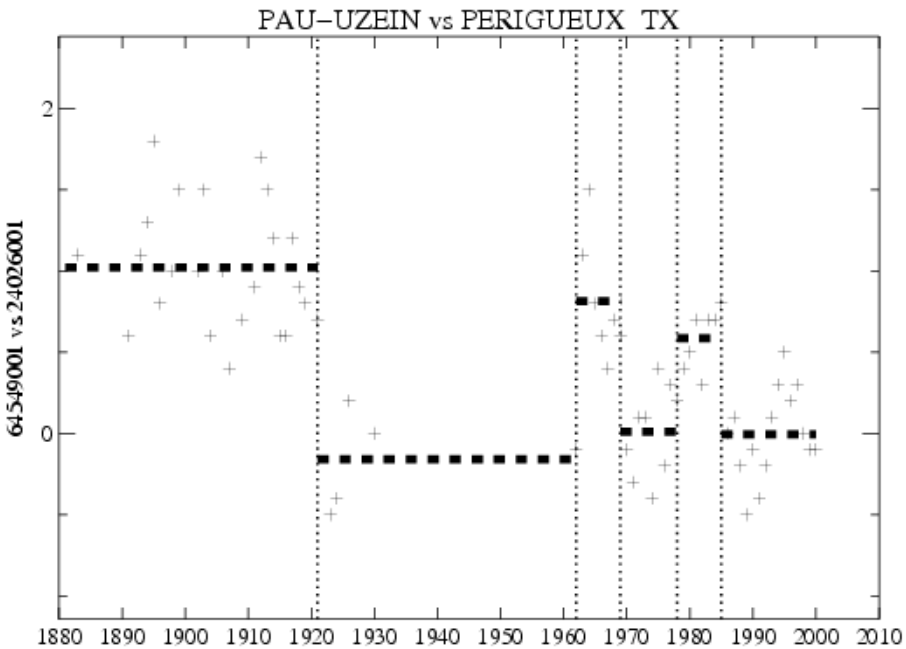
La méthode 1/2

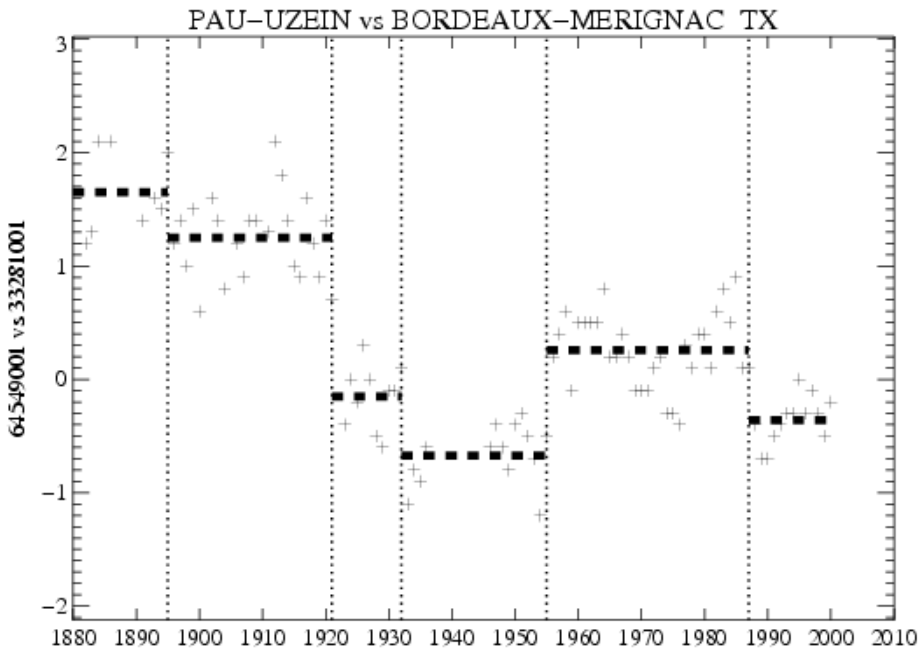
- PRINCIPE : enlever le signal climatique pour mettre en évidence les ruptures artificielles



La méthode 2/2

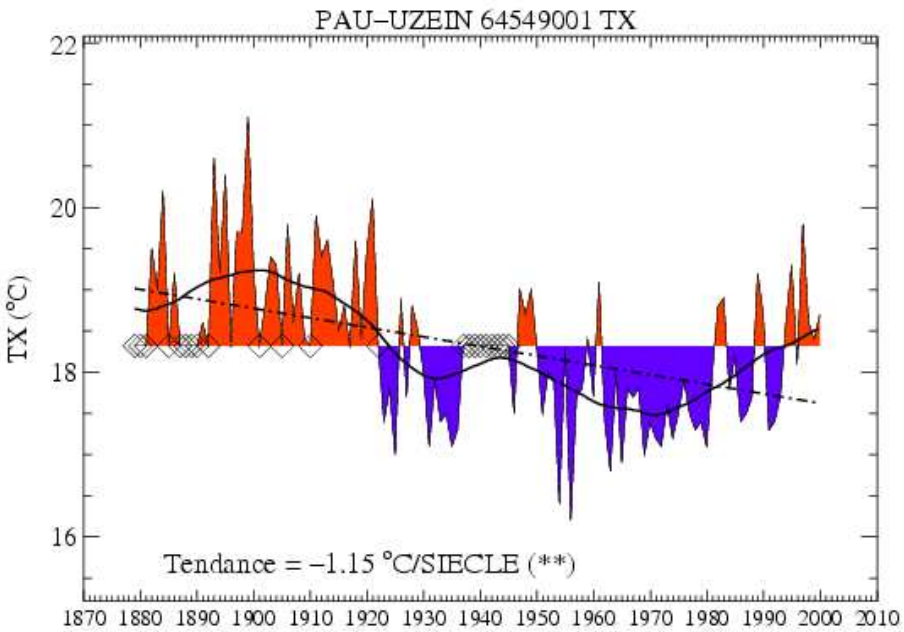
- Algorithme programmation dynamique + vraisemblance pénalisée
- Comparaisons multiples des séries non-homogènes, historique des postes



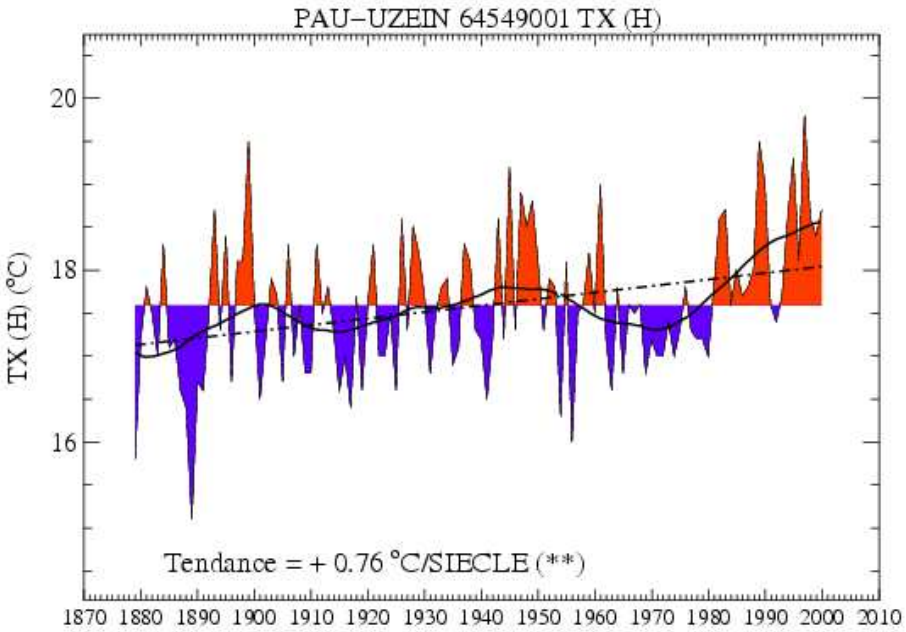


Le résultat 1/2

- **« AVANT »**

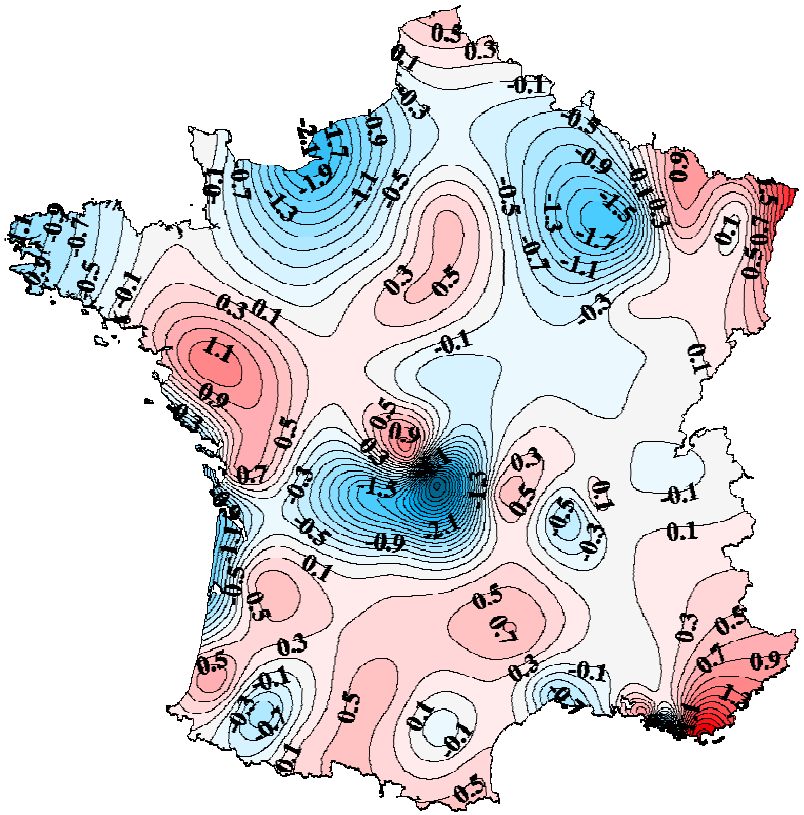


- **« APRES »**

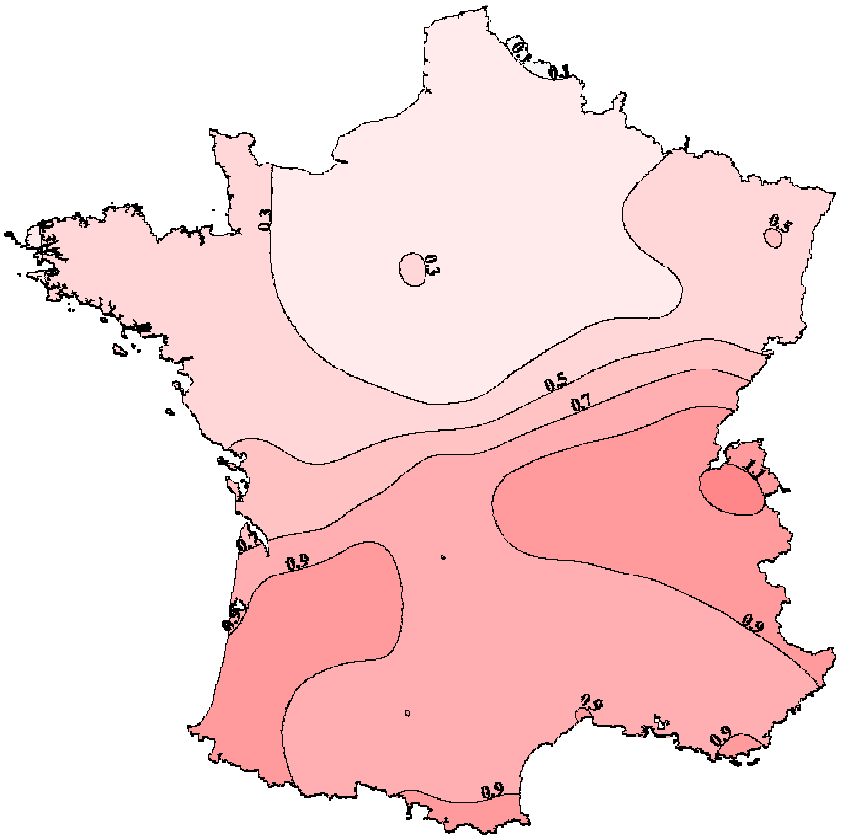


Le résultat 2/2

- **« AVANT »**



- **« APRES »**



Plan

Analyse spectrale

$X(t)$ variable aléatoire (température, précipitation, etc.).

$X(t)$ a-t-elle des échelles de temps caractéristiques ?

Analyse spectrale

$X(t)$ variable aléatoire (température, précipitation, etc.).

