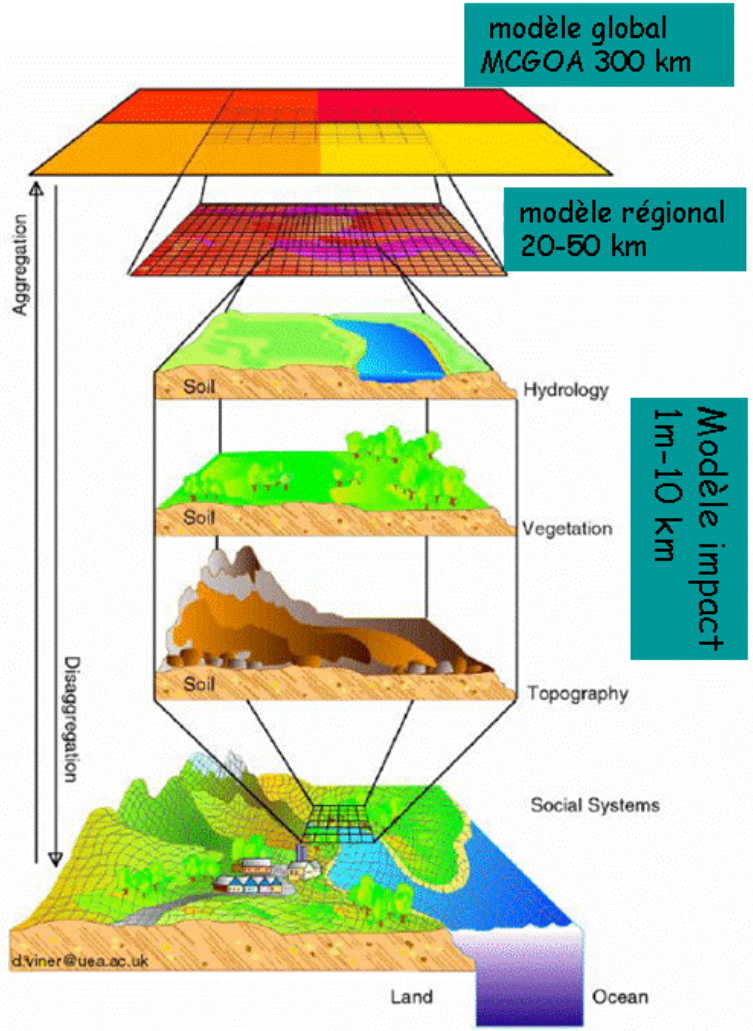
Désagrégation spatiale: les différentes approches avec des applications dans l'étude des impacts du changement climatique

Julien Boé

CNRS/CERFACS

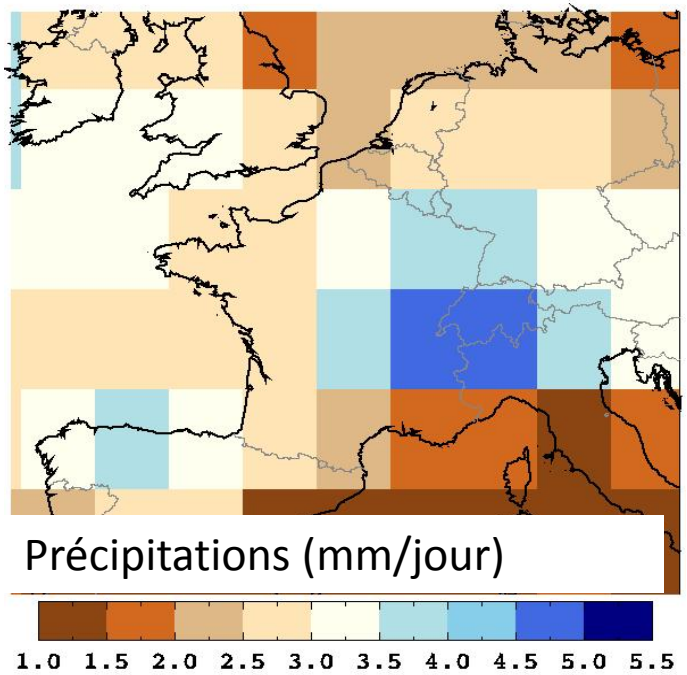
Formation Modélisation:
ECCOREV – 25-26 septembre 2014, CEREGE, Aix-en-Provence



Plan de la présentation

- Problématique, les différentes approches
- Désagrégation dynamique et correction de biais
- Désagrégation statistique
- Calibration: difficultés dans un cadre non stationnaire
- Conclusions

Problématique



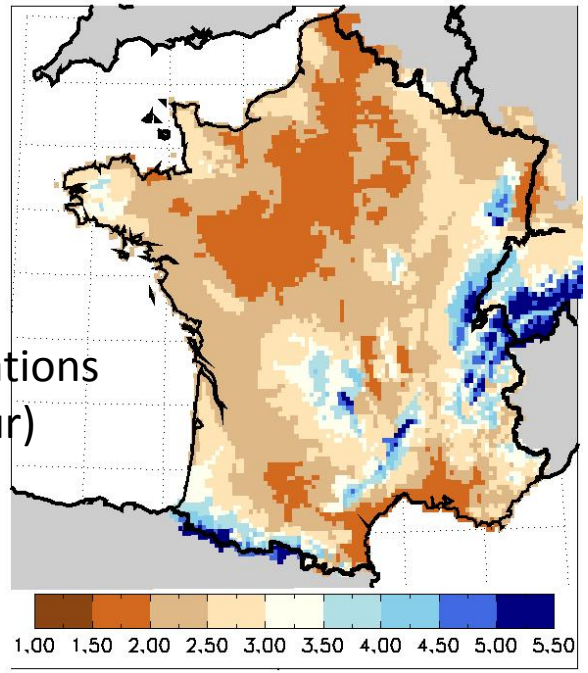
Information à basse résolution e.g. **Champs météo. d'un modèle climatique**
~ 100-250km

Variables météorologiques de forçage <10km, stations etc.

Downscaling

Désagrégation

Précipitations (mm/jour)



Modèles requérant des données à haute résolution e.g. **hydrologie**

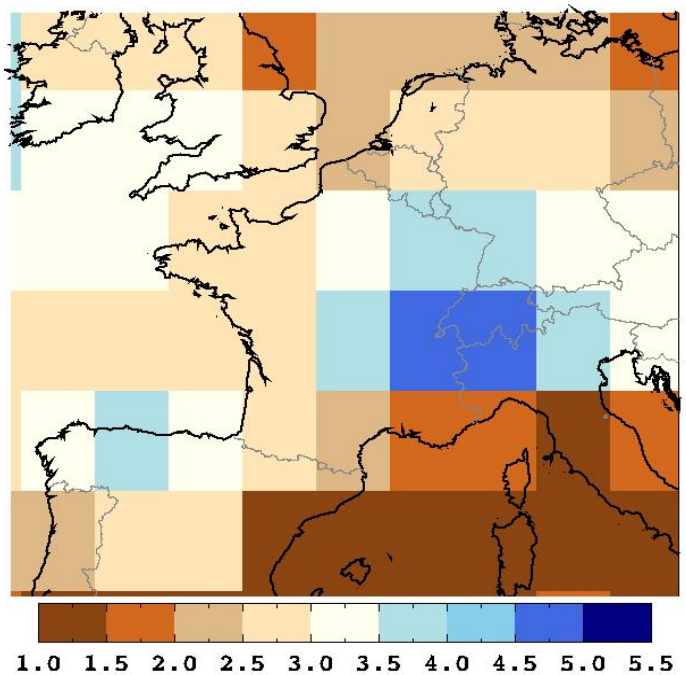
MODELE

Différentes approches

- Interpolation basique
- Interpolation « améliorée » avec prise en compte de caractéristiques physiographiques
- Désagrégation statistique
- Désagrégation dynamique

Interpolation basique

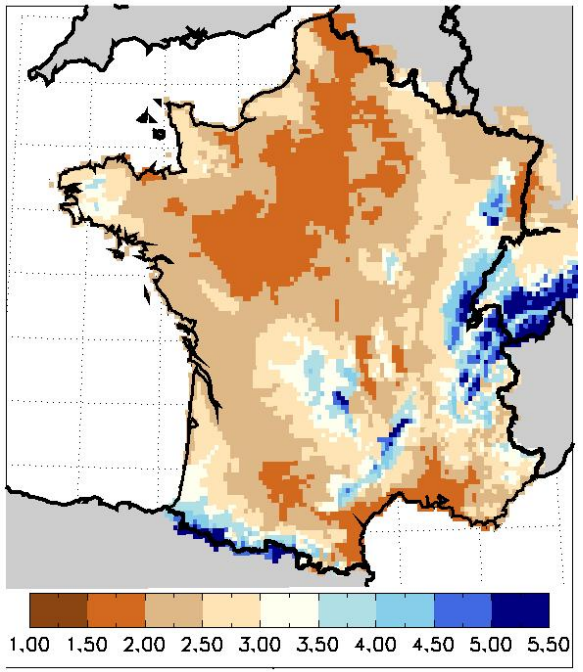
Modèle climatique 250km



Interpolation basique

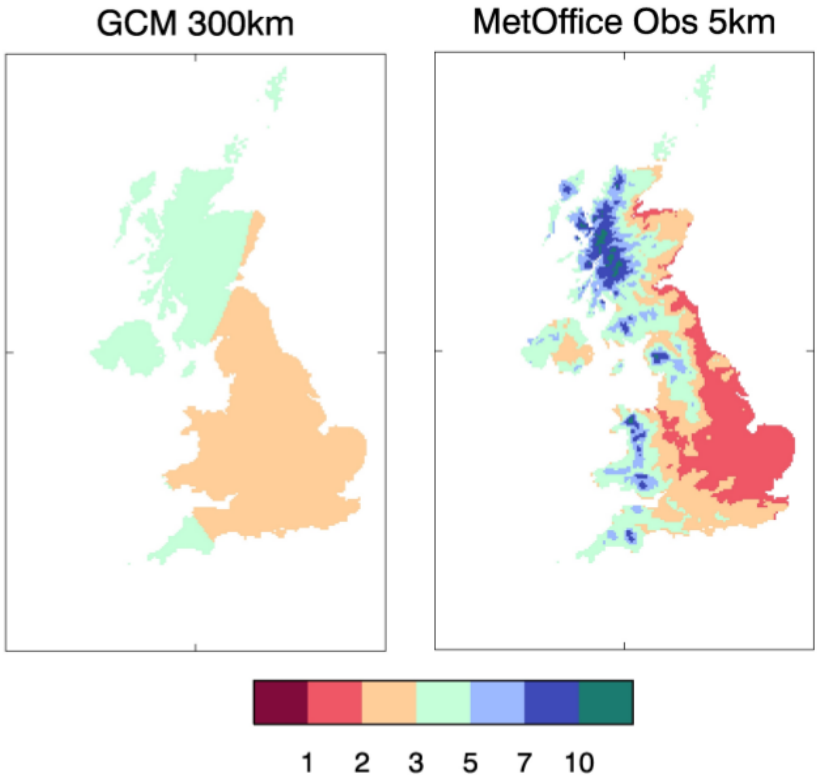
Non !!

Obs 8 km



Précipitations climatologiques (mm/jour)

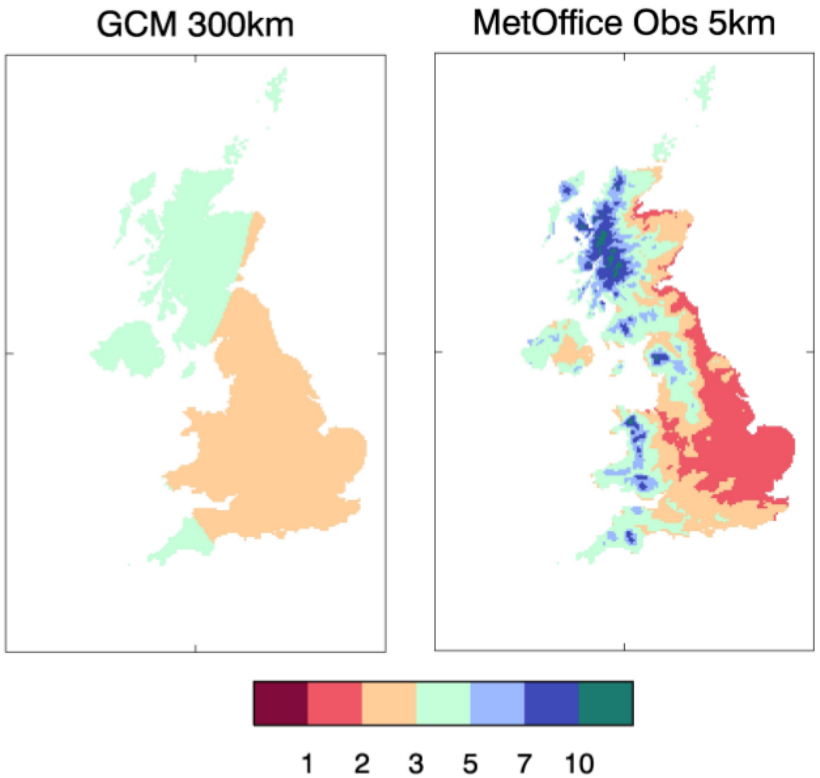
Interpolation basique



Précipitations moyennes (mm/jour), 1961-2000
Maraun et al. 2010

On peut toujours interpoler les données à basse résolution sur une grille à haute résolution
⇒On ne va pas créer de l'information à petite échelle
⇒**Problème majeur si de la variabilité à petite échelle existe…**

Interpolation « améliorée »



Précipitations moyennes (mm/jour), 1961-2000
Maraun et al. 2010

Orographie UK
wikipedia

Interpolation « améliorée »

- Prise en compte de facteurs physiographiques (e.g. topographie) influençant de manière significative la distribution spatiale de la variable d'intérêt

E.g. Modèle de régression

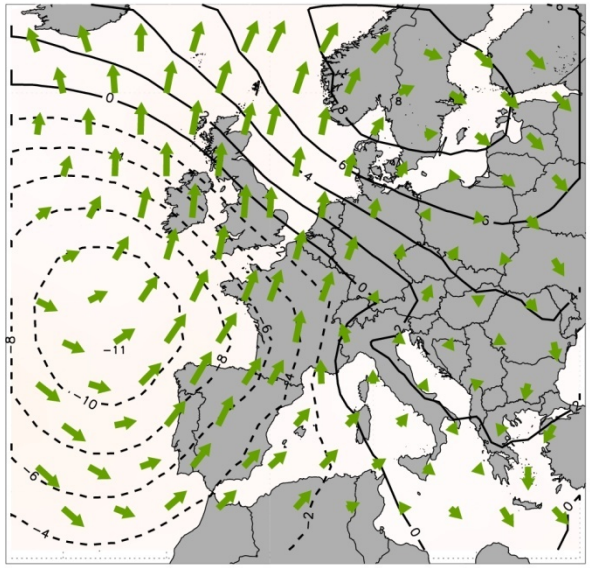
Vi(x,y)=a0+a1.Altitude(x,y)+a2.Latitude(x,y)+…

Disponible à haute résolution

Vi(x,y): carte climatologique ? carte mensuelle? carte journalière? => le rôle des facteurs physiographiques varient... Selon Vi aussi.