$X(t)$ a-t-elle des échelles de temps caractéristiques ?

1. Transformée de Fourier, passage dans le domaine fréquentiel:

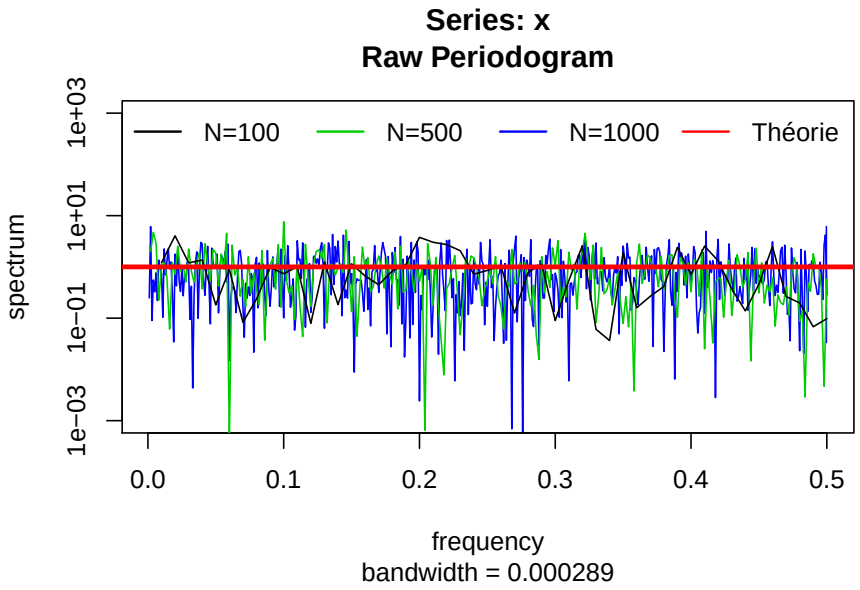
$$\hat{X}(f) = \int_t X(t) e^{-2i\pi ft} \, dt$$

2. Spectre de puissance:

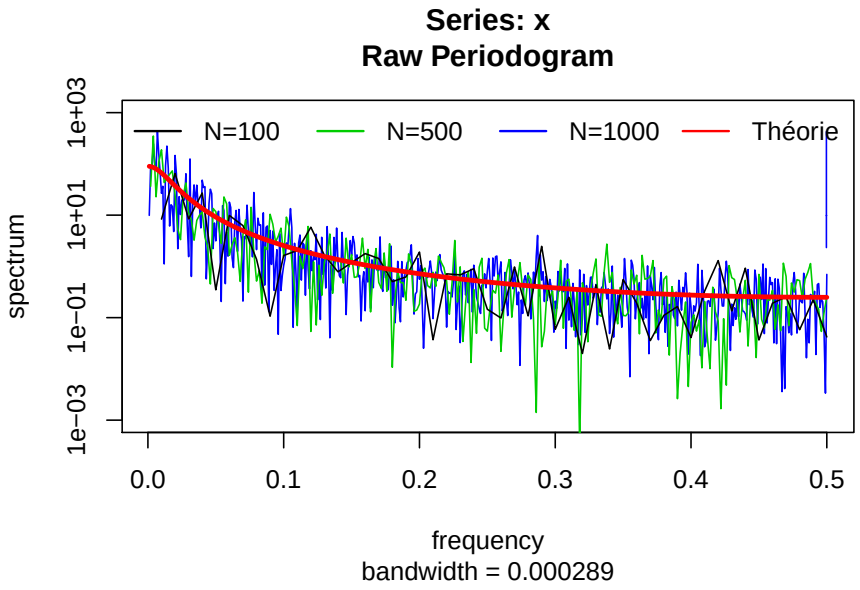
$$P_X(f) = \left| \hat{X}(f) \right|^2$$

Propriété : $P_X(f)$ est maximal à la fréquence f_0 quand $X(t)$ est périodique de période $1/f_0$.

Analyse spectrale Exemple des bruits blanc/rouge



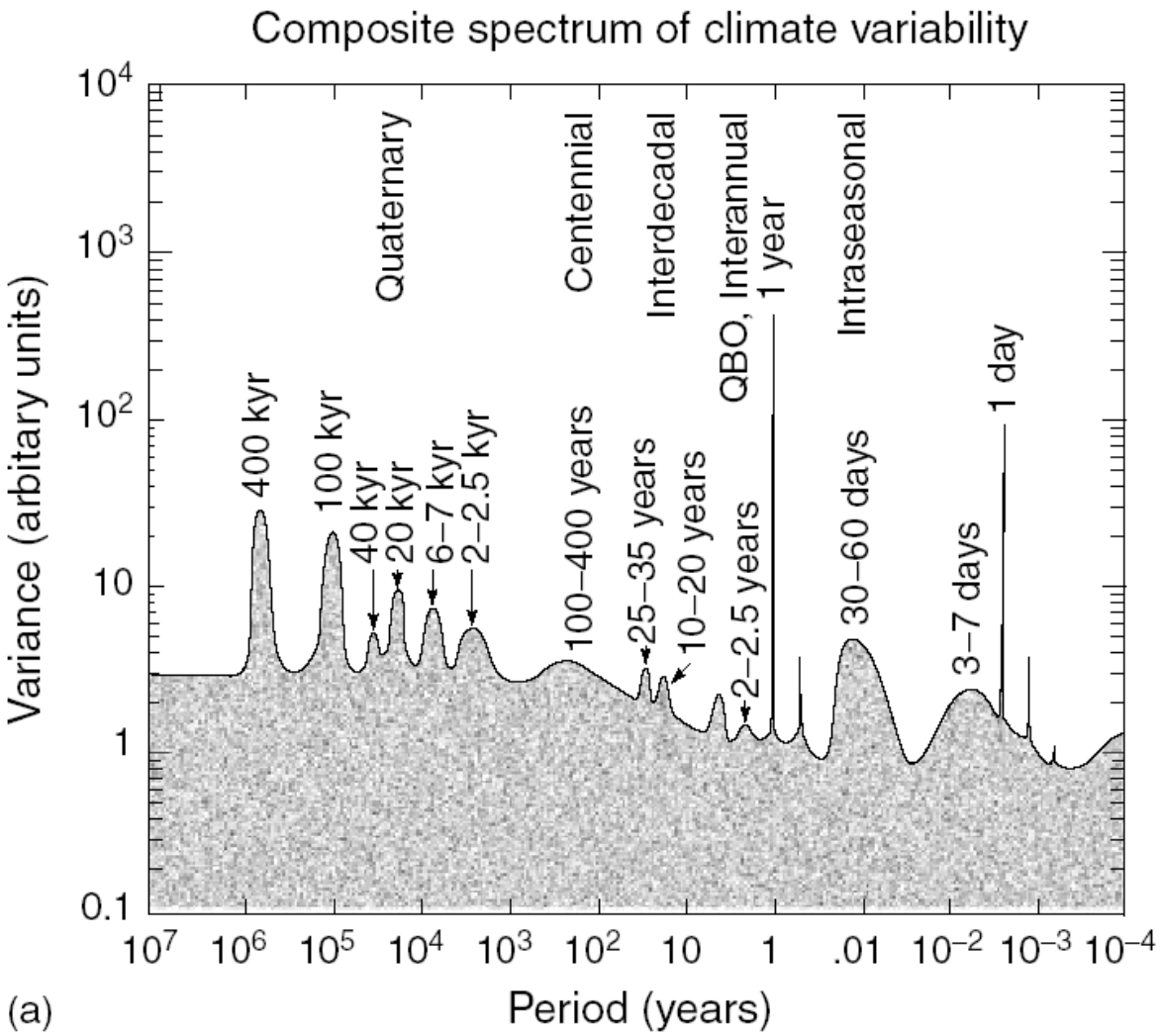
Bruit blanc : spectre constant.



Bruit rouge (exemple $\alpha = 0.9$) :

$$P_X(f) \approx \frac{\alpha}{1 - 2\alpha\cos 2\pi f + \alpha^2}$$

Analyse spectrale Exemple *idéalisé* du climat



Spectre de la température moyenne de surface.

Attention !
Une telle série temporelle n'existe pas.

Source : Ghil (2002).

Analyse spectrale Limites

- Principale limite de la méthode :
  - la variance de l'estimateur croît avec le nombre d'observations ;
  - la résolution (pouvoir de séparer deux pics proches) aussi.

Il y a donc un compromis à trouver.

- Principale limite de l'application au climat :

Aucune raison de penser que la variabilité du climat est cyclique, excepté aux périodes correspondant aux cycles des forçages externes (e.g., forçage astronomique).

Analyse en composantes principales

$X(s,t)$ champ $N_t \times N_s$ (température, précipitation, etc.).

Quels sont les principaux modes de variabilité spatio-temporelle de $X(s,t)$?

Analyse en composantes principales

$X(s,t)$ champ $N_t \times N_s$ (température, précipitation, etc.).

Quels sont les principaux modes de variabilité spatio-temporelle de $X(s,t)$?

▶ Séparation espace-temps (s et t) par décomposition en valeurs propres de la matrice de covariance de X (notée C, symétrique et semi-définie positive) :

$$X(s,t) = \sum_{k=1}^{K} p_k(t)\ e_k(s)\ .$$

Composantes Principales (PCs)

- coefficients temporels pour recombiner les $E_k(s)$;
- variances = valeurs propres de C ;
- non-corrélées, $\mathrm{cor}(p_k, p_{k'}) = \delta_{kk'}$.

Fonctions Orthogonales Empiriques (EOFs)

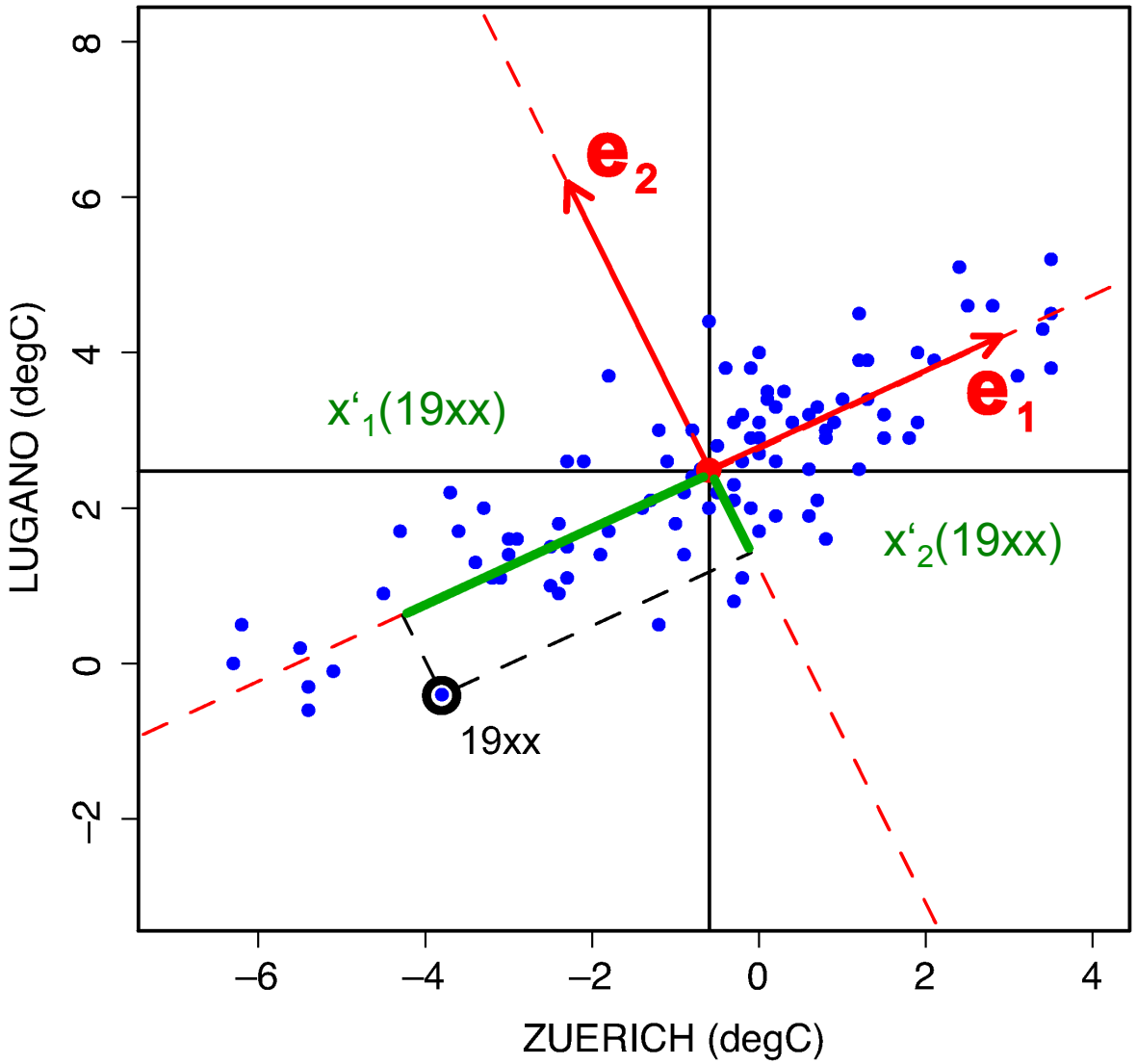
- vecteurs spatiaux (cartes) à partir desquels les données sont combinées ;
- vecteurs propres de C ;
- orthonormales, $e_k' e_{k'} = \delta_{kk'}$.

Analyse en composantes principales Calcul

$$X(s,t) = \sum_{k=1}^{K} p_k(t)\ e_k(s)\ .$$

1. On cherche e_1 unitaire qui maximise la variance de X expliquée par e :
$$\mathrm{Max}\left(V_e^X\ =\ e'\,X'\,X\,e\ =\ e'\,C\,e\right)\ \ \mathrm{tq.}\ \ e'\,e = 1\ .$$
2. Algèbre : si on note $\lambda_1 > \lambda_2 > \ldots > \lambda_N$ les valeurs propres de C, on montre que e_1 est le vecteur propre de C associé à λ_1 :
$$C\,e_1 = \lambda_1 e_1\ .$$
3. On cherche e_2 de façon similaire (contraintes $e'\,e = 1$ et $e'\,e_1 = 0$).
4. On montre que e_2 est le vecteur propre de C associé à λ_2, et ainsi de suite.
5. On en déduit les PCs par projections orthogonales :
$$\forall k\ \ p_k = X\,e_k\ .$$

Analyse en composantes principales Exemple 2D

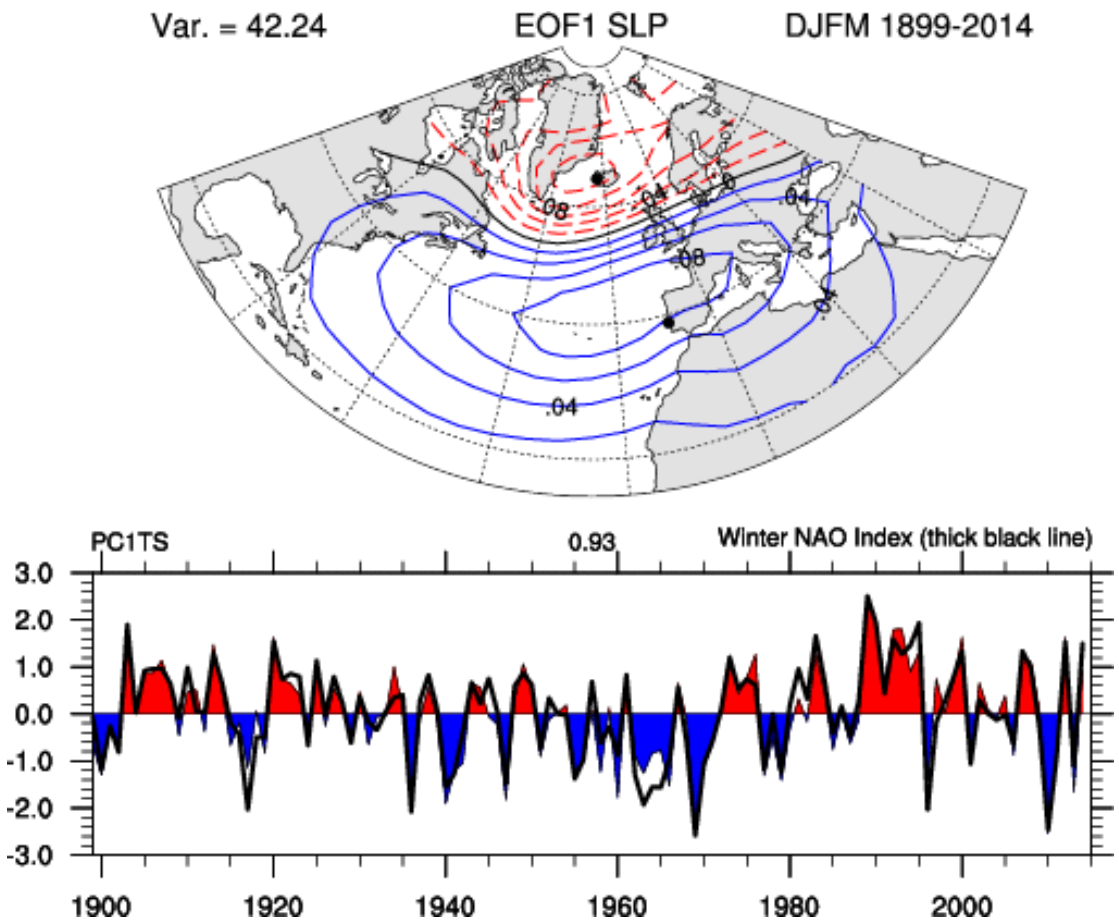


- e_1 direction *principale* du nuage de points ;
- e_2 direction orthogonale ;
- p_1, p_2 les coordonnées dans la base (e_1, e_2) (*espace des phases*).

© Cours de Cristoph Frei (ETH Zürich).

Analyse en composantes principales Illustration classique

Vecteurs e_1 et p_1 de la pression de surface hivernale : la NAO.



© Site de Jim Hurrell.

Analyse en composantes principales Remarques générales

- Hiérarchisation des modes de variabilité :
  - réduction de la dimension tout en conservant suffisamment de variance ;
  - signification *physique* éventuelle des premiers modes.

- Utilisations classiques en climat :
  - étude de la variabilité spatio-temporelle d'une variable ;
  - recherche d'un signal commun à n variables.

- Limites de l'application au climat :
  - sur-interprétation éventuelle ;
  - n'explique pas tout. Ex: la NAO n'explique "que" 25 % de la variance des températures européennes hivernales.

Plus d'infos : cours de Pascal Yiou (LSCE) et Christoph Frei (ETH Zürich).

Classification

$X(s,t)$ variable aléatoire (température, précipitation, etc.).

$X(s,t)$ s'agglomère-t-elle autour d'un petit nombre d'états *préférentiels* ?

Classification

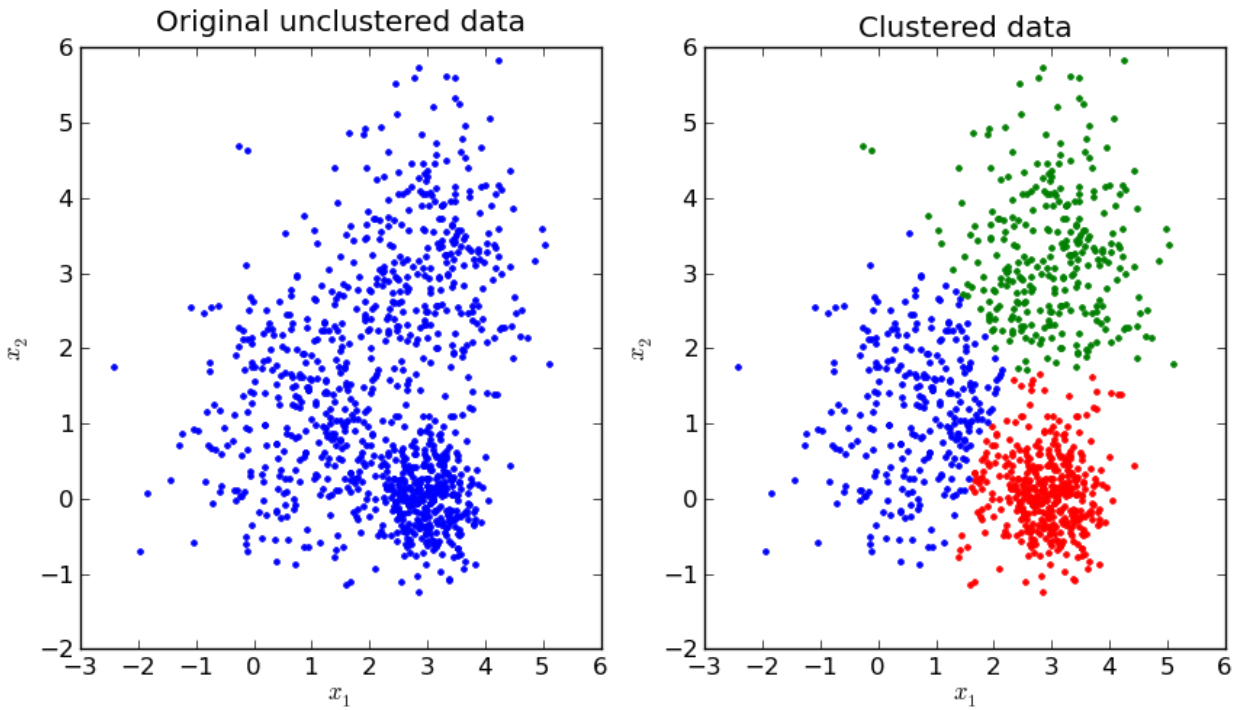
$X(s,t)$ variable aléatoire (température, précipitation, etc.).

$X(s,t)$ s'agglomère-t-elle autour d'un petit nombre d'états *préférentiels* ?

- On cherche à regrouper les $X(s,t)$ en k classes en tâchant de :
  - minimiser la variance intra-classes ;
  - maximiser la variance inter-classes ;
  - optimiser le nombre de classes.
- Cela revient à déterminer les maxima de la distribution de X.
- Exemples de techniques :
  - Algorithme *k-means*, groupement de X par itérations dynamiques ;
  - *Mixture modeling*, modélisation de X par juxtaposition de gaussiennes.

Classification Exemple de l'algorithme *k-means*

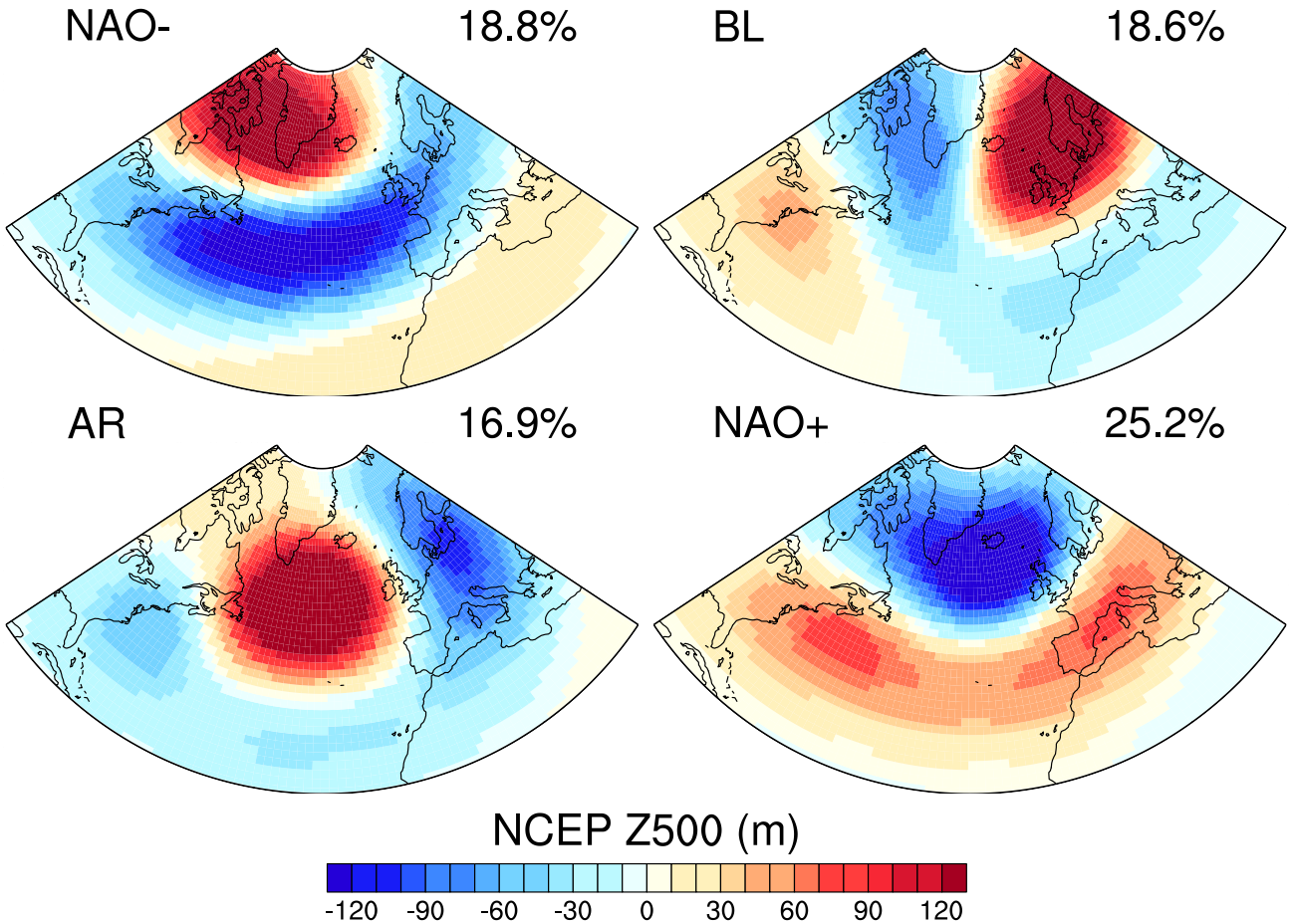
1. Choix du nombre k classes a priori.
2. Initialisation (aléatoire ou non) des k centres de classe (*centroïdes*).
3. Itérations :
   - Chaque observation est rangée avec le centroïde *le plus proche* ;
   - Les centroïdes sont recalculés (e.g., moyenne de la classe) ;
   - Itérations jusqu'à convergence du centroïde.