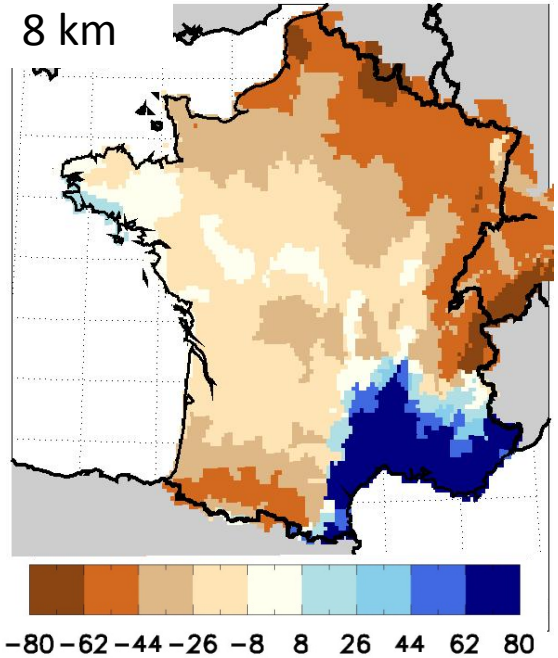
=> Approche « statique »

Illustration



Type de circulation atmosphérique, flux de sud
Flèches: Vent 850 hPa
Lignes: Anomalies pression

Anomalies relatives des précipitations associées (%)



Observations

=> Le lien entre précipitations et orographie est « dynamique »: e.g. dépend de la circulation atmosphérique.

=> Prise en compte de cet aspect dynamique: **désagrégation statistique ou dynamique**

La désagrégation

=>Aller de l'échelle du modèle climatique vers celle du modèle d'impact.

Deux grandes familles de méthodes:

Etablir une relation statistique entre variables locales et prédicteurs de grande échelle

Désagrégation statistique

Résoudre explicitement la physique et la dynamique du système climatique régional

Désagrégation dynamique

Utilisées de façon indépendante ou combinée

DESAGREGATION DYNAMIQUE

Modèle climatique (GCM: global climate model)

$$\frac{d\vec{V}}{dt} = \frac{-\nabla p}{\rho} - 2\vec{\Omega} \wedge \vec{V} + \vec{g} + \vec{F}$$

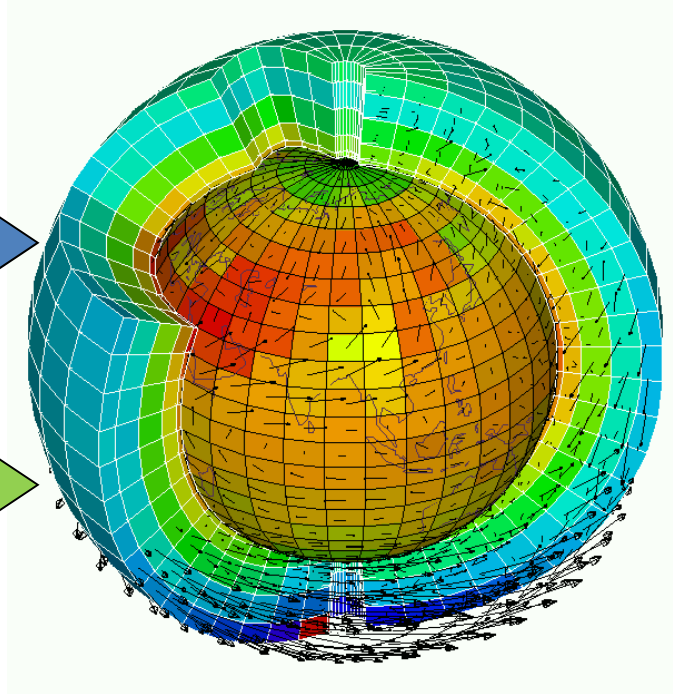
$$C_p \frac{dT}{dt} = \frac{RT}{p}\frac{dp}{dt} + Q$$

$$\frac{d\rho}{dt} = -\rho \nabla \vec{V}$$

$$\frac{dq}{dt} = Q'$$

Découpage en maille

Découpage en temps



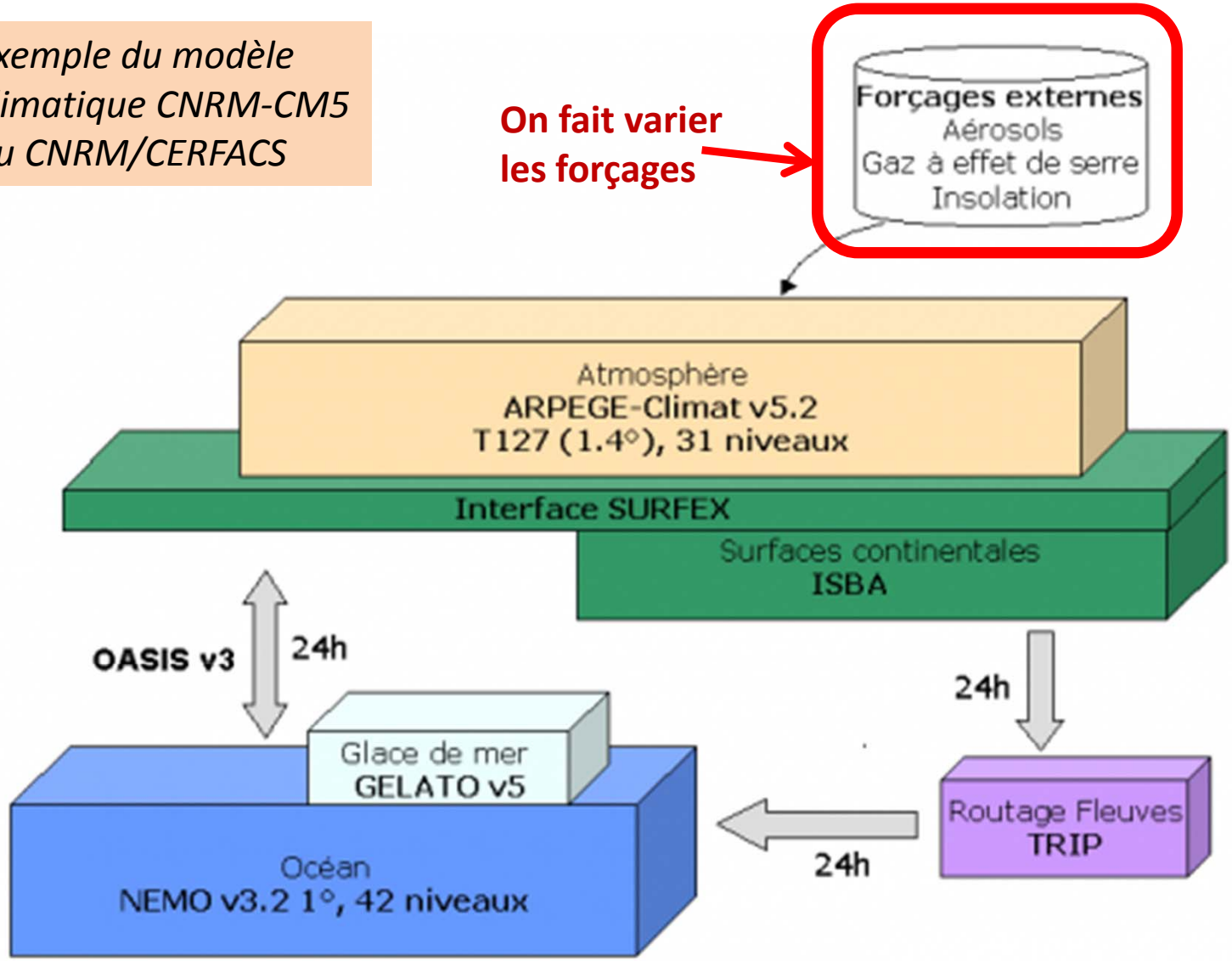
Typiquement: maille de ~ 200km x 200km!

Représentation des phénomènes d'échelle spatiale inférieure à la taille de la maille

=> Modèles climatiques globaux couplés océan/atmosphère: très couteux en temps de calcul. Limite leur résolution.

Modèle climatique (GCM: global climate model)

Exemple du modèle climatique CNRM-CM5 du CNRM/CERFACS



Différentes approches

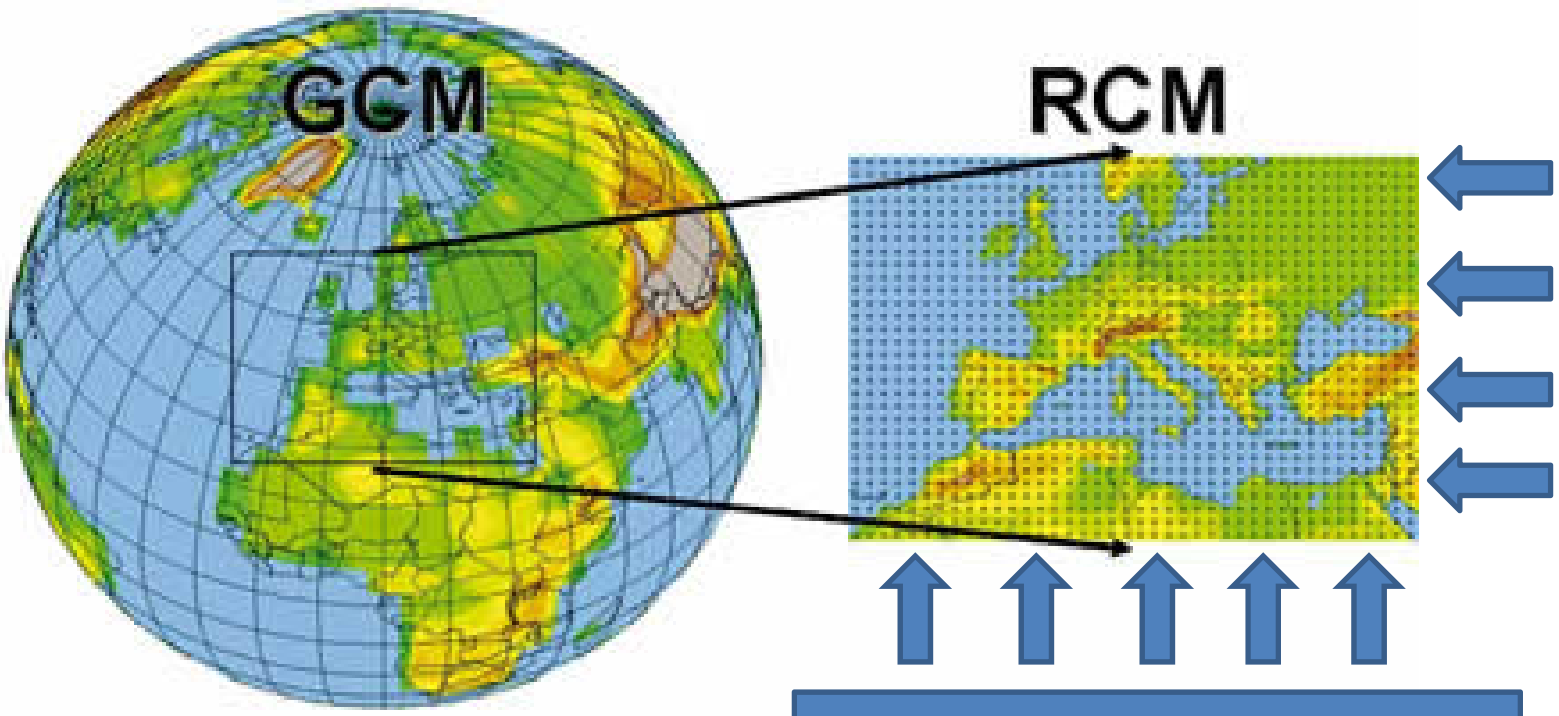
- **Désagrégation dynamique**: augmenter la résolution tout en maintenant des couts de calcul acceptable.

=> Ne simuler que l'atmosphère (conditions aux limites océaniques), à plus haute résolution

=> Ne simuler que l'atmosphère et utiliser une résolution variable, avec un zoom sur une zone d'intérêt (conditions aux limites océaniques)

=> Réduire le domaine simulé (conditions aux frontières latérales + limites océaniques si non couplés)

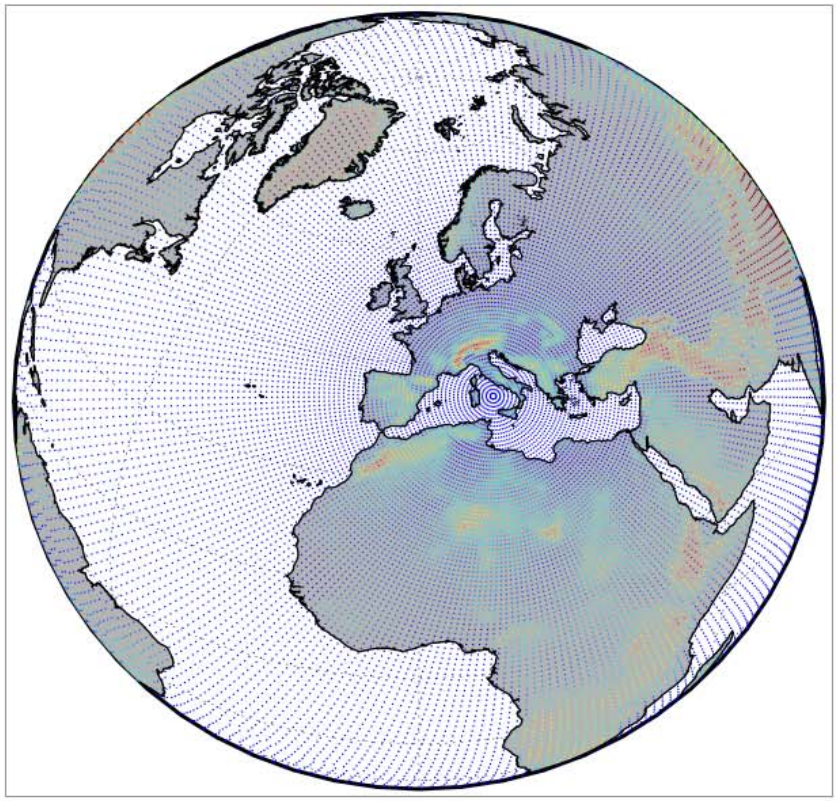
Modèle à aire limitée (souvent « Regional Climate Model »)

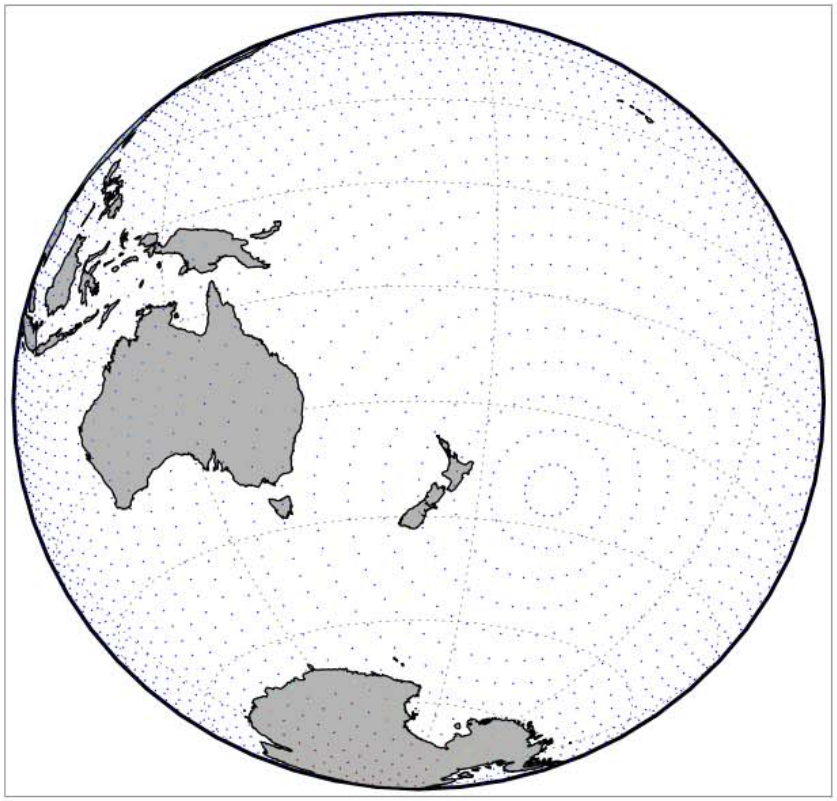


Giorgi, WMO (2008)

Conditions aux frontières du GCM
Surface: température de surface de la mer
Frontières latérales (vent, température, humidité etc.)