Code source et infos ici.

Classification Régimes de temps nord-atlantiques 1/3

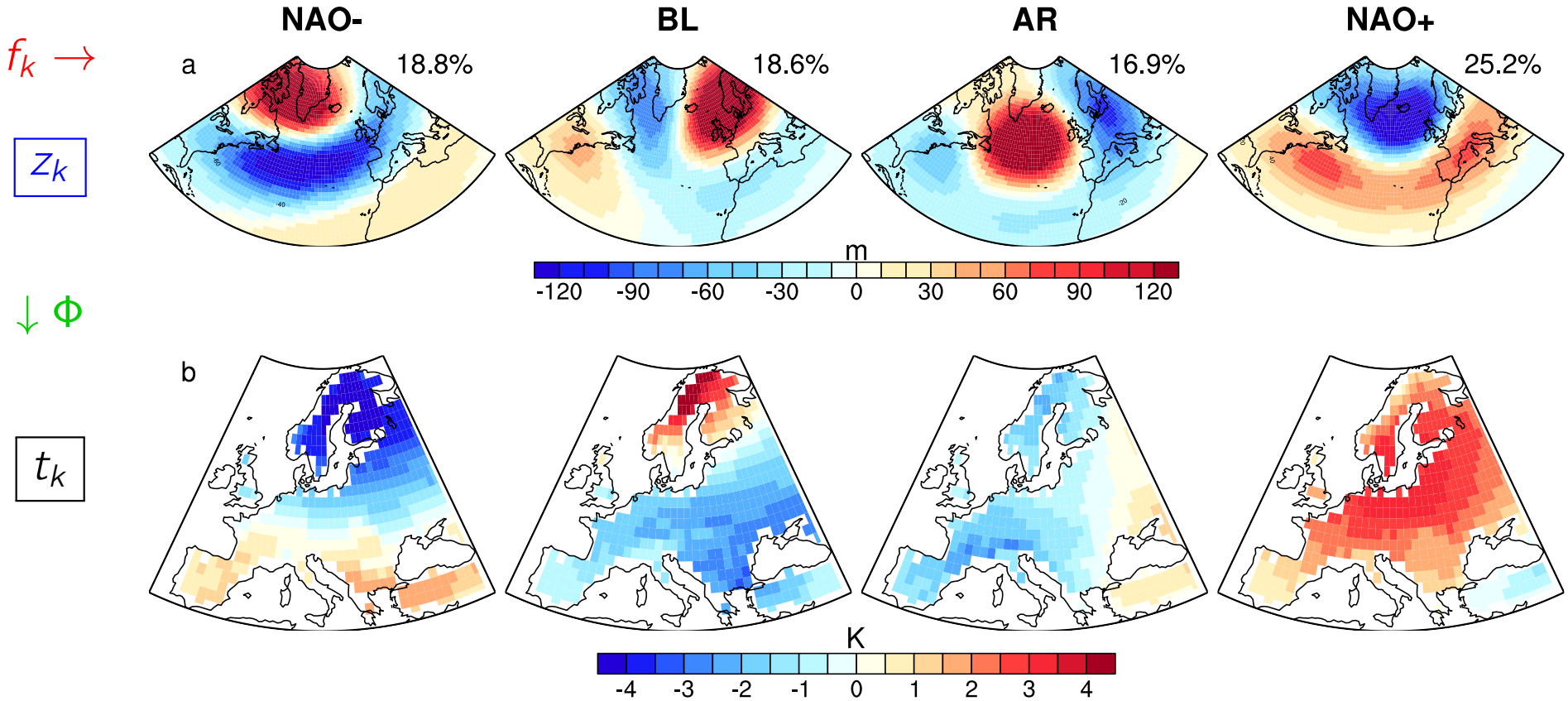
k-means des hauteurs quotidiennes de géopotentiel à 500mb (Z500).



Données : Z500 NCEP2 (DJFM 1979–2008) | Source : Cattiaux et al. (2013).

Classification Régimes de temps nord-atlantiques 2/3

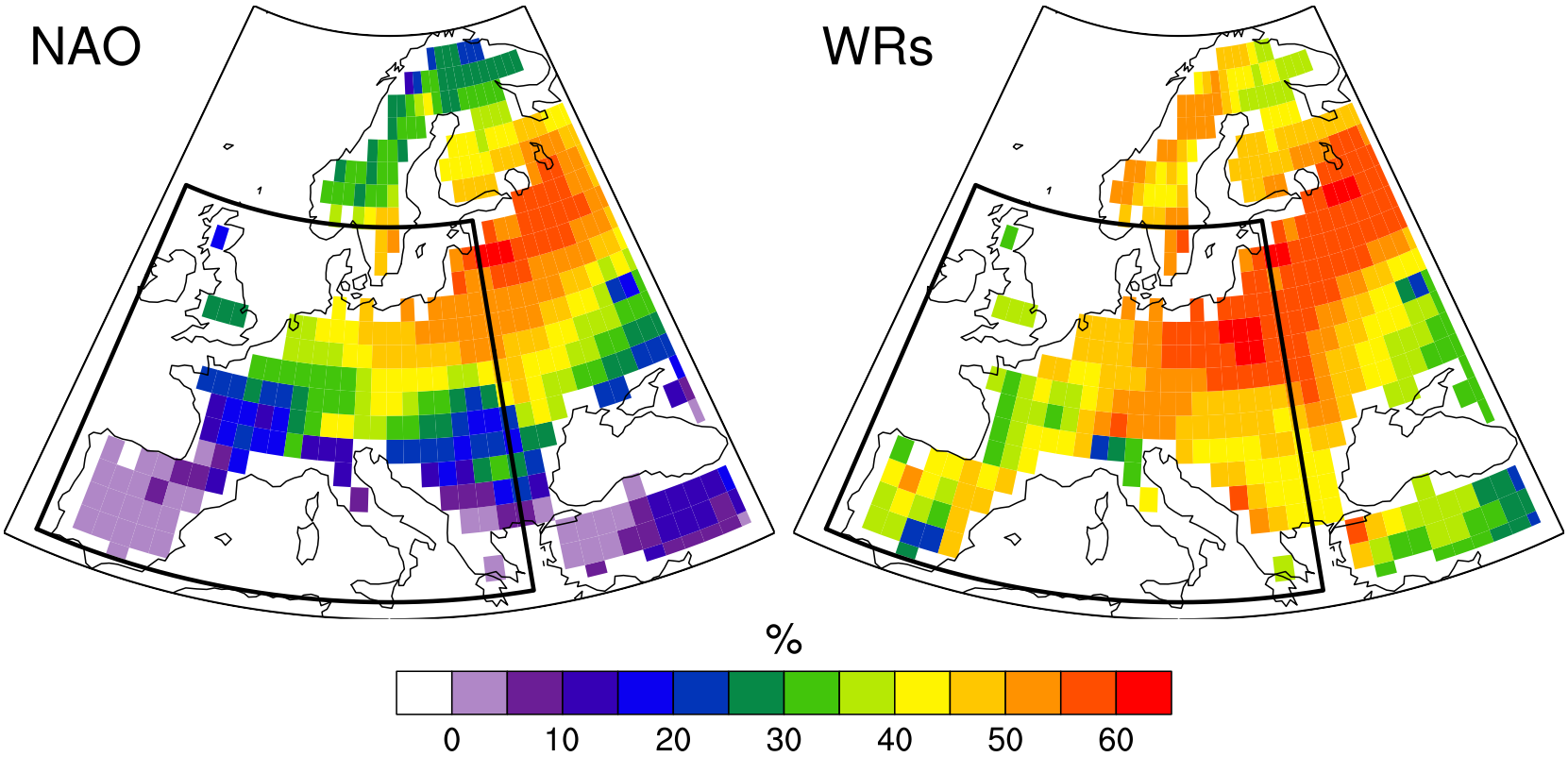
Description *discrète* des T européennes : $\overline{T} = \sum_k f_k \cdot t_k = \sum_k f_k \cdot \Phi(z_k)$.



Données : Z500 NCEP2 & T EOBS (DJFM 1979–2008) | Source : Cattiaux et al. (2013).

Classification Régimes de temps nord-atlantiques 3/3

- Les régimes expliquent davantage de variance des T hivernales que la NAO.



Données Z500 NCEP & T EOBS – Estimation sur DJFM 1979–2008.

Classification Remarques

- Régimes de temps :
  - choix arbitraire du nombre de classes ;
  - définis séparément pour hiver et été ;
  - *k-means* sensible à l'initialisation ;
  - nécessite de réduire au préalable la dimension par ACP ;
  - éventuellement, critères pour attribuer à une classe (persistance, distance au centroïde, etc.) et création d'une classe "poubelle".

- Plus généralement, ces techniques sont basées sur la variance, donc permettent de décrire la variabilité *courante longrightarrow* autres méthodes pour étudier les événements les plus rares (voir fin du cours).

Plus d'infos : cours de Pascal Yiou (LSCE), site de Christophe Cassou (CERFACS).

Plan

Formalisme

$X(t)$ et $Y(t)$ variables aléatoires (température, précipitation, etc.).

$X(t)$ et $Y(t)$ présentent-elles des tendances ?

$X(t)$ et $Y(t)$ sont-elles corrélées ?

Formalisme

$X(t)$ et $Y(t)$ variables aléatoires (température, précipitation, etc.).

$X(t)$ et $Y(t)$ présentent-elles des tendances ?

$X(t)$ et $Y(t)$ sont-elles corrélées ?

1. Modélisation (statistique) des données (e.g., i.i.d., loi normale).
2. Formulation de l'hypothèse nulle H_0 (e.g., corrélation $= 0$).
3. Formulation de l'hypothèse alternative H_1 (e.g., corrélation $\neq 0$).
4. Choix d'une statistique r pertinente (e.g. coefficient de corrélation).
5. Estimation de la distribution sous H_0 de cette statistique r.
6. Définition d'une région dans laquelle on pense pouvoir accepter H_0 avec un certain niveau de confiance α.
7. Calcul de r à partir des données et *décision*.

Formalisme C'est toujours pareil !

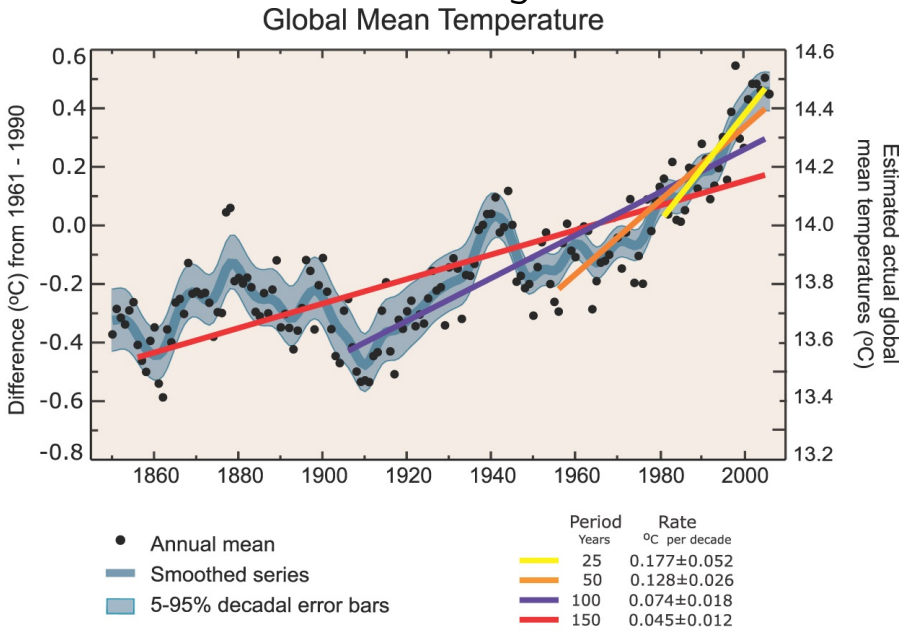
Types d'erreur

- Deux types d'erreur peuvent arriver lors d'un test.

	Rejet de H_0	Non-rejet de H_0
H_0 est vraie	Erreur de 1e espèce $proba = \alpha$ Niveau de significativité	Pas d'erreur proba $= 1 - \alpha$ Niveau de confiance
H_1 est vraie	Pas d'erreur $proba = 1 - \beta$ Puissance	Erreur de 2e espèce $proba = \beta$

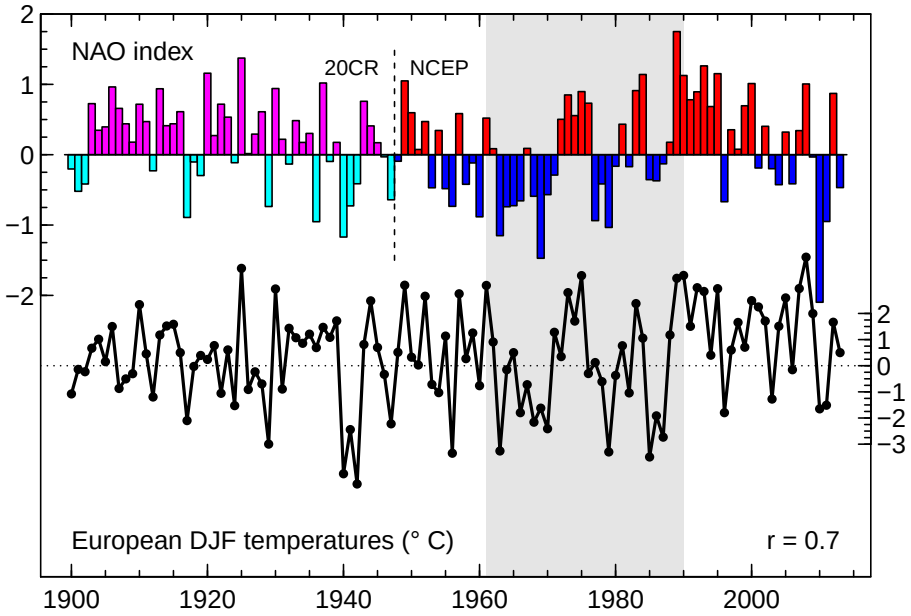
Illustrations

Tendances T globale



Source : IPCC AR4 (2007), FAQ 3.1, Figure 1.

Corrélation NAO et T Europe du Nord



Données : Z500 20CR & NCEP / T HadCRUT4.

Remarques :

- Distribution de r sous H0 liée à une loi de Student (nombre de degrés de liberté à définir, cf. TP).
- Corrélation n'implique pas *causalité*.

Remarques

- Remarques générales :
  - hypothèses cruciales sur les données (cf. TP) ;
  - pas toujours possible d'estimer analytiquement la distribution sous H_0, recours à des simulations numériques (e.g., *bootstrap*) ;
  - niveau de confiance déterminé arbitrairement ;
  - formalisme contre-intuitif, potentiellement source de confusion.

- Limites de l'application au climat :
  - en général, une seule série de données (pas possible de *resampler*) ;
  - fortes dépendances en temps et en espace.

Plus d'infos : cours de Pascal Yiou (LSCE) et de Slava Kharin (Environnement Canada).

Plan

Problématique

Une tendance n'est pas un changement !

La variabilité **interne** peut générer des tendances *par hasard*.
Un changement désigne une modification des forçages **externes**.

Comment détecter un changement climatique ?

Comment attribuer ce changement à des causes *i* ?

Problématique

Une tendance n'est pas un changement !

La variabilité **interne** peut générer des tendances *par hasard*.
Un changement désigne une modification des forçages **externes**.

Comment détecter un changement climatique ?

Comment attribuer ce changement à des causes i ?

1. Montrer que le signal n'est pas cohérent avec le seule variabilité interne.
2. Montrer que le signal est cohérent avec la réponse attendue à un ensemble de causes qui contient i (condition suffisante).
3. Montrer que le signal n'est pas cohérent avec la réponse attendue sans les causes i (condition nécessaire).

Formalisme 1/3

Hypothèse 1 : additivité.

$\longrightarrow$ On écrit que les observations Y_ℓ $(Y_{s,t})$

$=$ la moyenne m_ℓ

$+$ la variabilité interne ε_ℓ

$+$ la somme des *vraies* réponses $X_\ell^{*(i)}$ aux forçages externes i.

Remarque : en pratique on ne connaît pas X^* mais X, la réponse simulée par les modèles de climat.

Hypothèse 2 : les réponses vraies et simulées ne diffèrent qu'en amplitude.

Modèle statistique standard

$$Y_\ell = m_\ell + \sum_{i=1}^{N} \beta_i X_\ell^{(i)} + \varepsilon_\ell$$

Formalisme 2/3

Modèle statistique standard

$$Y_\ell = m_\ell + \sum_{i=1}^{N} \beta_i X_\ell^{(i)} + \varepsilon_\ell$$

$\longrightarrow$ On veut tester les facteurs d'amplitude β_i.

1. Détection (*on cherche à rejeter* H_0):
$$H_0 : \beta = 0_N \text{ vs. } H_1 : \beta \neq 0_N .$$

2. Attribution, cohérence avec toutes causes (*on cherche à accepter* H_0) :
$$H_0 : \beta_i = 1_N \text{ vs. } H_1 : \beta \neq 1_N .$$

3. Attribution, incohérence sans i (*on cherche à rejeter* H_0) :
$$H_0 : \beta_i = 0 \text{ vs. } H_1 : \beta_i \geq 0 .$$

Formalisme 3/3

Modèle statistique standard

$$Y_\ell = m_\ell + \sum_{i=1}^{N} \beta_i X_\ell^{(i)} + \varepsilon_\ell$$

Hypothèse 3 : la distribution de ε_ℓ est connue (modèles de climat).

$\longrightarrow$ On estime les β_i par *Optimal Fingerprinting* (Hasselmann, 1979).

- Cas $N = 1$:

$$\hat{\beta} = \frac{Y' C^{-1} X}{X' C^{-1} X} \ .$$

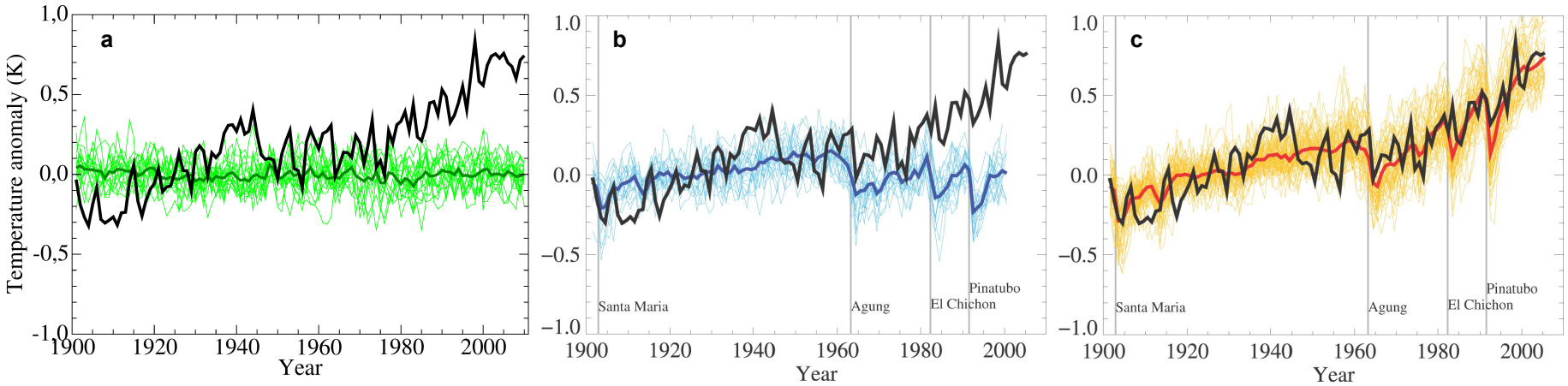
- Cas $N > 1$:

$$\hat{\beta} = (X' C^{-1} X)^{-1} X' C^{-1} Y \ .$$

Où $C = \mathrm{Cov}(\varepsilon_\ell)$.

Résultat majeur 1/2

T globale vs. variabilité interne + causes naturelles + causes anthropiques

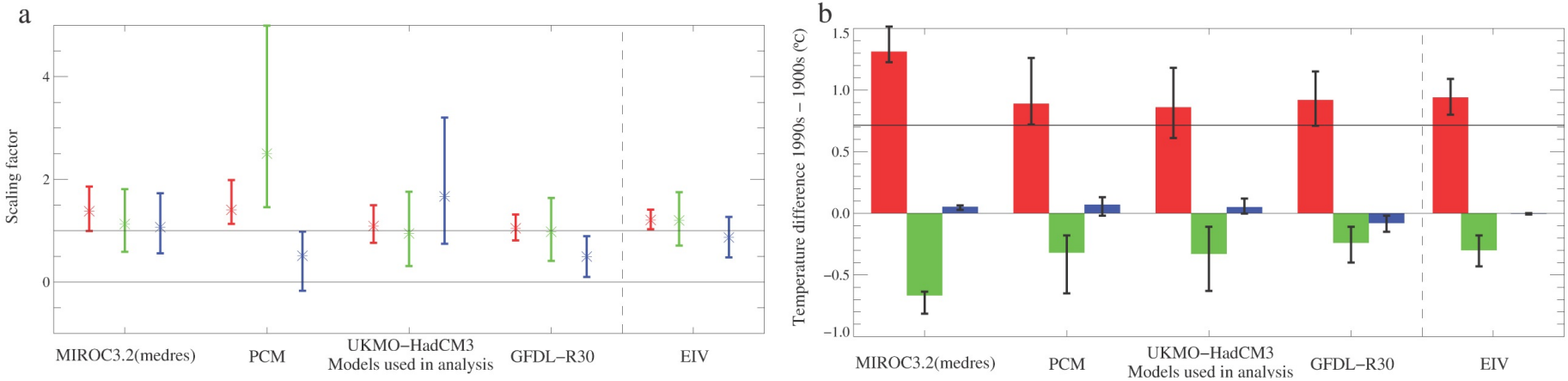


Adapté de l'IPCC AR4 (2007) Figure 9.5.

Résultat majeur 2/2

Reconstruction $T_{\mathrm{OBS}} = \beta_{\mathrm{GES}} T_{\mathrm{GES}} + \beta_{\mathrm{OA}} T_{\mathrm{OA}} + \beta_{\mathrm{NAT}} T_{\mathrm{NAT}} + \varepsilon$

Estimations des facteurs β et des contributions respectives au ΔT 1990s−1900s.



Source : IPCC AR4 (2007) Figure 9.9.

GES : Gaz à effet de serre – OA : Autres anthropiques (e.g. aérosols) – NAT : Naturels.

Remarques

- De plus en plus de résultats sur autres variables, autres régions etc.
- Limites :
  - hypothèses assez fortes ;
  - besoin de réduire en espace et en temps (ACP, projections sur harmoniques sphériques, moyennes décennales, etc.) ;
  - difficulté de séparer des réponses colinéaires (e.g., GES vs. aérosols anthropiques) ;
  - rapport signal sur bruit parfois fort ;
  - ne peut s'appliquer qu'à des variables bien représentées par les modèles de climat (peut être diagnostiqué a posteriori).

Plan

Prévision et échelles de temps

Principe de la prévision

Déterminer parmi l'ensemble des *possibles* le sous-ensemble des *plausibles*, en fonction des conditions initiales (observations).

Attention ! Différent de la *projection* climatique qui consiste à déterminer l'ensemble des possibles lui-même, et qui n'est donc pas *initialisée*.

- Différentes échelles de temps (et d'espace) :
  - prévision immédiate (une ville) ;
  - prévision *météo* classique (un pays) ;
  - prévision saisonnière (un continent) ;
  - projection centennale (le globe).