Modèle à résolution variable

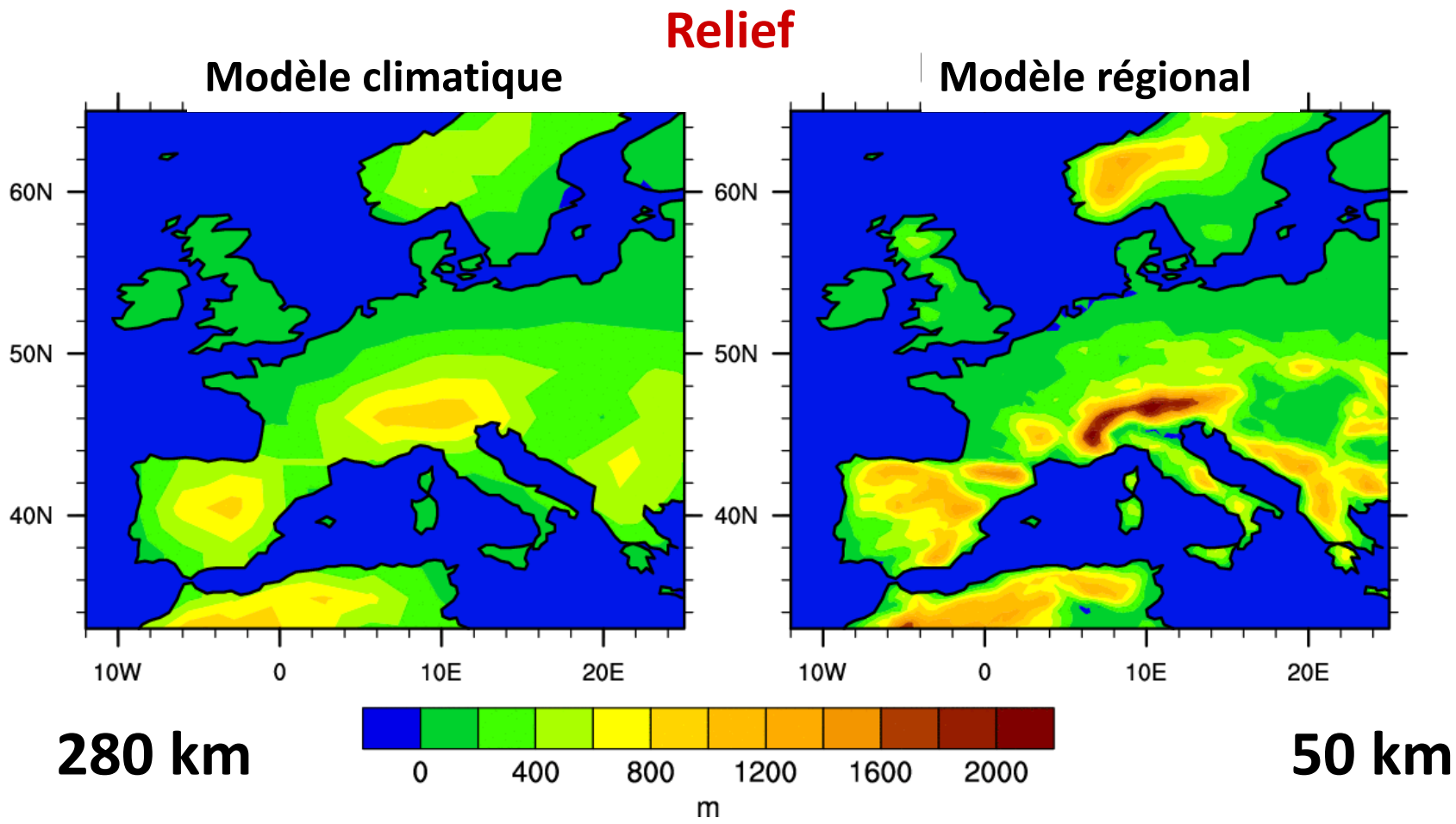




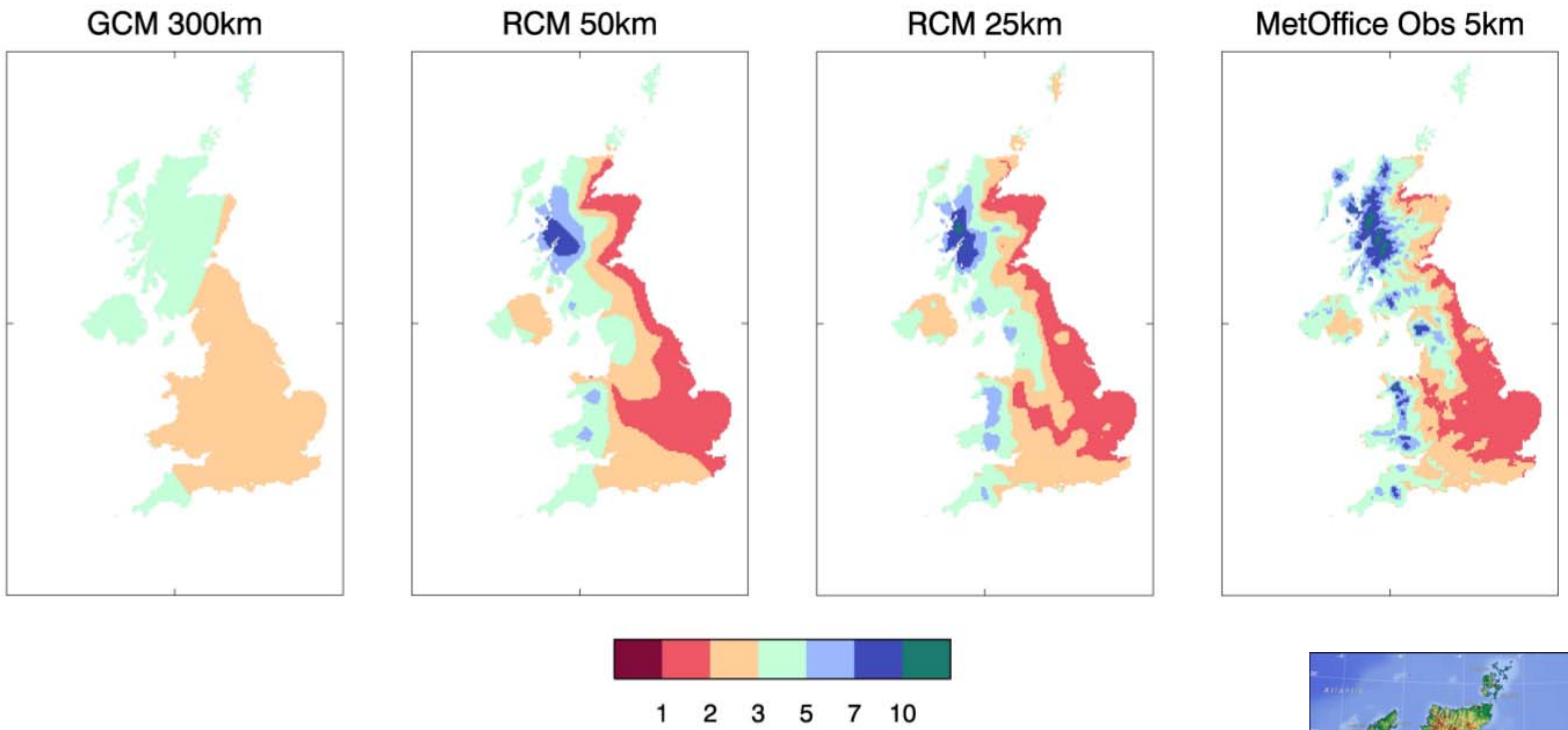
Grille du modèle ARPEGE climat dans une configuration à résolution variable

La désagrégation dynamique

Augmentation de la résolution d'un modèle atmosphérique
=> Meilleure représentation du relief, des traits de côtes etc.
Contraints par les modèles climatiques pour les conditions aux limites.



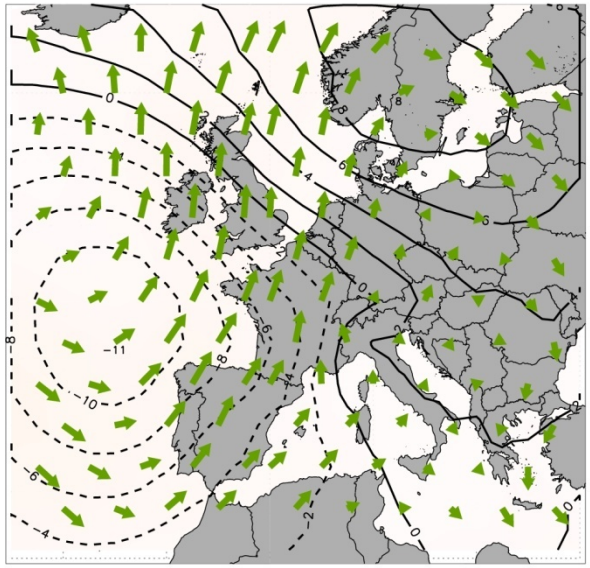
Apports de la haute résolution



Précipitations moyennes, 1961-2000. Modèle global et régional à différentes résolutions
Maraun et al. 2010

Apports /limitations de la haute résolution

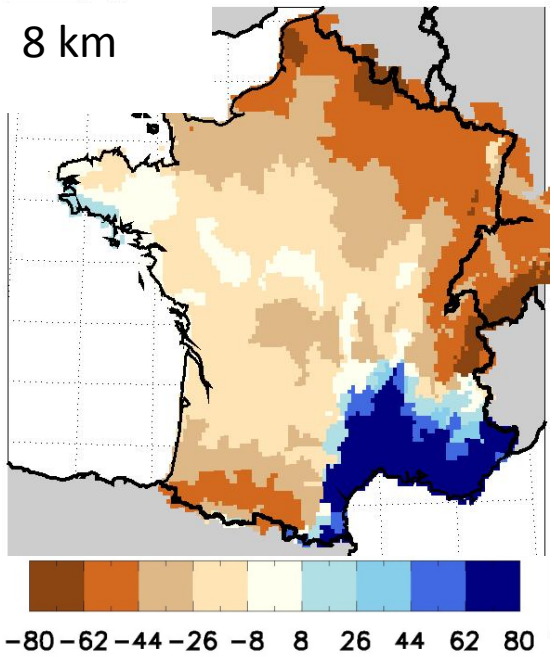
Anomalies relatives des précipitations associées (%)



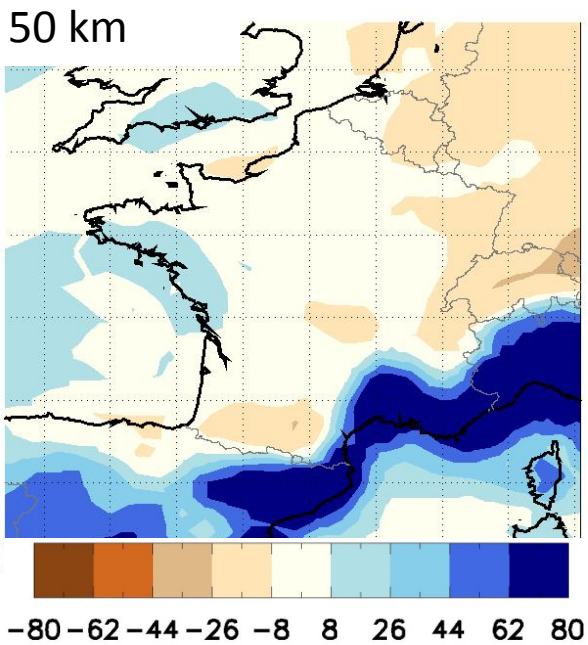
Type de circulation atmosphérique, flux de sud