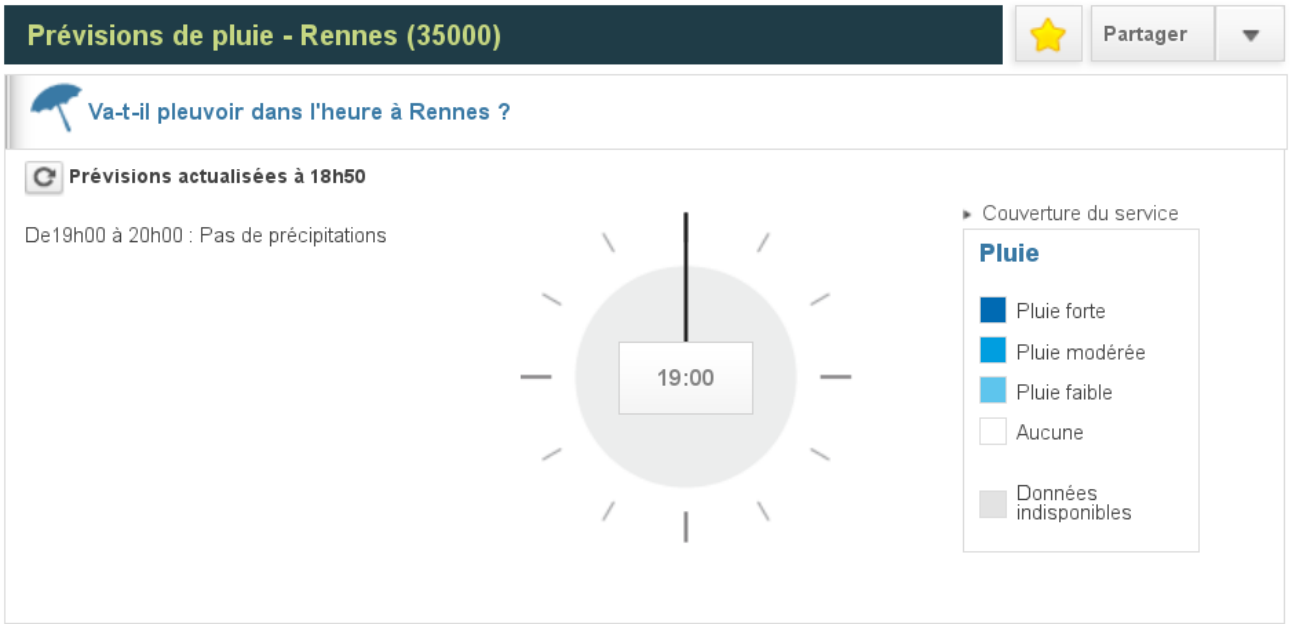
Remarque : la *prévisibilité* n'est pas toujours la même.

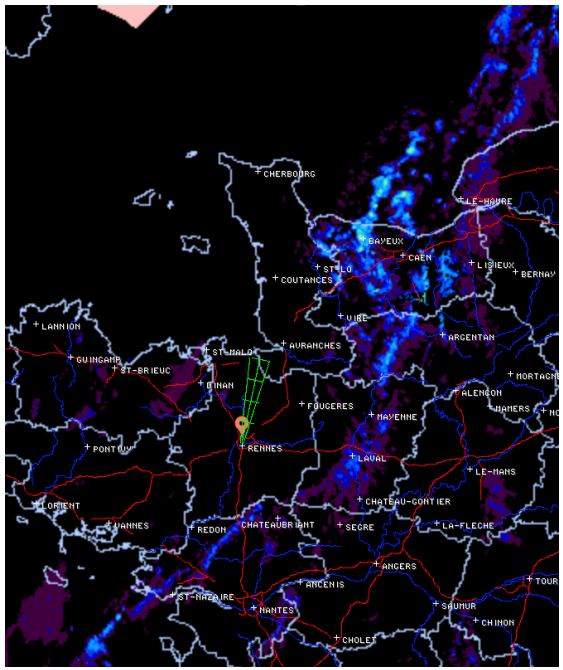
C'est parti pour quelques illustrations...

La prévision immédiate

On bâche le Court Central ? On passe en pneus slick ? Je rentre à vélo maintenant ou j'attends un peu ?

$\longrightarrow$ Prévision fine à très petite échelle, essentiellement basée sur les observations (traitement d'image, peu de stats).



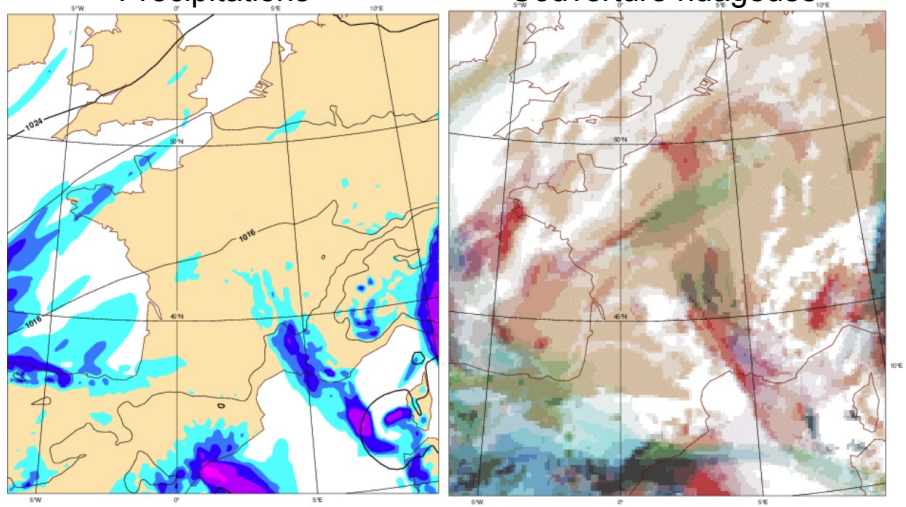


© Météo-France.

La prévision météo *classique*

On s'habille comment demain ? Quelle rando ce WE ?

$\longrightarrow$ Prévision du temps *sensible* et de sa chronologie, basée sur des simulations de modèles numériques, initialisées à partir d'observations (éventuellement adaptation statistique en aval).



© TF1 & Météo-France.

La prévision saisonnière

Doit-on s'attendre à un hiver plutôt chaud/froid, humide/sec ?

$\longrightarrow$ Prévision des grandes tendances des quelques mois à venir, basée sur des simulations de modèles numériques, initialisées à partir des anomalies observées de température de surface de la mer. Qq alternatives statistiques.

Prévisions pour l'hiver 2014/15

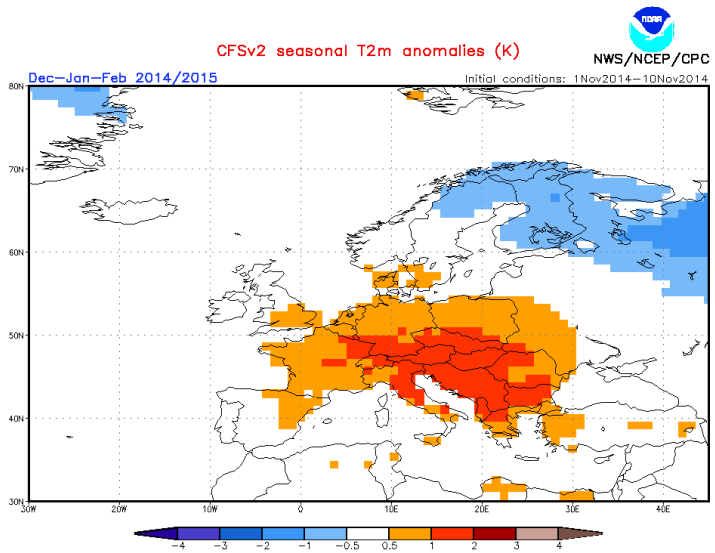
	Température	Précipitation
MODELES	France Métropole	France Métropole
MF		
CEP		
Met Office		
NCEP		
JMA		
Synthèse		
EuroSIP		
LC-MME		
Scénario privilégié par Météo-France	*chaud*	*pas de scénario privilégié*

T : Chaud
P : Sec

Normal

T : Froid
P : Humide

Pas de scénario privilégié

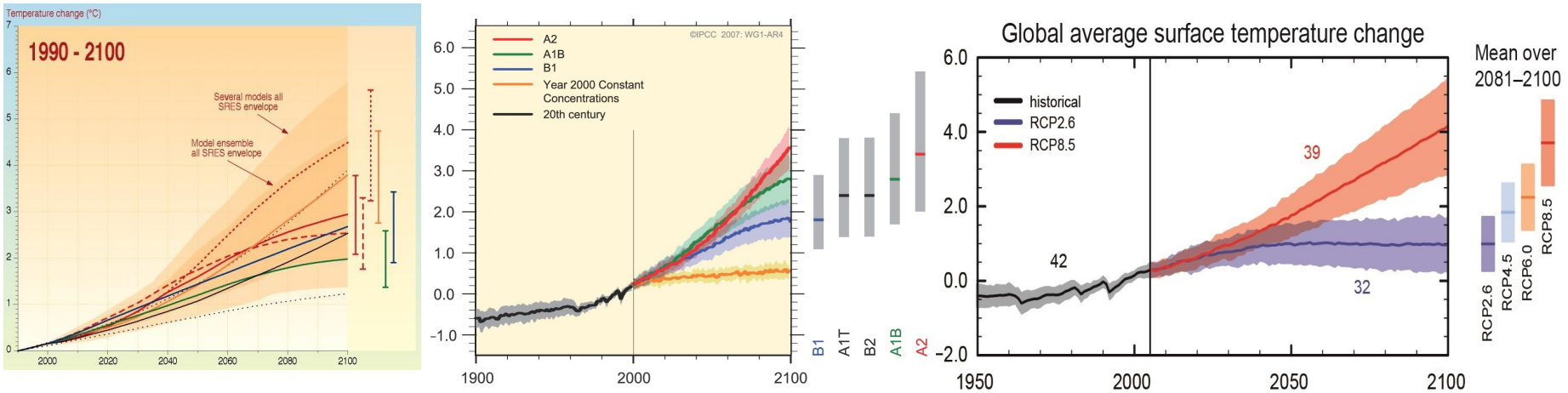


La projection centennale

Quel climat pour le XXIe siècle ?

$\longrightarrow$ Evaluation des distributions de probabilité des variables climatiques, basée sur des simulations de modèles numériques, "non-initialisées", auxquelles on impose seulement les conditions aux limites (forçages externes).

Projections pour la T globale d'ici 2100

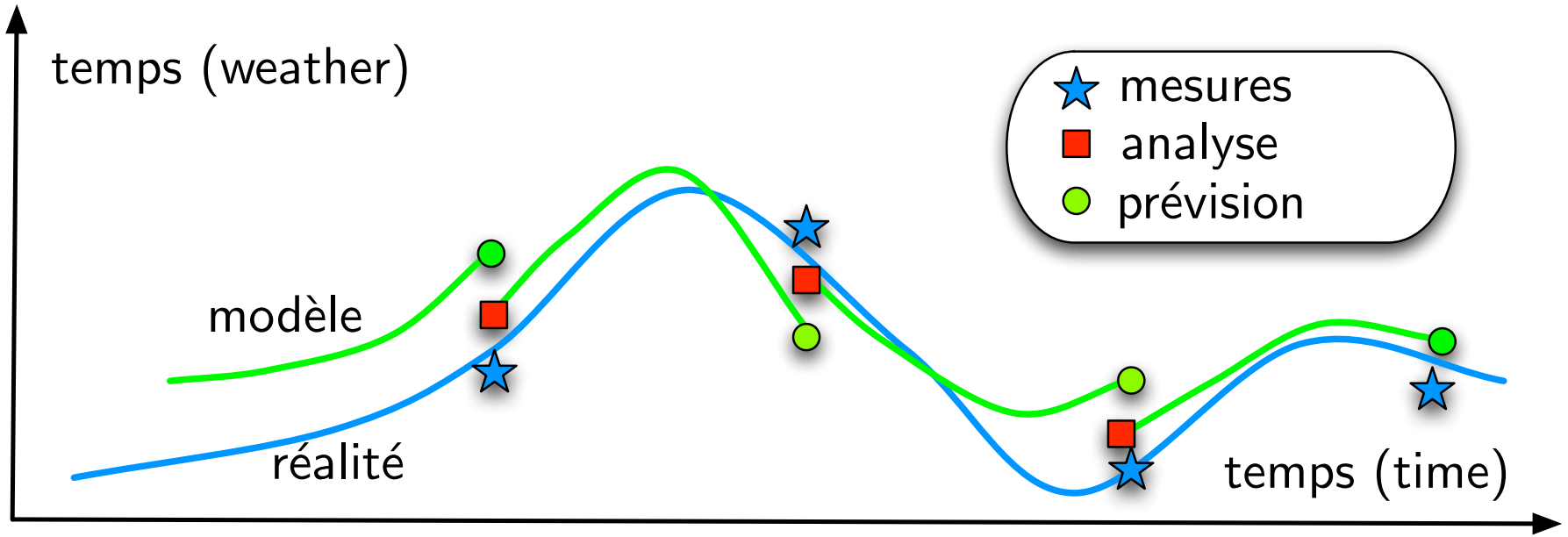


© IPCC TAR (2001), AR4 (2007) et AR5 (2013).