Flèches: Vent 850 hPa

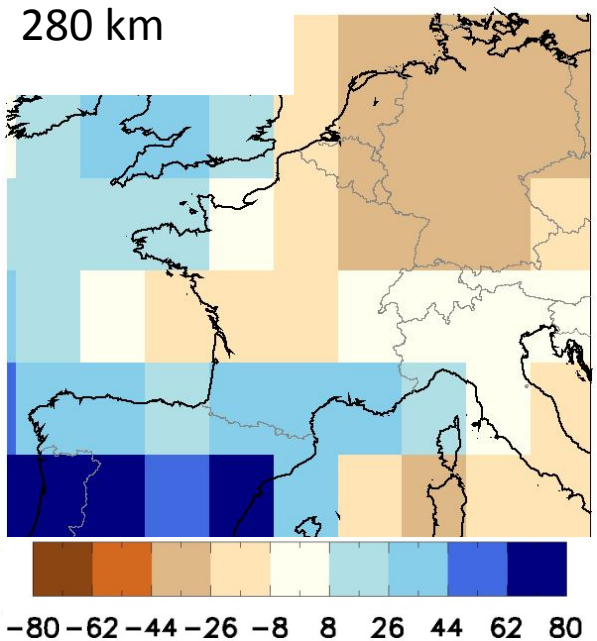
Lignes: Anomalies pression



Observations

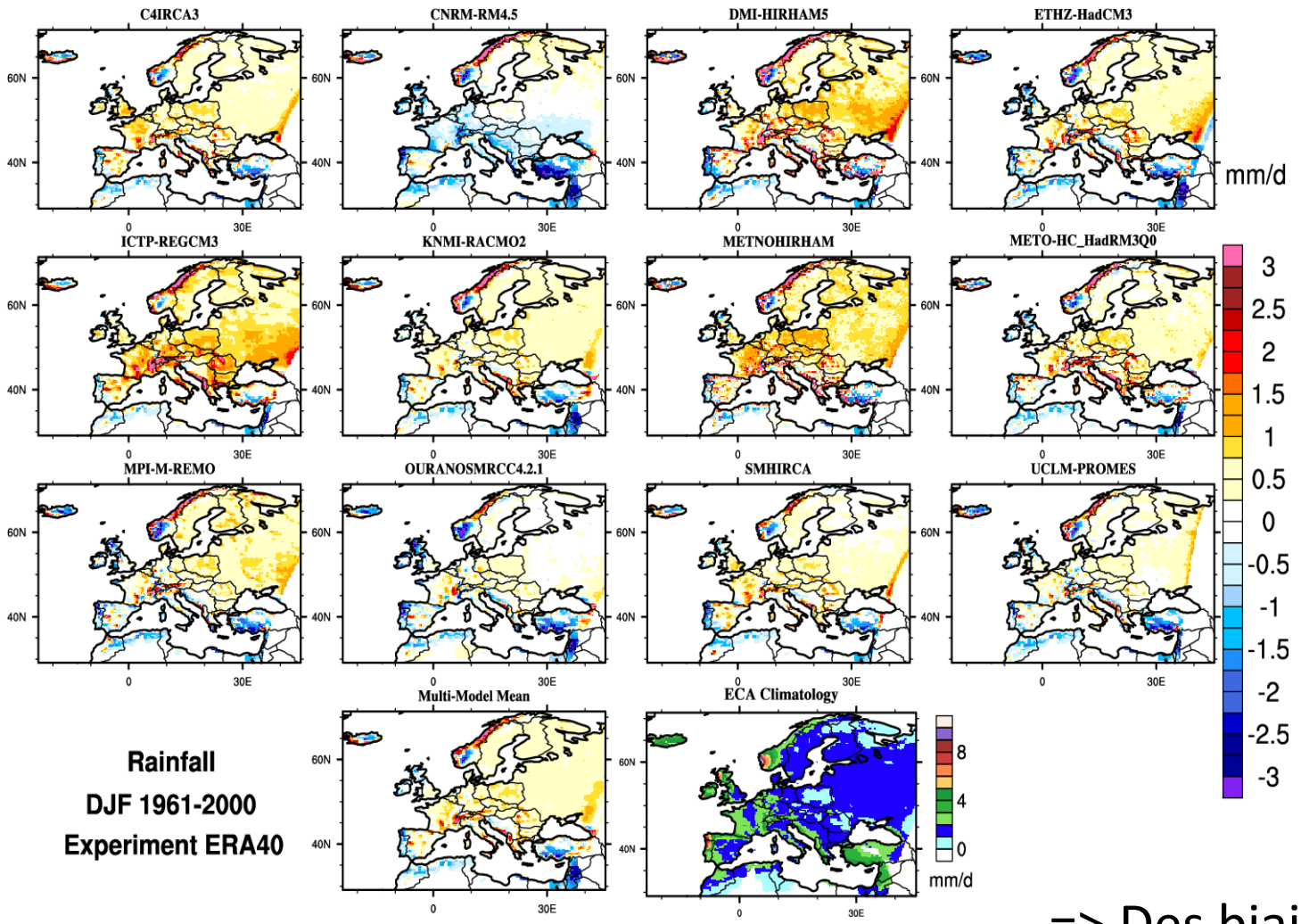


Modèle régional



Modèle climatique

Apports / limitations de la haute résolution

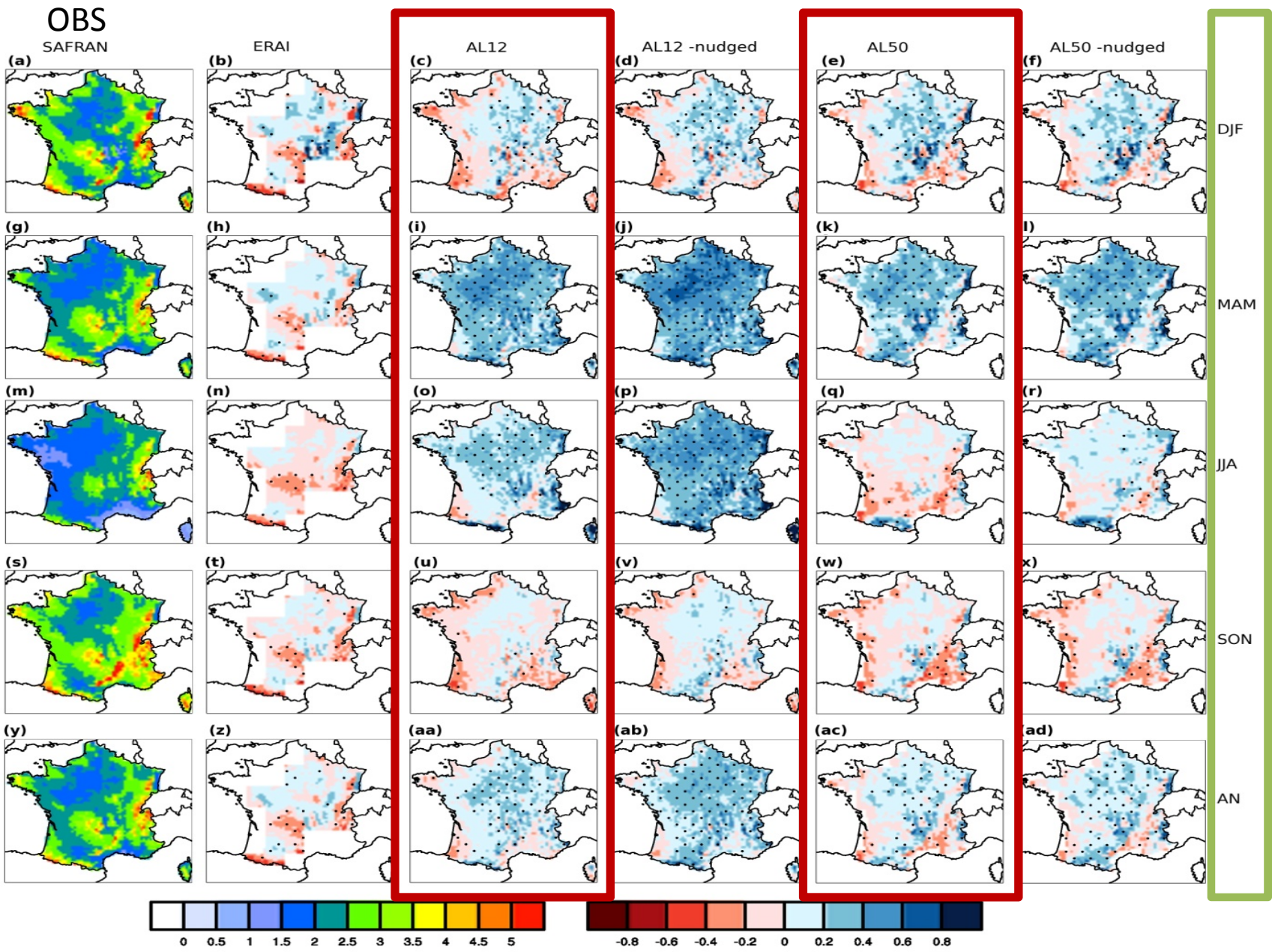


Biais des précipitations en hiver (mm/jour) des RCMs ENSEMBLES (25 km)

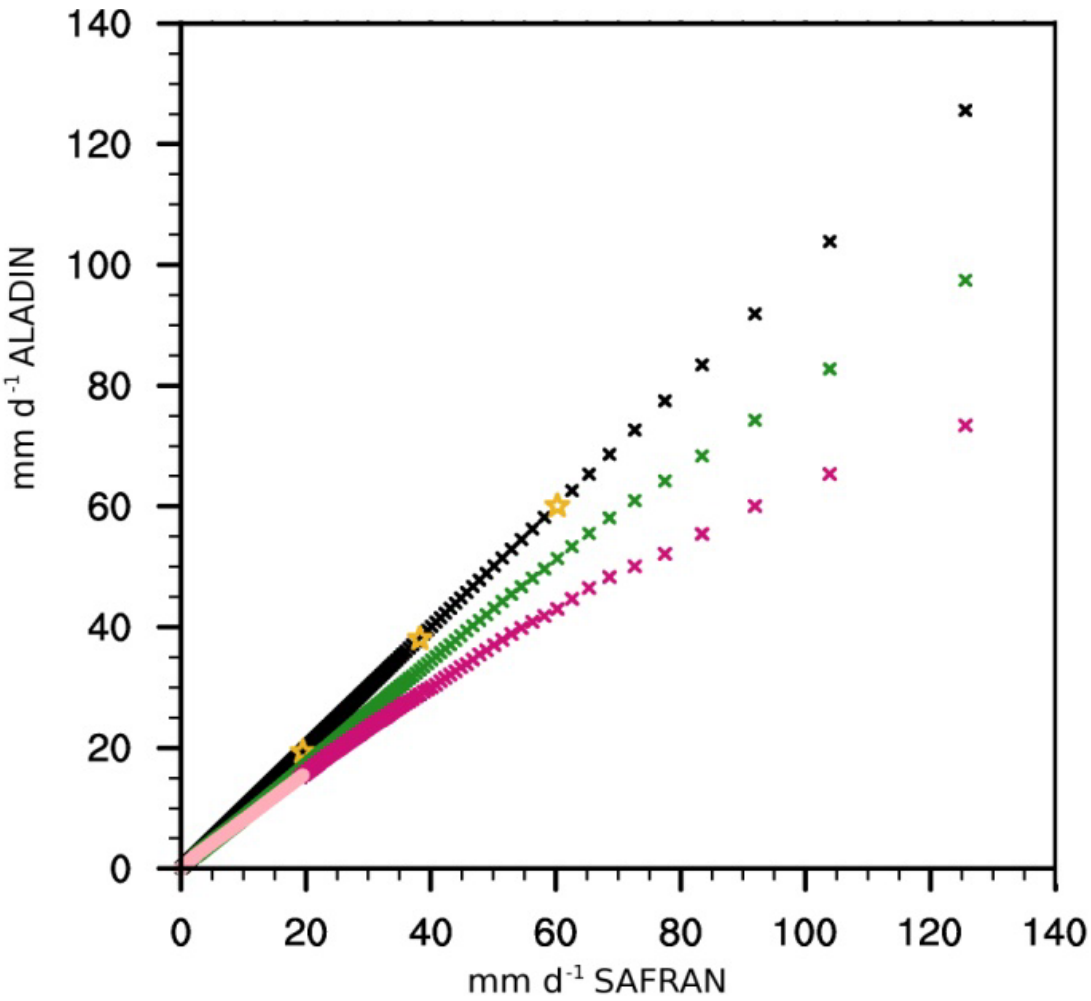
Source: Cyril Caminade

OBSERVATIONS

=> Des biais importants demeurent.



Climatologie des précipitations et biais (mm/jour) , 1981-2010. Desagrégation ERAI par différentes versions ALADIN, Harader et al. 2014



SAFRAN (OBS)
ALADIN 12km
ALADIN 50km

Clair: différences non significatives:
12km/50km
12km/SAFRAN

Milliles de précipitations ALADIN vs SAFRAN (obs), SOND, 1981-2010, Sud-est de la France. Harader et al. 2014

Downscaling dynamique

- **De claires améliorations: variabilité spatiale bien plus réaliste, extrêmes etc.**
- **Mais**:

⇒Approche couteuse en temps de calcul (mais bases de données multi-modèles, plus ou moins complètes…)

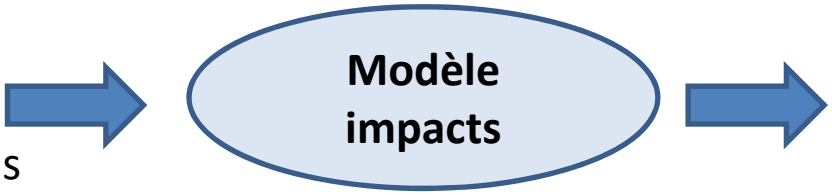
⇒Augmenter la résolution ne résout pas tous les problèmes: des paramétrisations physiques restent nécessaires, pour les mêmes processus que dans les GCMs (tant qu'on ne résout pas la convection humide explicitement: ~sous 5 km). Des biais demeurent.

⇒**Correction de biais est (presque toujours) nécessaire**

Correction de biais. Pourquoi?

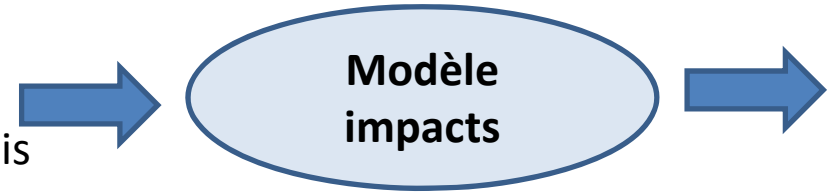
ex. changement climatique

Variables
entrées présentes
« justes » (Vp) + biais



Sorties présentes (Sp)
f(Vp,biais)

Variables
entrées futures
« justes » (Vf) + biais



Sorties futures (Sf)
f(Vf,biais)

Changements sorties => **Sf-Sp = f(**Vp,Vf**,biais)** (modèle non linéaire)

=> Que corrige-t-on? Etat moyen uniquement? Distribution statistique du modèle? Prise en compte des changements de variabilité ou non?

Deux méthodes pour corriger les biais moyens
(sous forme additive: e.g. température)

Méthode du Delta

Delta(x,y)= [Mf(x,y,t)]-[Mp(x,y,t)]

Mfc(x,y,t)=O(x,y,t)+Delta(x,y)

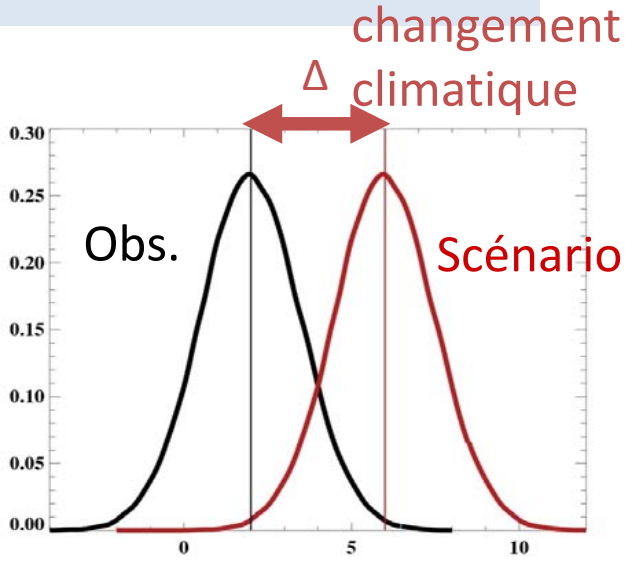
Mpc(x,y,t)=O(x,y,t)

Débiaisage

B(x,y)=[Mp(x,y,t)]-[O(x,y,t)]

Mfc(x,y,t)=Mf(x,y,t)-B(x,y)

Mpc(x,y,t)=Mp(x,y,t)-B(x,y)



Fonctions de densité de probabilité

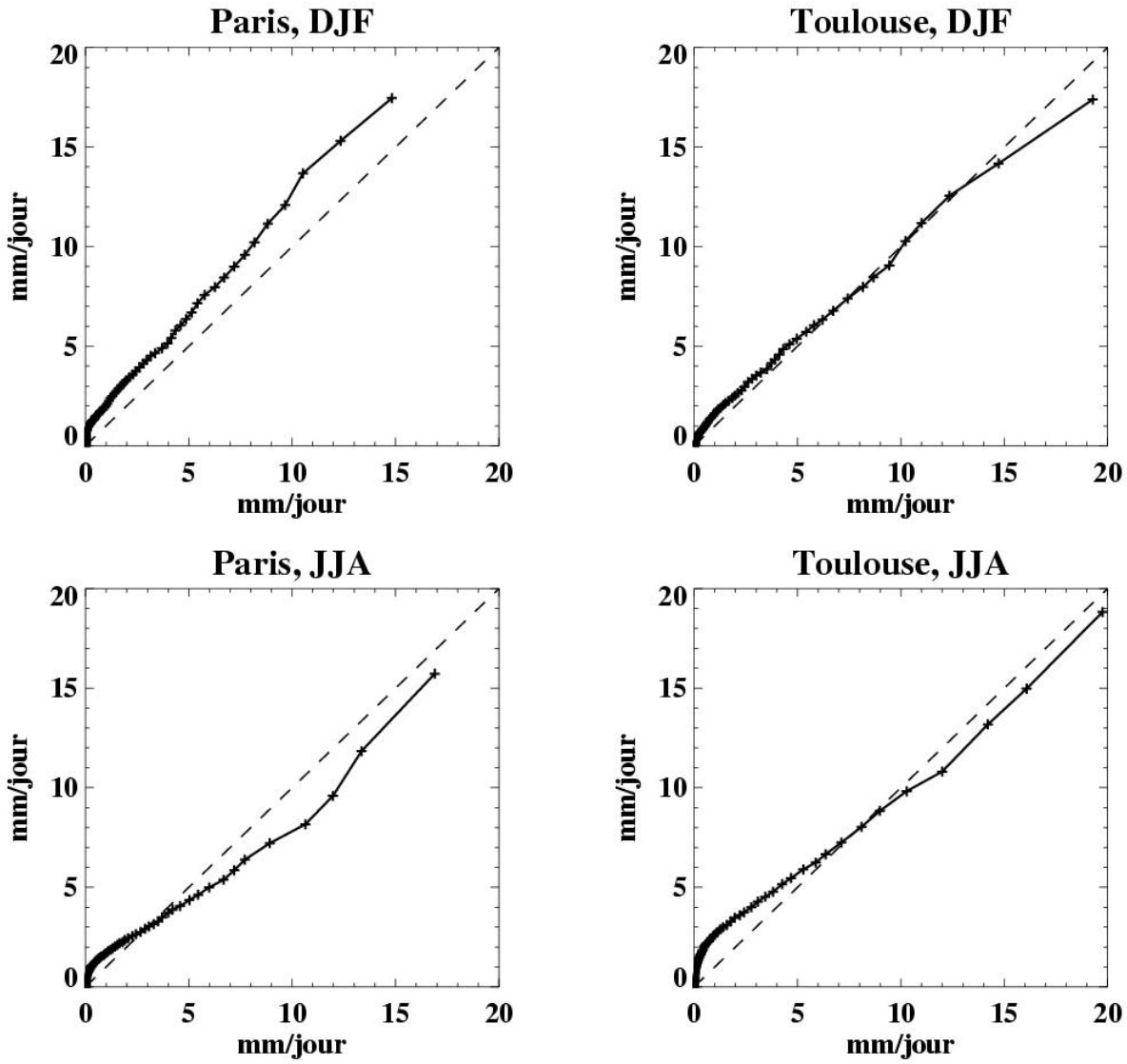
M: modèle
p,f: présent/futur
c: corrigé
B: biais
O: observation
[]: moyenne clim.

Biais dans la distribution

Correction de la distribution entière:

Méthode de correction quantile-quantile

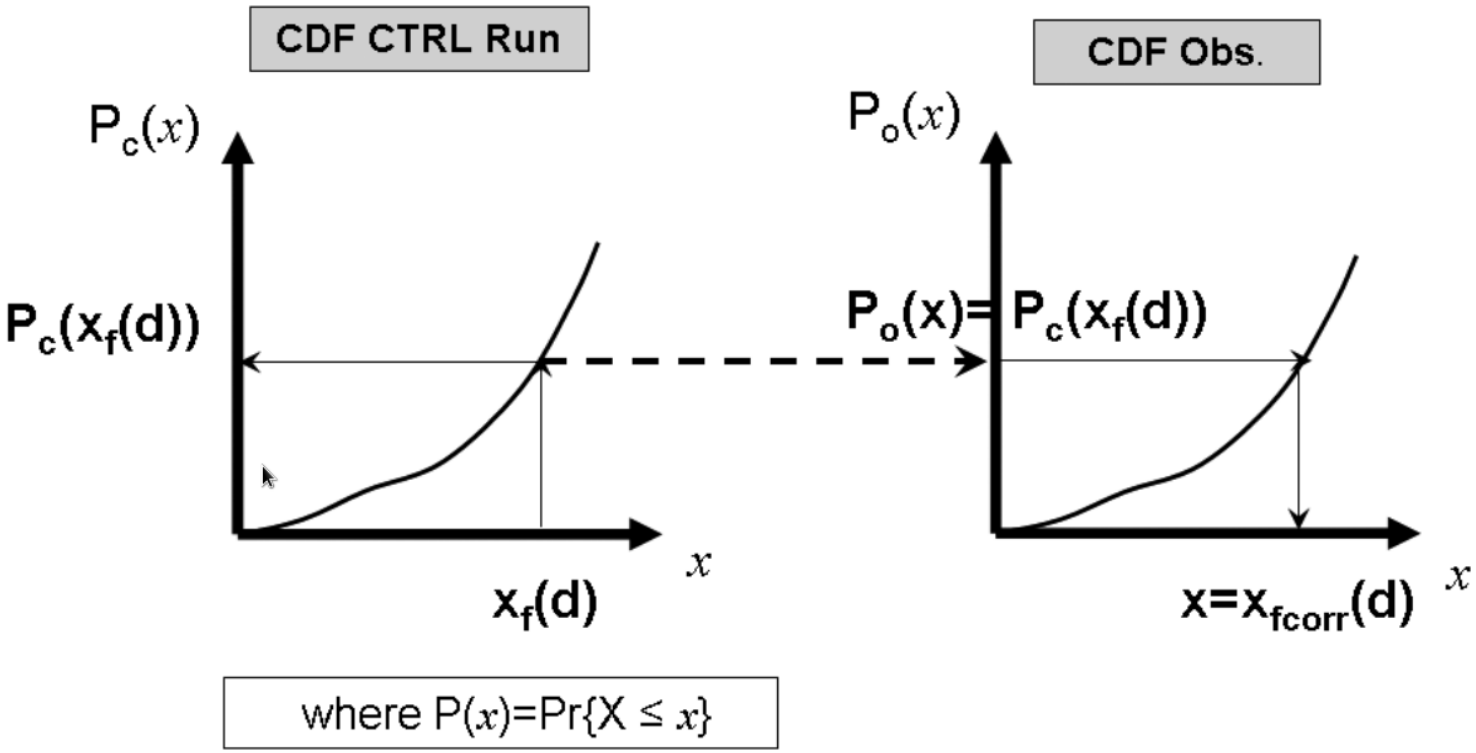
Déqué 2007, utilisée dans Boé et al. 2007 dans le contexte hydrologique. Différentes variantes:
Michelangeli et al 2009; Li et al. 2010, Thèse Jeanne Colin etc.



Centiles des précipitations simulées (y) vs.
Centiles des précipitations observées (x) (ARPEGE 50 km sur France)

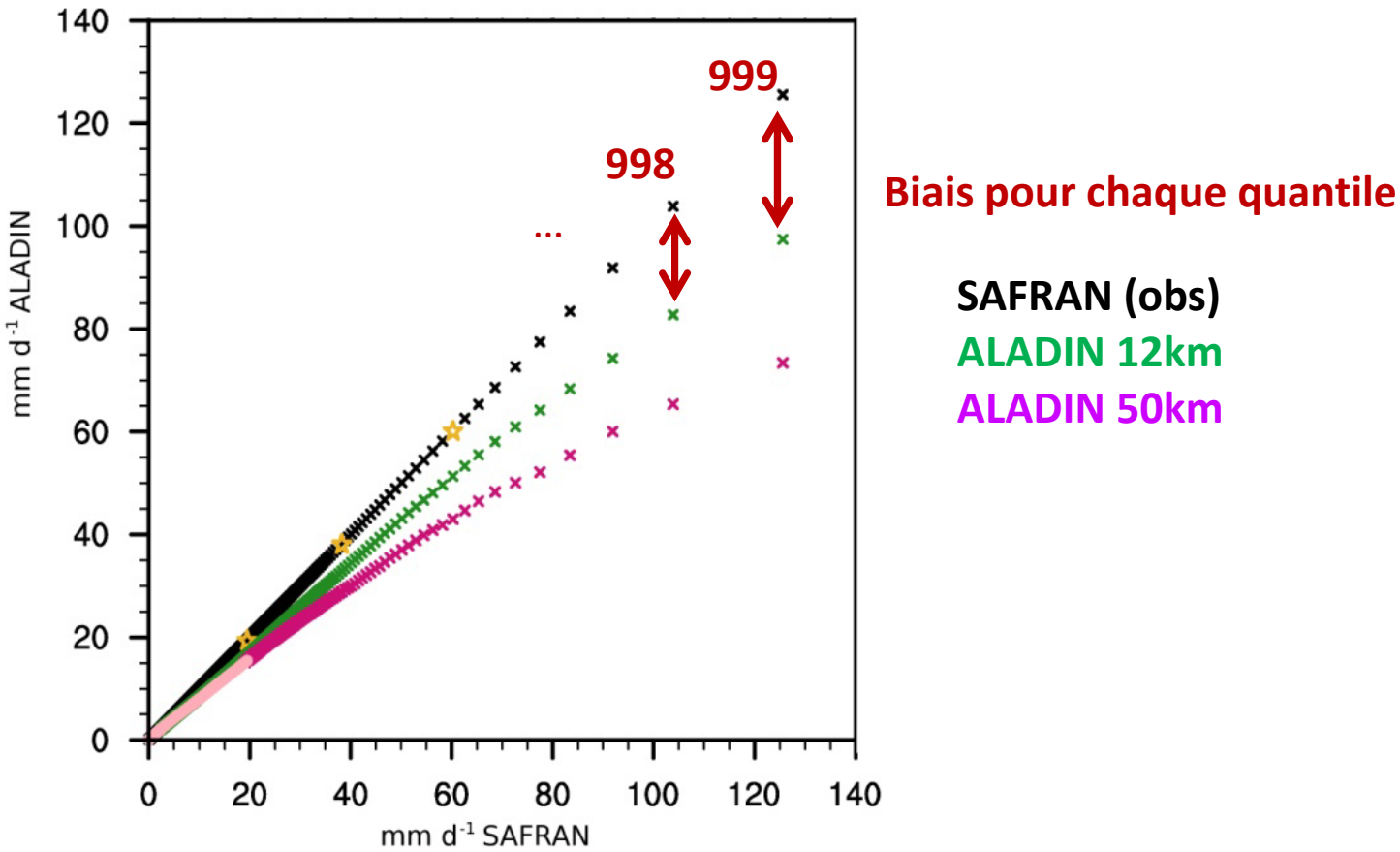
Méthode de correction quantile-quantile

Fonction de densité de probabilité cumulée



Boé et al. 2007

Autres corrections quantile-quantile



Milliles de précipitations ALADIN vs SAFRAN (obs), SOND, 1981-2010. Sud-est France
Harader et al. 2014

=> Méthode de débiaisage, par quantile plutôt que sur la moyenne
=> Egalement possible: Delta pour chaque millile (futurisation: thèse J. Colin)

Correction de biais

- Peut être utilisée sans désagrégation…
- Meilleure méthode?? => Dépend de la région / variable / application / qualité des observations etc.
- En plus, par définition, elles marchent toutes bien sur la période de calibration (présent)…
- **Difficultés**:

=>Hypothèse: les biais ne dépendent pas de la période (ne dépendent pas du climat moyen par exemple). Vrai?

=>Corrections indépendantes pour chaque variable: liens physiques après corrections? (e.g. température / humidité spécifique)

DESAGREGATION STATISTIQUE

La désagrégation statistique

Idée de base

$\mathbf{C_{pe} = F(C_{GE}, \Phi_s)}$

avec:

C_{pe} = le climat aux petites échelles (régionales/locales)

C_{GE} = le climat aux grandes échelles (globales/continentales)

Φ_s = les caractéristiques de surface aux petites échelles (orographie, contraste terre-océan, usage des sols)

⇒ **La désagrégation statistique consiste à établir une relation statistique entre les variables locales et les prédicteurs de grande échelle.**

$$\boxed{C_{pe}(t) = G[C_{GE}(t)] + \varepsilon(t)}$$

G ne change pas dans le climat perturbé

C_{GE} est simulé de façon réaliste dans les modèles

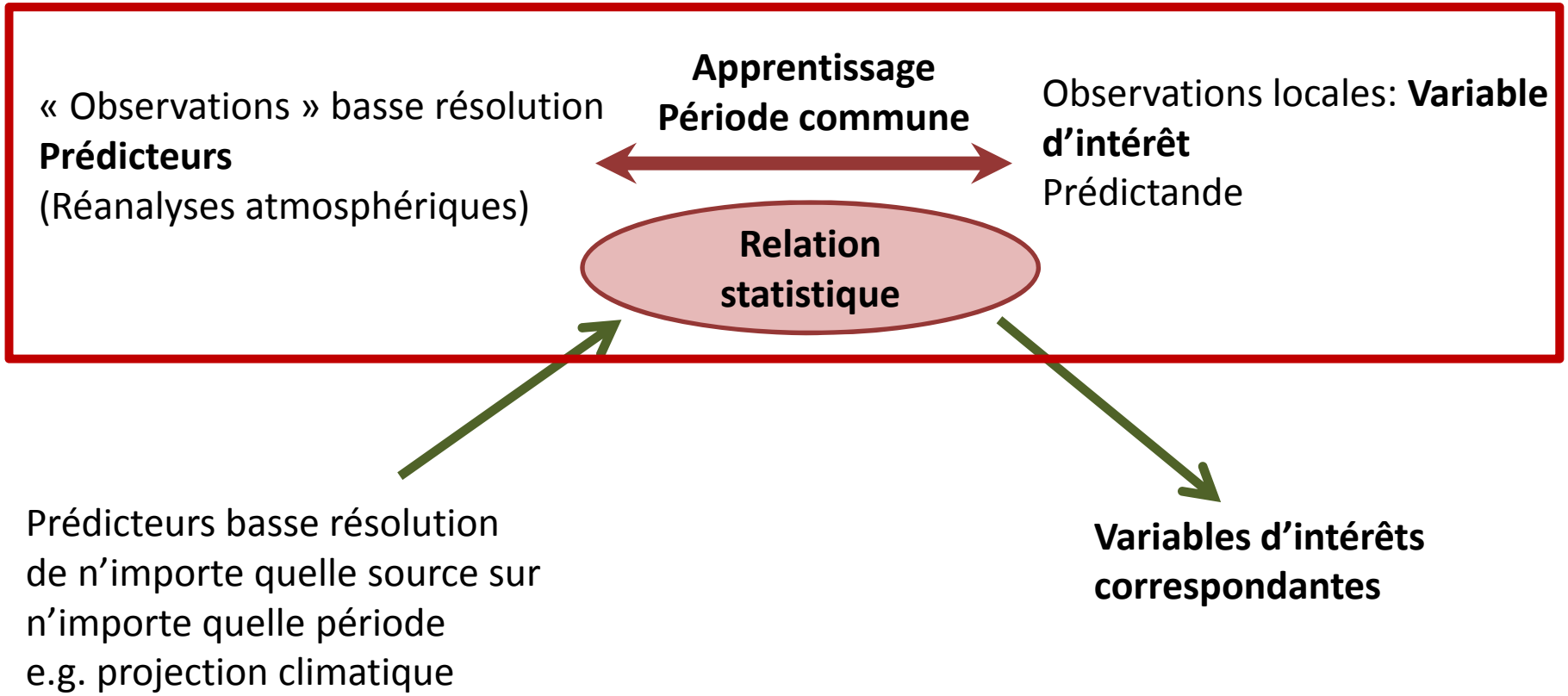
Ce qui n'est pas expliqué par GE.
=> Le changement climatique ne joue pas sur ε

G: différent en chaque point
(car Φ_s différent…)

Différentes approches statistiques pour définir G

Hypothèses (cadre changement climatique)

Principe



=> De multiples approches pour la relation statistique

Les principales familles de méthodes de désagrégation statistique

- Analogues et types de temps
- Régressions
- Générateurs de temps

Analogues / types de temps

- Type de temps = ensemble de jours ayant une circulation de grande échelle similaire. Souvent obtenus par classification automatique.