THE problème de la prévision

Comment initialiser ?

- Assimilation pour coller aux observations (en temps réel).
- Analyse pour ne pas “brusquer” le modèle numérique.



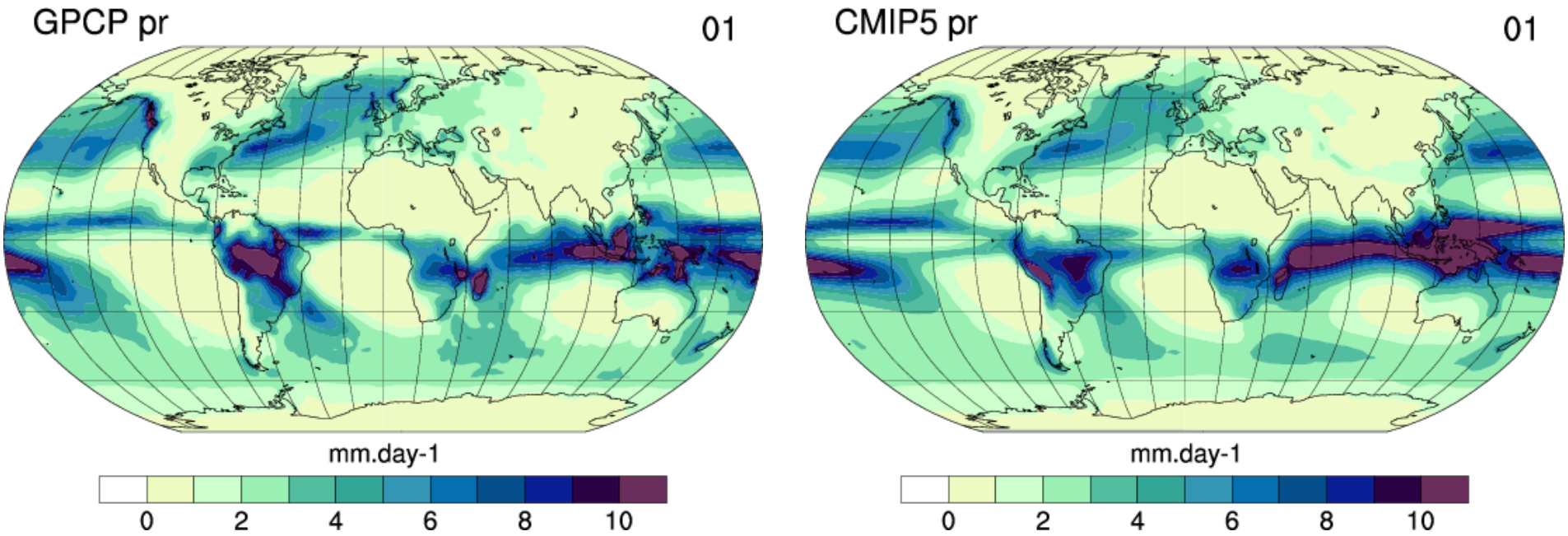
© O. Thual (Météo-France).

THE problème de la modélisation

Comment avoir confiance dans le modèle ?

$\longrightarrow$ Comparaison permanente aux observations (évaluation) : beaucoup de stats (méthodes d'interpolation, calcul de scores, etc.)

Exemple de la précipitation moyenne dans les modèles de climat



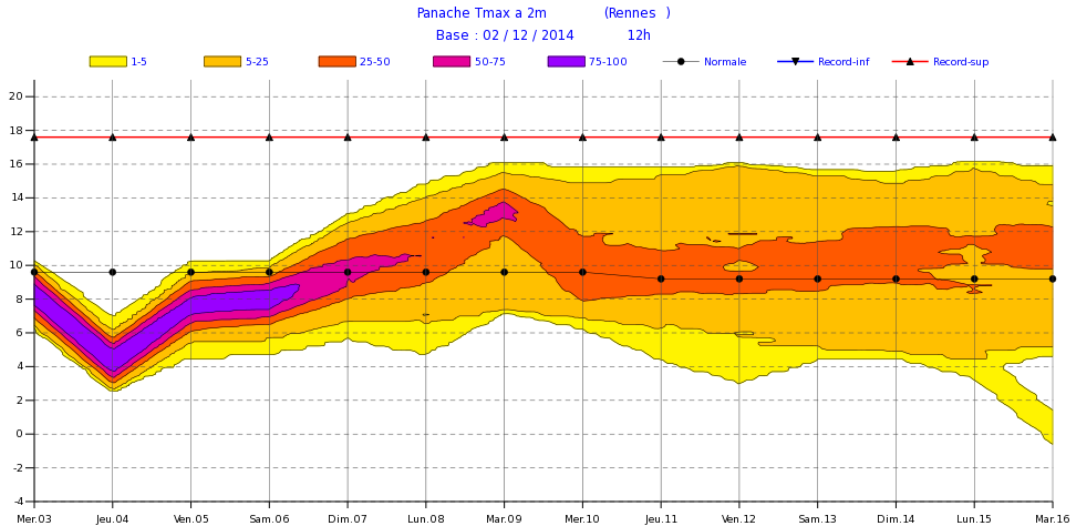
Données GPCP 1981–2010 et CMIP5 (ensemble de 38 GCMs).

Sources d'incertitudes et utilisation d'ensembles

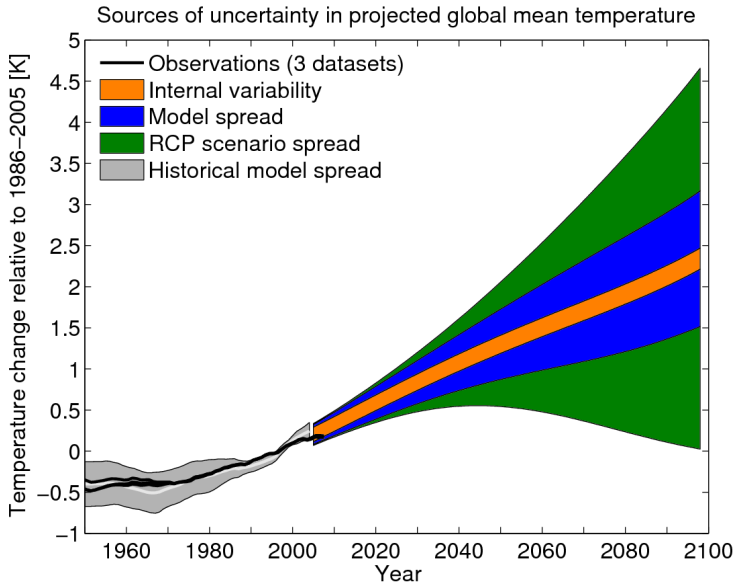
En prévision / projection :

- Conditions initiales : plusieurs initialisations *raisonnables* à une date donnée.
- Modélisation : plusieurs modèles et/ou versions d'un même modèle.
- Variabilité interne : initialisations à différentes dates d'une simulation de climat stationnaire (spectre bien plus large que pour prévision).
- Conditions aux limites : plusieurs scénarios de forçages externes.

Prévision : exemple incertitude C.I.



Projection : les 3 incertitudes



Remarques

- Prévision vs. projection :
  - météo vs. climat ;
  - réalisation vs. distribution ;
  - initialisation vs. grands équilibres physiques.

- Sources de prévisibilité (et donc méthodes) dépendent de l'échéance spatio-temporelle.

- Les prévisions qui montent :
  - la prévision mensuelle (revisiter la limite d'Ed Lorenz) ;
  - la prévision décennale (étudier les sources de prévisibilité, e.g., l'océan profond).

Plan

Motivations

On a observé X pendant N années.

Quelle est la probabilité de battre le record de X à l'année $N+1$?

Quel est le niveau de retour à $2 * N$ de X ?

Motivations

On a observé X pendant N années.

Quelle est la probabilité de battre le record de X à l'année $N+1$?

Quel est le niveau de retour à $2 * N$ de X ?

- Pour étudier les événements observés :
  - distribution empirique (quantiles, nombre de déviations standards) ;
  - théorie des valeurs records.
- Pour étudier les événements *pas encore* observés :
  - théorie des valeurs extrêmes (EVT).

Records Théorie

Pour X_n ($n \in 1..N$) une variable stationnaire :

- Probabilité que X_n soit record, i.e. $\forall i < n\ ,\ X_i < X_n$:

$$p_n = \frac{1}{n}\ .$$

- Espérance du nombre de records de X observés pendant 1 et n :

$$e_n = \sum_{i=1}^{n} p_i = \sum_{i=1}^{n} \frac{1}{i} \longrightarrow ln(n) + \gamma\ .$$

- Espérance de la valeur r_k du k^e record :

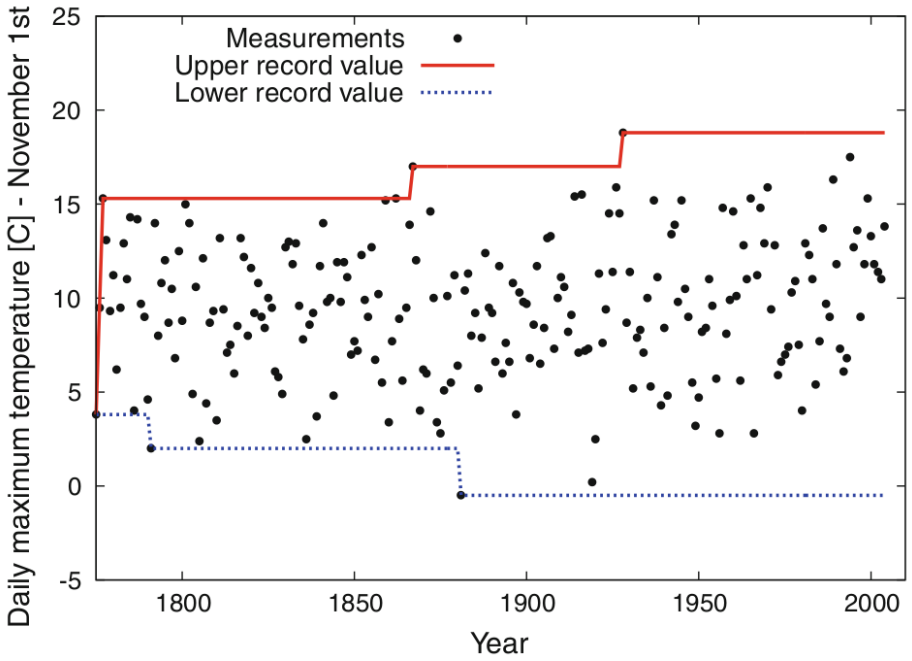
$$r_1 = \int_{-\infty}^{+\infty} X\ p(X)\ dX \quad \text{et} \quad r_k = \frac{\int_{r_{k-1}}^{+\infty} X\ p(X)\ dX}{\int_{r_{k-1}}^{+\infty} p(X)\ dX}\ .$$

$$\text{Cas } X \sim \mathcal{N}(\mu = 0, \sigma)\ :\ r_1 = \mu = 0\ ,\quad r_2 = \sqrt{\frac{2}{\pi}}\ \sigma \quad \text{et} \quad r_k = \frac{r_2\ e^{-r_{k-1}^2/2\sigma^2}}{\text{erfc}(r_{k-1}/\sqrt{2\sigma^2})} \longrightarrow \sqrt{2k\sigma^2}\ .$$

Propriété : p et e indépendants de la distribution de X, mais pas r.

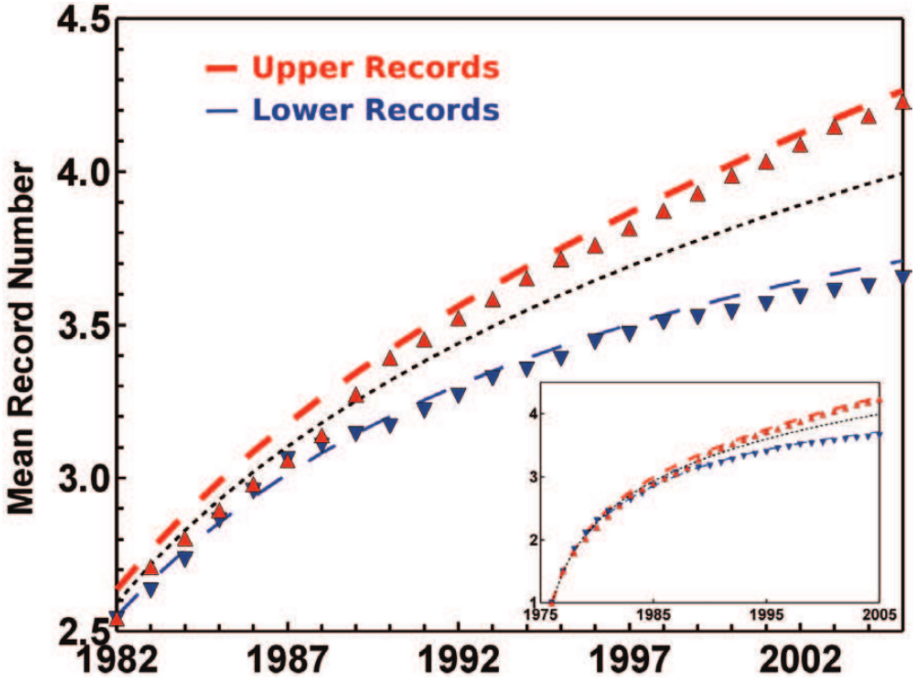
Records Illustration

Records de T le 1/11 à Prague



Source : Wergen et al. (2012)

Moyenne européenne de e_n



Source : Wergen and Krug (2010)

- On retrouve le comportement logarithmique sur e_n.
- On bat *significativement* plus de records chauds que froids.