=>Les mêmes causes (de grande échelle) produisent les mêmes effets (de petite échelle).

Type de temps

=> Définition de types de temps discriminants pour précipitations en France

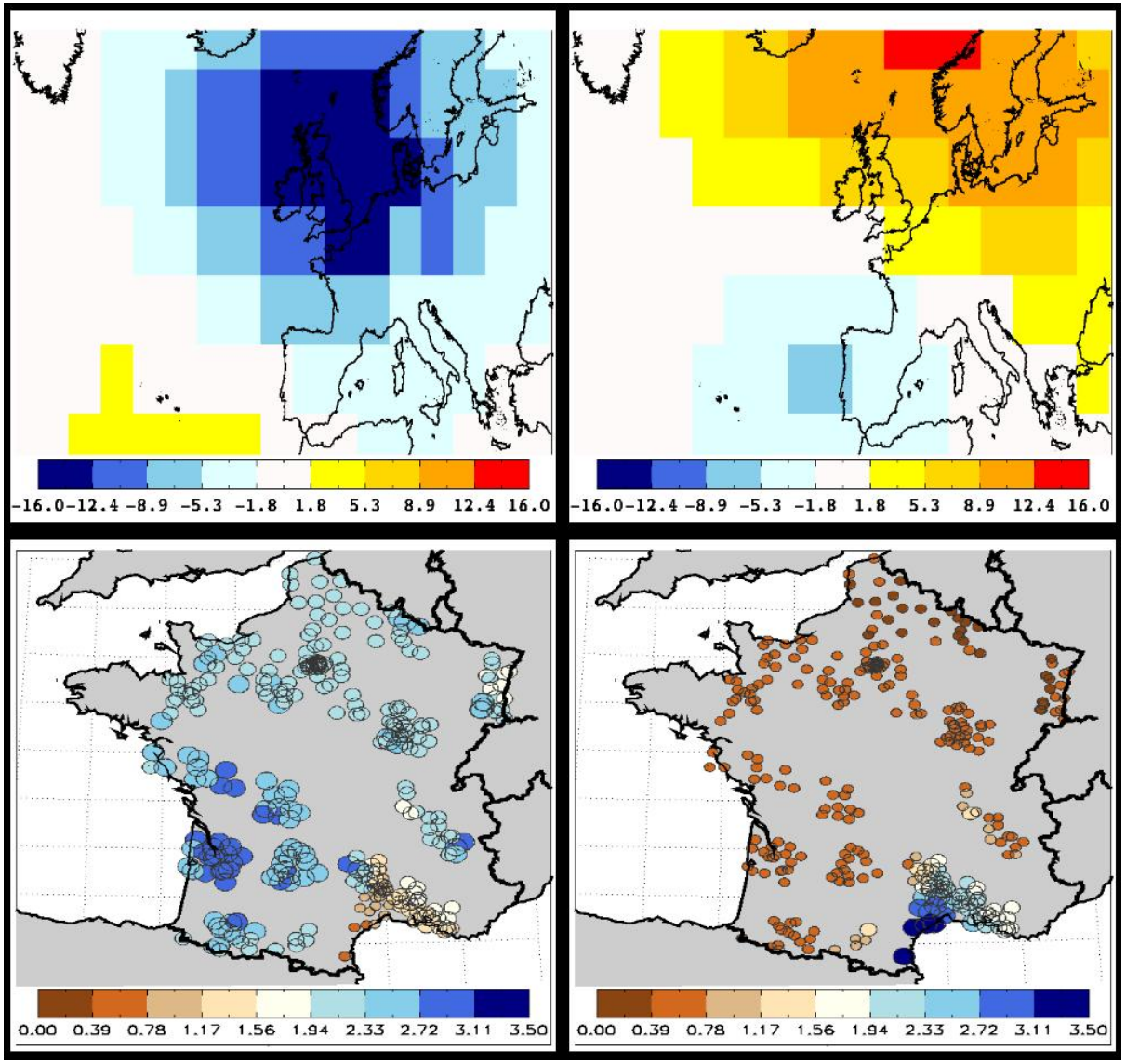
•Variable de circulation de grande échelle: Pression (MSLP)

=>8 types de temps

Boé et Terray 2008

Hiver NDJFM
Composites de MSLP (hPa)

Données Emulate



Composites de Précipitation (rapport entre $P_{type}/P_{moyenne}$)

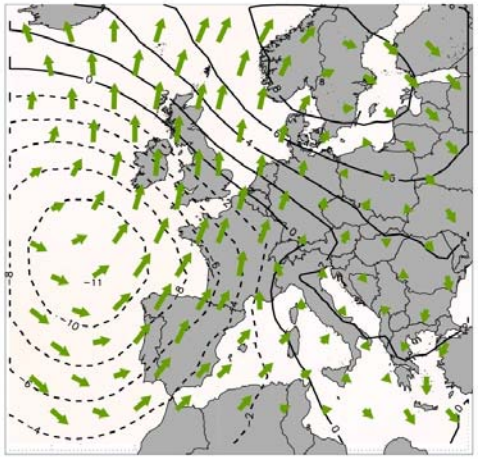
Données Météo-France

Analogues / types de temps

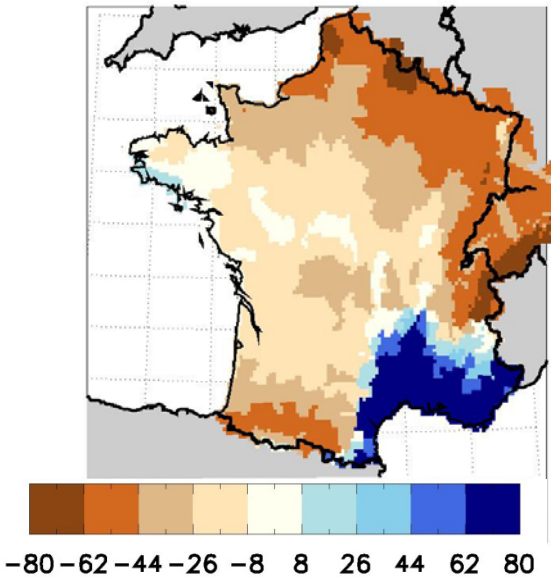
- Type de temps = ensemble de jours ayant une circulation de grande échelle similaires. Souvent Obtenus par classification automatique.
- =>Les mêmes causes (de grande échelle) produisent les mêmes effets (de petite échelle).
- **Analogues: cas limite où chaque jour est un type de temps à lui tout seul**

Flèches: Vent 850 hPa
Lignes: Anomalies pression
OBSERVATIONS

Période apprentissage



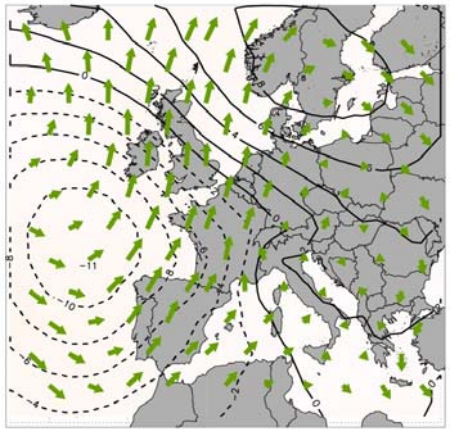
Même jour(s)



Observations

Similarité

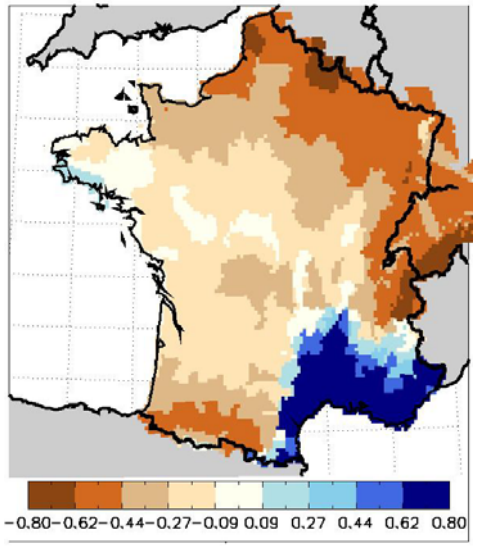
(même cause de grande échelle)



Circulation jour j
(e.g. modèle futur)

Même effet
(de petite échelle)

Anomalies relatives des précipitations associées (%)



Données désagrégées correspondant au jour j

Analogues / types de temps

- Définition de la similarité: mesure de distance & variables (prédicteurs)
- Basé sur du « resampling » de la période d'apprentissage:

=> Maintien de l'auto-corrélation spatiale et le lien entre variables (beaucoup plus compliqué avec les autres approches)

=>Mais: les plus fortes valeurs de la période d'apprentissage ne sont jamais dépassées …

(température, problème climat futur, même pour représenter les changements moyens par exemple)

Régression

- **Régression linéaire**

=> $\mathbf{E(V_i)=a_0+a_1.X_1+a_2.X_2+\ldots}$ X_1, X_2 prédicteurs de grande échelle, V_i variable d'intérêt à l'échelle locale. Dépendent du temps.

- Pb: Variance non expliquée
- Pb: Hypothèse gaussienne. E.g. pas le cas pour les précip journalières…

- **Modèle linéaire généralisé**

$\Rightarrow$ $\mathbf{g(E(V_i))=a_0+a_1.X_1+a_2.X_2+\ldots}$ Avec g une fonction lien, transformation pour que l'influence des prédicteurs devienne linéaire.

- **Modèle additif généralisé**

- **Réseau de neurones artificiels**

…

$\Rightarrow$ **Très simple dans le cadre gaussien, moins quand on en sort...**

$\Rightarrow$ **Prise en compte des dépendances spatiales et inter-variables: compliqué…**

Générateurs de temps

- **Modèles statistiques qui génèrent des séquences de temps aléatoires avec des propriétés statistiques similaires au temps observé** **=>** Séries temporelles de taille illimitée, cout de calcul très faible.

Exemple le plus simple. En deux temps. Occurrence pluie puis intensité

(i) Occurrence de pluie: chaine de Markov de 1er ordre à deux états. **Proba. de transition $p_{ij}(t)$** avec I(t): sec ou humide => **$p_{ij}(t)=Pr(I(t)=j \mid I(t-1)=i)$**, i,j=0,1

(ii) Intensité des précipitations: valeur aléatoire sur distribution statistique (e.g. distribution gamma). (=> indépendance d'un jour sur l'autre...)

(iii) Désagrégation changement climatique: simple perturbation des paramètres du modèle statistique (e.g. moyenne de la distribution)

En pratique, il faut compliquer:

e.g. **Conditionnement** de l'occurrence et de l'intensité par la circulation atmosphérique => meilleure auto-corrélation temporelle, représentation de la variabilité et du changement climatique…

Application spatiale: modéliser en plus statistiquement les dépendances spatiales

Application multi-variée: modéliser les interdépendances entre variables…

=> Compliqué…

Désagrégation statistique

- **Pas de méthode parfaite, ni universelle**.
- **Choix de la méthode dépend de l'application souhaitée**:

  1 seule station ou un champ?

  1 seule variable ou plusieurs variables?

  Variables? Journalier ou mensuel? (i.e. quelle distribution statistique?)

  Extrêmes, critiques ou pas?
- Aussi sophistiquée que soit la méthode stat., s'il n'y a pas la bonne physique (**i.e. prédicteurs**) derrière, ça ne marchera pas dans une période différente => crucial de comprendre la physique du problème.
- Beaucoup de littérature «théorique». En pratique, application spatiale et surtout multi-variée -ce dont on a besoin le plus souvent-: rare…
- **Difficulté d'évaluation**: c'est assez facile de faire une méthode qui marche correctement sur la période historique… Autre période? Climat futur?

Calibration: difficultés dans un cadre non stationnaire

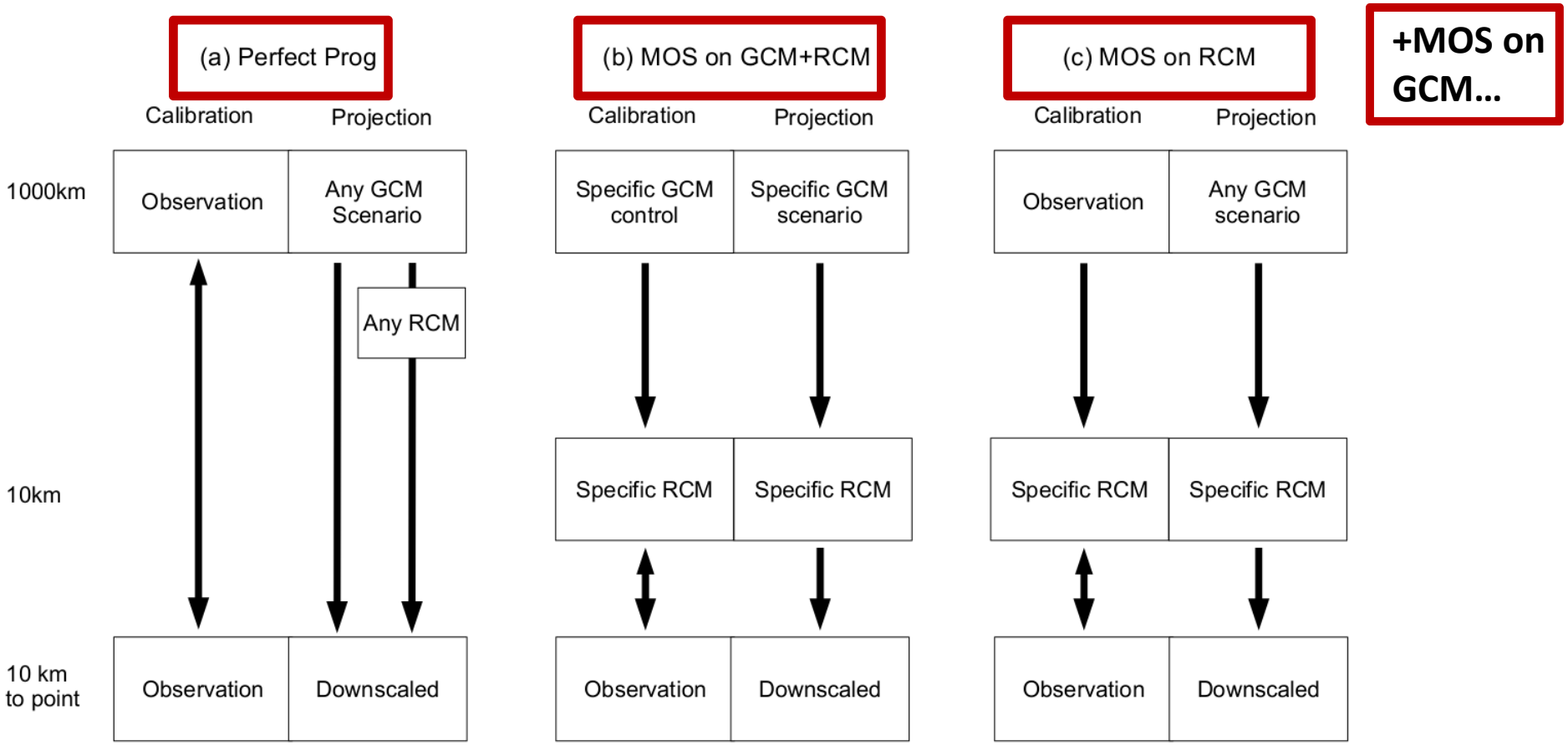
- **Séparation dynamique / statistique pas toujours si nette en pratique.**

=> On peut faire les deux à la suite: dynamique puis statistique (~ correction de biais)

=> Désagrégation dynamique: correction **statistique** des biais en général nécessaire.

=> Toutes les approches ont une phase de calibration aux observations

Classification alternative



Calibration
Désagrégation

Perfect Prog: Perfect prognosis
MOS: Model output statistics
Maraun et al. 2010

Problème commun

- **Calibration** se fait nécessairement dans le **climat présent**

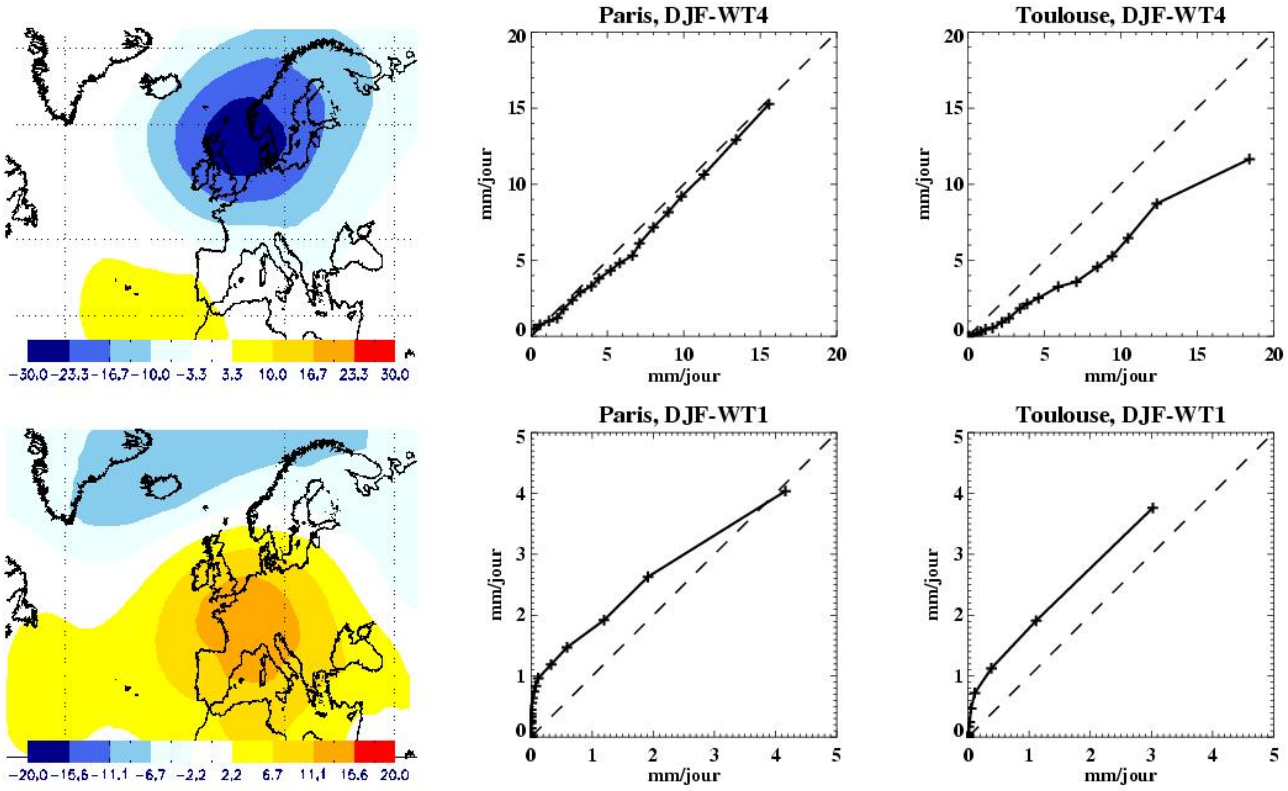
  => Reste valide hors de la période de calibration? Notamment dans le climat futur?

  => Dépend des mécanismes que l'on prend en compte.
- Difficulté: évaluation aussi dans le climat présent…

Correction de biais. Difficultés

Anomalie pression de surface

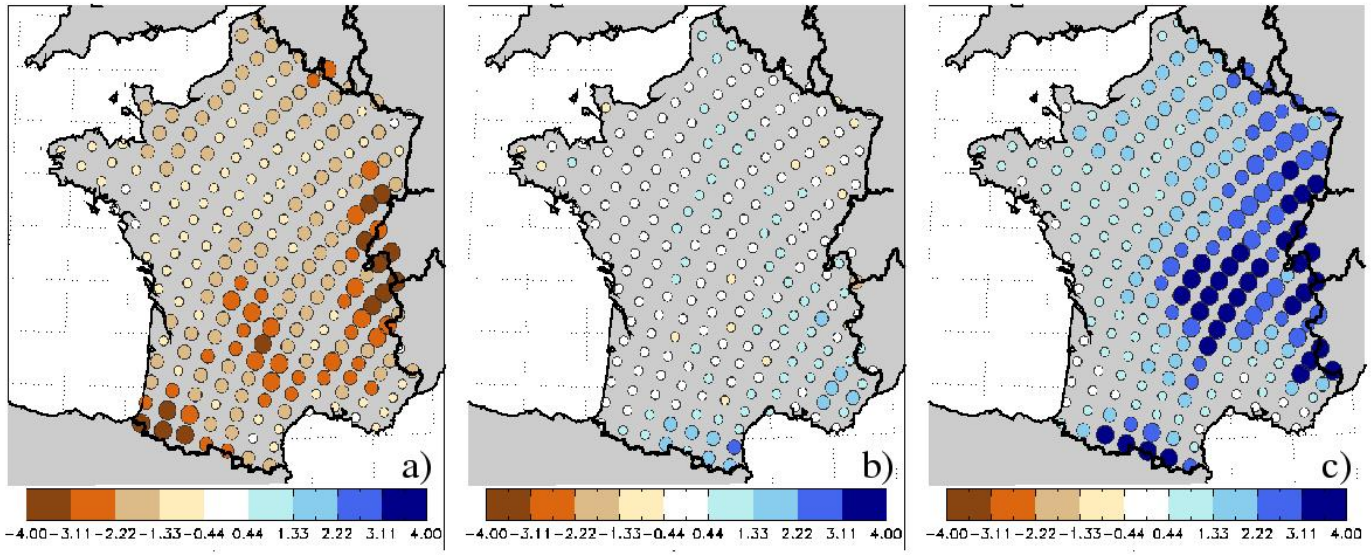
hPa



Centiles des précipitations simulées (y) vs.
Centiles des précipitations observées (x)
(ARPEGE 50 km sur France)
pour deux types de circulations en hiver (Boé 2007)

Correction de biais. Difficultés

mm/jour



1er décile | 5ème décile | dernier décile

Biais des précipitations en été
en fonction du décile de température
(ARPEGE climat). Boé 2007

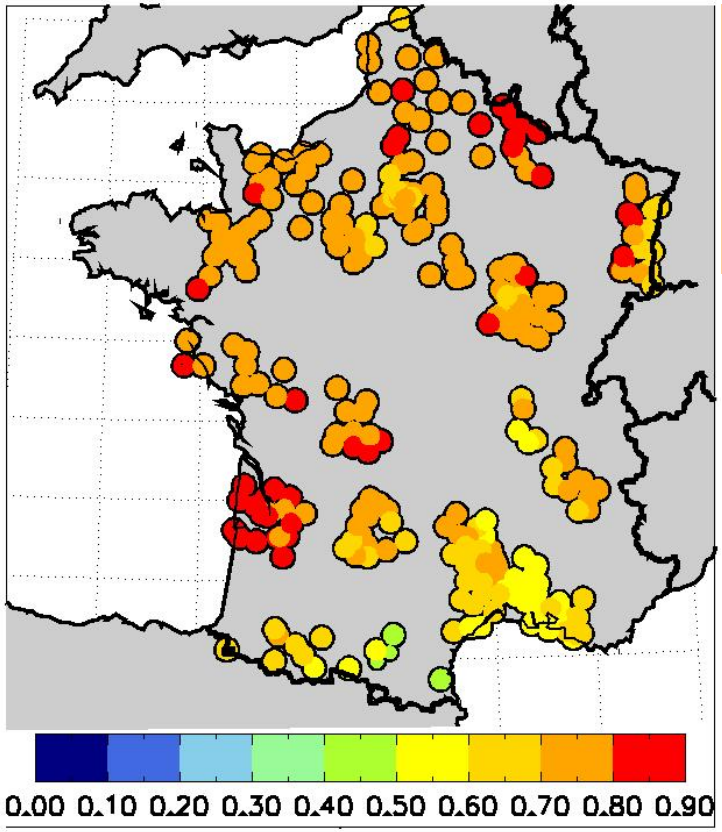
⇒Covariabilité entre variables (et entre leurs biais)
⇒Correction quantile-quantile conjointe température / précipitations

Désagrégation statistique: difficultés

⇒ Reconstruction du cumul saisonnier (NDJFM) des précipitations sur le 20ème siècle par régression multiple avec comme prédicteurs l'occurrence des types de temps et les distances au centroide

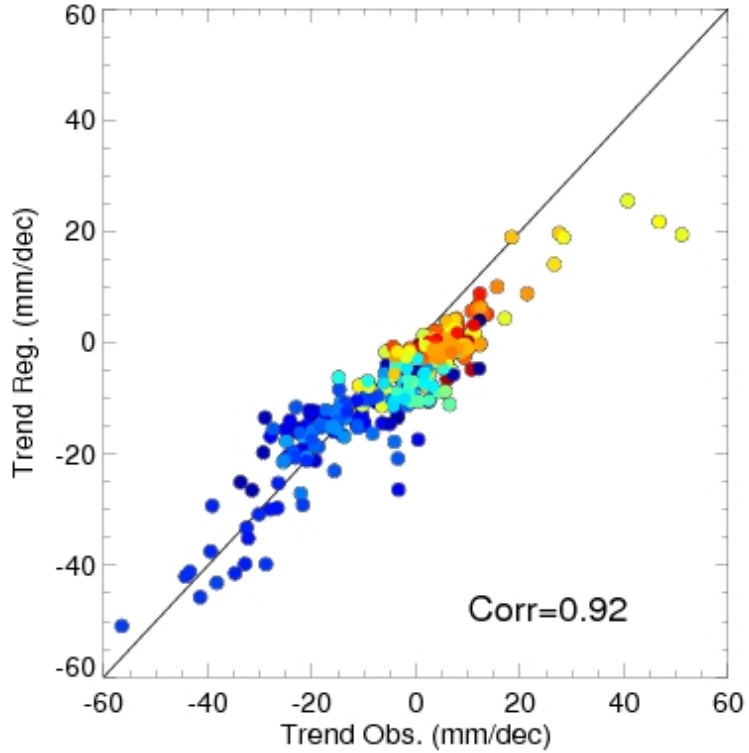
⇒ **Prédicteurs: circulation uniquement**

Données Météo-France



Corrélation observation /reconstruction 1900/2000

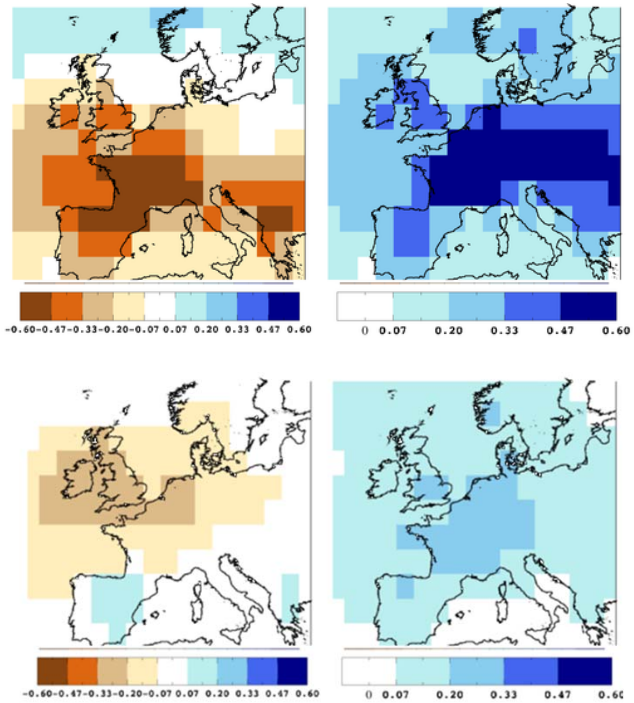
Données Météo-France



Tendances Pr 1951-2000 observation vs reconstruction

1 point=1station, couleur: latitude => bleu=sud

Désagrégation statistique: difficultés?