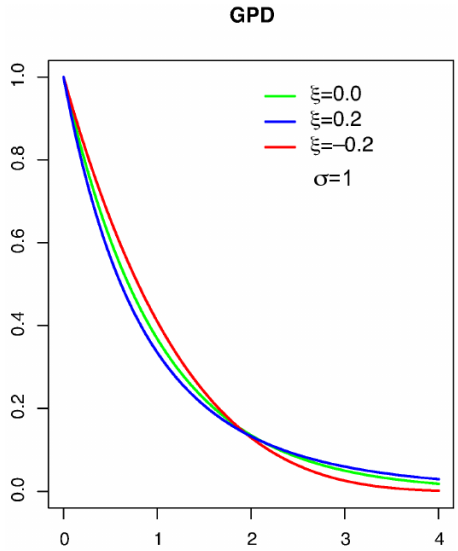
EVT Grandes lignes

- Modéliser la valeur maximale par *blocs* :

Generalized Extreme Value Distribution

$$\mathrm{GEV}(x;\mu,\sigma,\xi) = \exp\left\{-\left[1+\xi\left(\frac{z-\mu}{\sigma}\right)\right]^{-1/\xi}\right\} .$$

Regroupe 3 familles : Gumbel ($\xi = 0$), Fréchet ($\xi > 0$) et Weibull ($\xi < 0$).

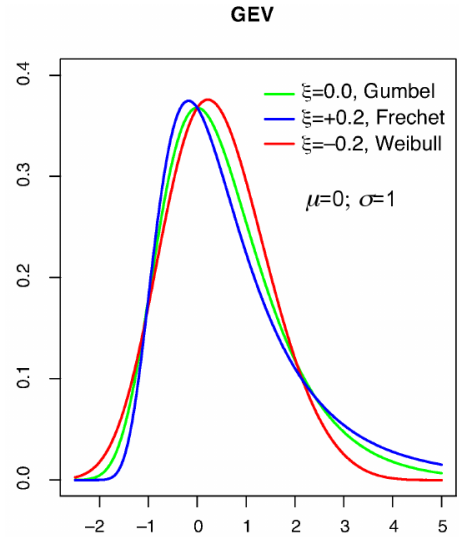


- Modéliser les dépassements d'un seuil :

Generalized Pareto Distribution

$$\mathrm{GPD}(y;\sigma,\xi) = 1-\left(1+\xi\frac{y}{\sigma}\right)^{-1/\xi} .$$

Contient la loi exponentielle ($\xi = 0$).



EVT Illustration 1/2

Ajustements
Gumbel - GEV

DUREES DE RETOUR DE FORTES PRECIPITA
Episode : 1 jour – Méthode de Gumbel

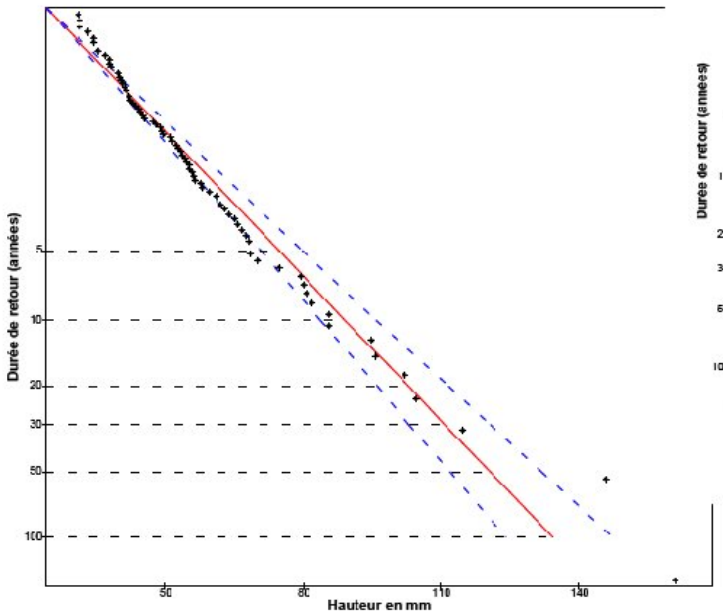
Statistiques sur la période 1922-2006
sous-période : du 01 Janvier au 31 Décembre

MARIGNANE (13)

Indicatif . 13054001, alt . 5 m., lat . 43°26'30'

GRAPHIQUE D'AJUSTEMENT

La droite donne la hauteur de précipitations estimée pour une durée de retour exprimée en années.
Les observations sont pointées. L'intervalle de confiance à 70 % est représenté en pointillés.



Page 2/2

DUREES DE RETOUR DE FORTES PRECIPITATIONS
Episode : 1 jour – Loi GEV

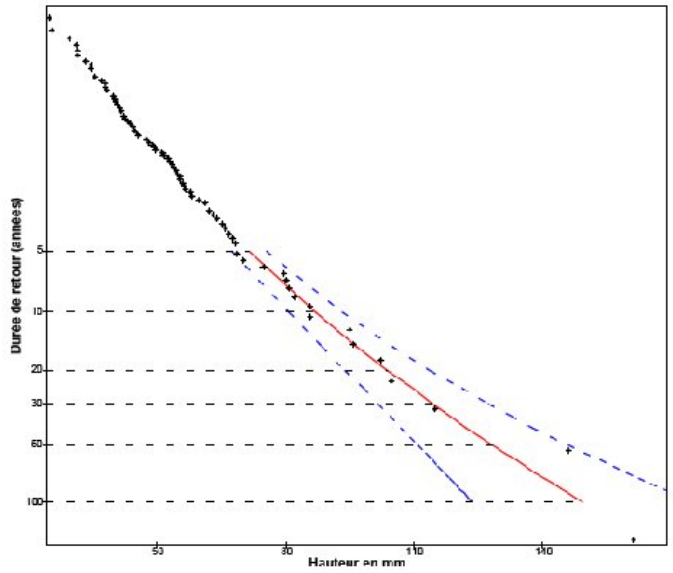
Statistiques sur la période 1922-2006

MARIGNANE (13)

Indicatif . 13054001, alt . 5 m., lat . 43°26'30"N, lon . 05°13'39"E

GRAPHIQUE D'AJUSTEMENT

La droite donne la hauteur de précipitations estimée pour une durée de retour exprimée en années.
Les observations sont pointées. L'intervalle de confiance à 70 % est représenté en pointillés.

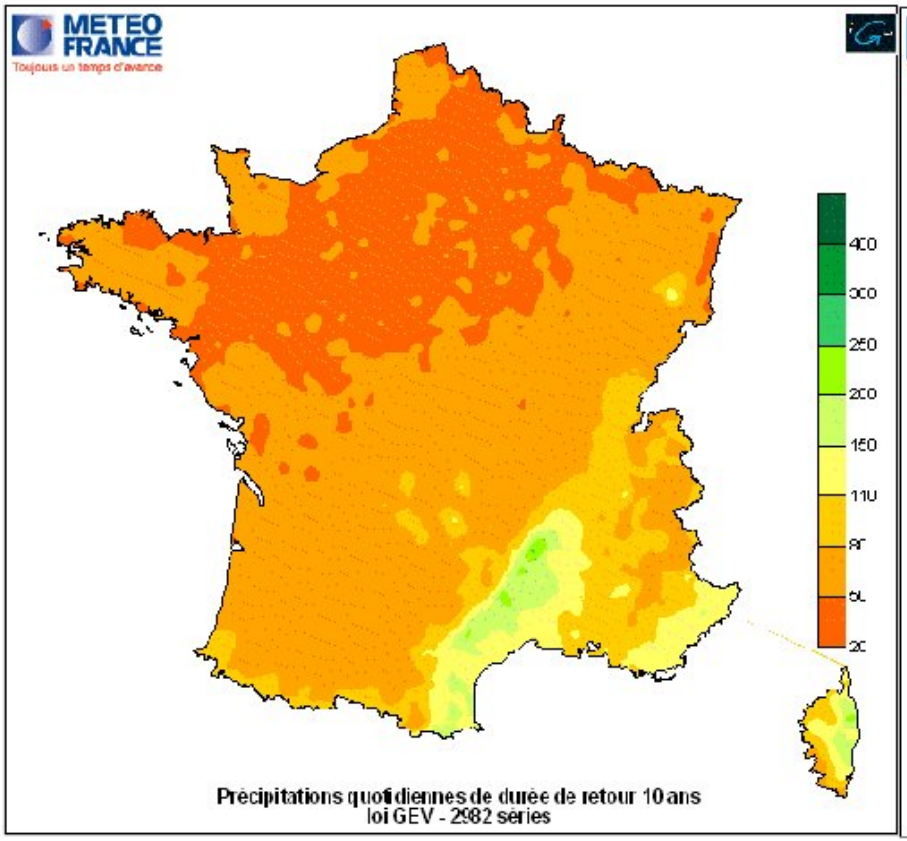


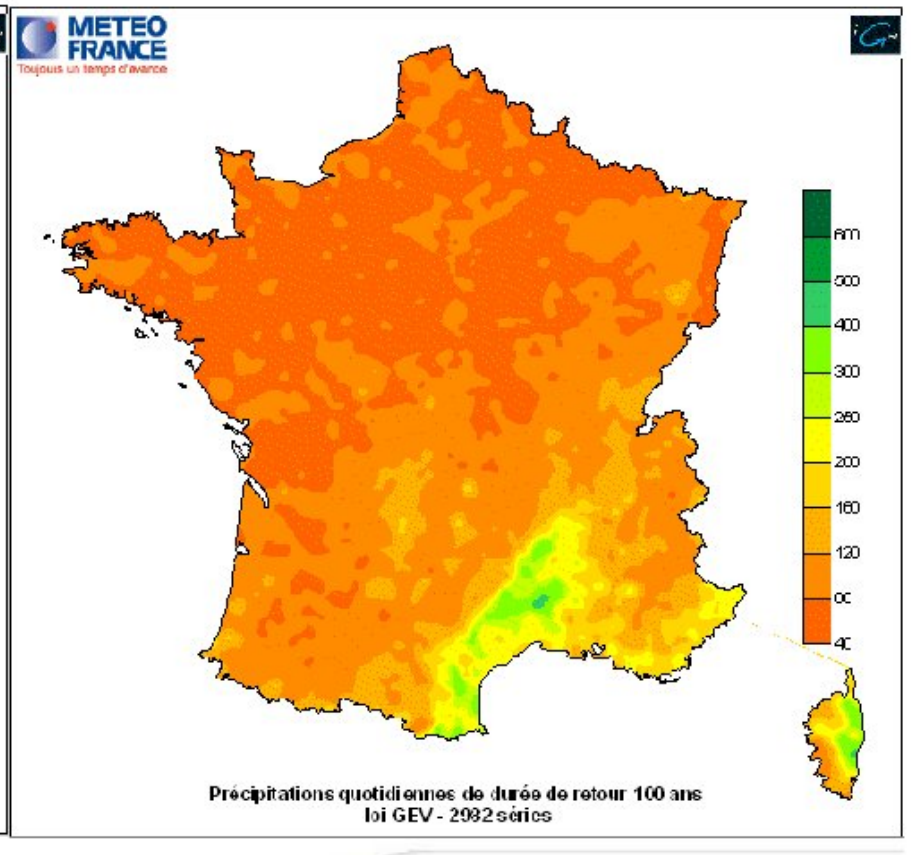
Page 2/2

METEO FRANCE
Toujours un temps d'avance

EVT Illustration 2/2

Niveaux de retour 10 ans et 100 ans des P quotidiennes (GEV + interpolation)





Remarques

- Records : problème d'actualité (France et Monde) !

- Les évolutions des records/extrêmes sont-elles uniquement liées aux changements de moyenne ? Changements de forme des distributions ?

- Attribution d'événements : notion de fraction de risque attribuable, par exemple évaluée à partir de GEV non-stationnaires.