-0.6 0.6 0 0.6

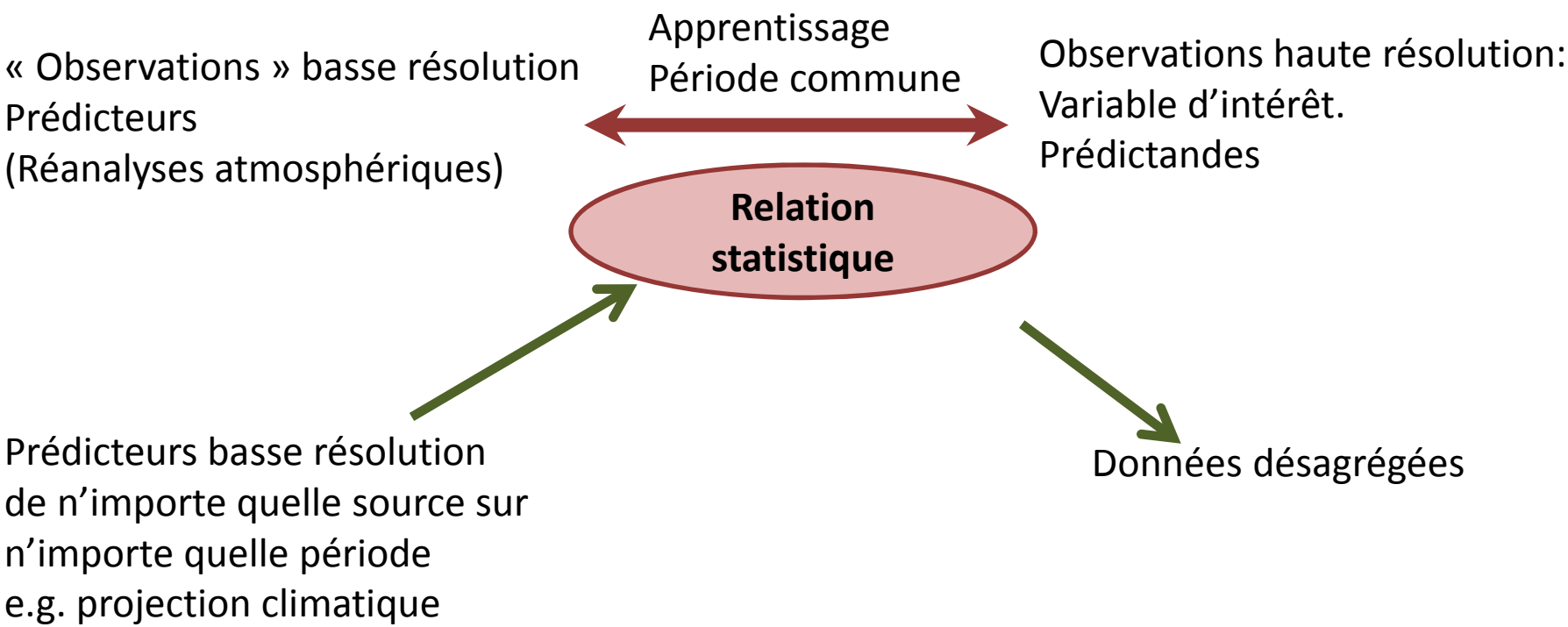
Boé et al. 2009

(gauche) Moyenne d'ensemble du changement de **précipitations en été** (mm/jour ; 2081-2100 vs. 1961-2000)
(droite) Dispersion (1 sigma) correspondante.
15 modèles CMIP3

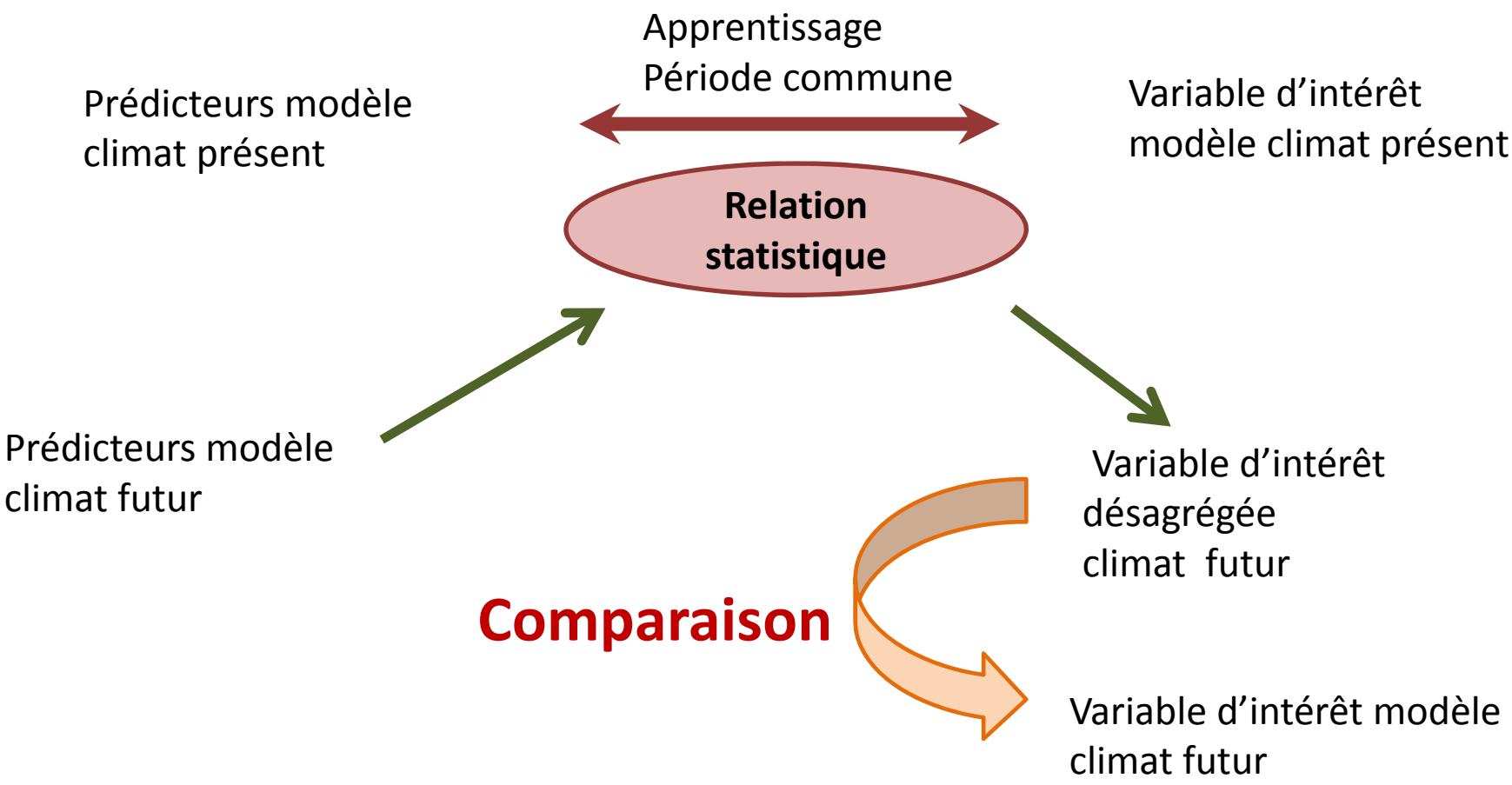
Changements de précipitations **dus aux changements de circulation** (estimer avec la méthode des analogues sur la pression réduite au niveau de la mer)

⇒Choix des prédicteurs : choix des processus physiques représentés…
⇒Comment savoir s'ils sont pertinents pour un climat différent?

Test modèle parfait

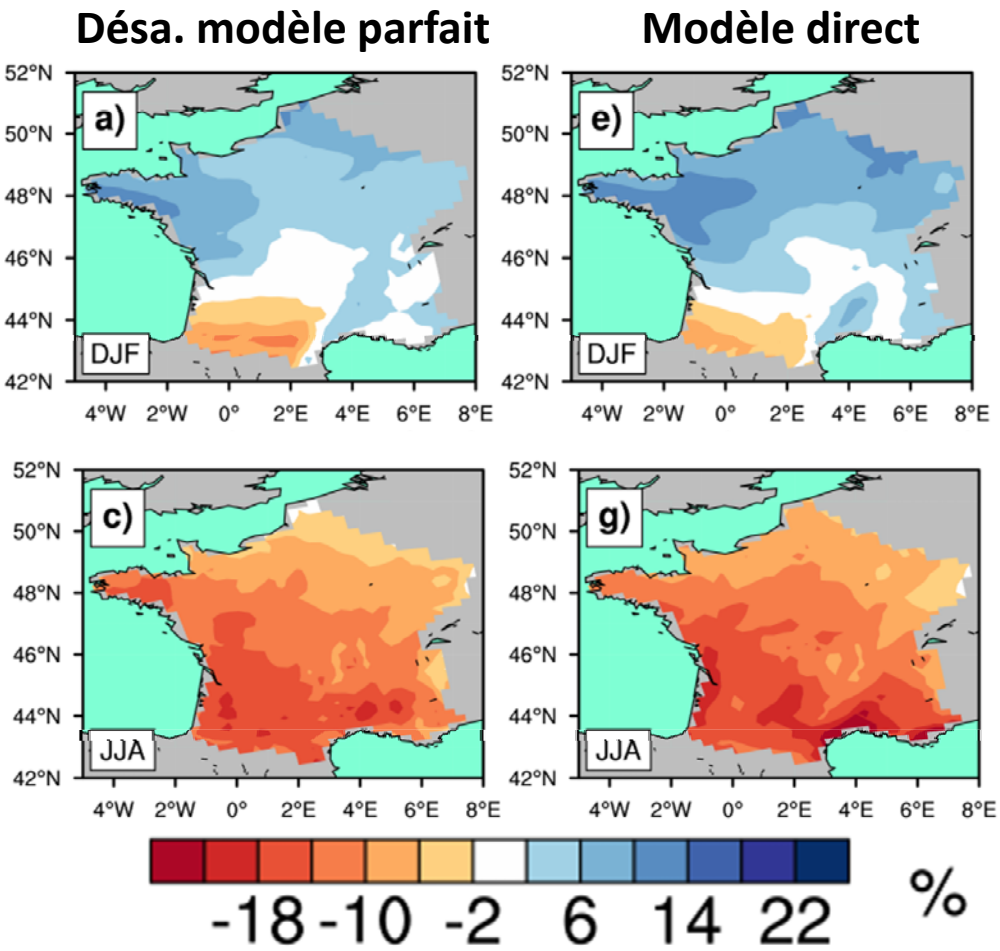


Test en modèle parfait



=> Méthode applicable dans le climat futur (pour ce modèle…): on doit obtenir la même chose

Test en modèle parfait.



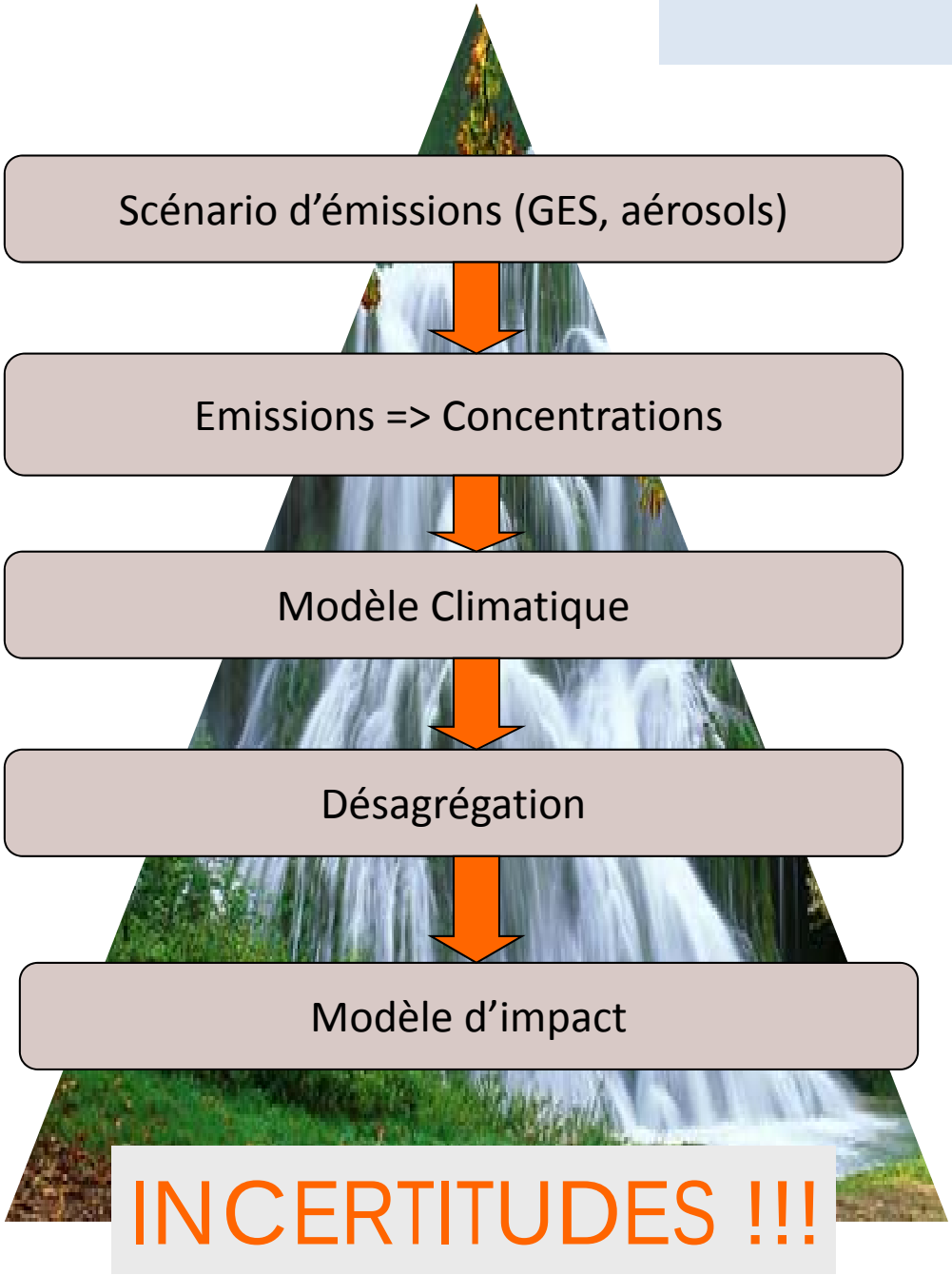
Moyenne d'ensemble
12 modèle régionaux

Méthodes des **analogues** avec comme prédicteurs:
-Pression niveau mer
-Température air surface
-TTI
-Humidité spécifique à 850 hPa
-Flux d'humidité à 850 hPa

Changements précipitations [2030-2050] – [1961-1990]
Dayon et al. (2014)

CONCLUSIONS

Attention…



Désagrégation = juste une étape d'une étude des impacts du changement climatique.
e.g. même avec une méthode de désagrégation "parfaite", si un seul modèle climatique utilisé: =>problème.
Mieux vaut une méthode simple appliquée à de multiples modèles (e.g. Delta) qu'une méthode plus complexe mais appliquée à un seul modèle…

Conclusions

- Correction de biais et désagrégation statistique =>**besoin d'observations** (bonne **couverture spatiale**, **longue période** disponible, **homogénéité** temporelle etc.)

- Choix dynamique vs statistique? Cout de calcul

=>Aujourd'hui, bases de données de projections climatiques régionales (ENSEMBLES, CORDEX), jusqu'à 12km. Mais moins complètes que les bases globales…

=>Choix de l'approche: dépend de l'application et des objectifs

Conclusions

•Une différence est la confiance qu'on l'on accorde aux modèles climatiques (globaux et régionaux)
Statistique: confiance dans la grande échelle (mais quid des autres processus dans le futur, des rétroactions locales?) =>*On représente "parfaitement" un nombre limité de processus*
Dynamique: confiance dans toutes les variables simulées (mais quid des biais et surtout de leur importance dans les changements simulés?) =>*On représente tous les processus, mais certains de façon imparfaite?*
=> Les **deux approches** sont justement **complémentaires**: Intérêt d'utiliser les deux. Convergence des résultats => confiance.

•Calibration puis utilisation sur une période différente, dans un autre climat: problème majeur de la désagrégation statistique et de la désagrégation dynamique avec correction de biais. Ne pas oublier que **c'est un problème physique autant que statistique**

Conclusions

- Attention aux utilisateurs de données désagrégées: elles ont été faites en pensant à une application, et sont plus ou moins adaptées à d'autres applications (est-ce que les extrêmes sont cruciaux, dépendance entre les variables…)

- Attention aux estimations de l'incertitude due à la désagrégation

- Dans tous les cas, on fait des hypothèses implicites sur les mécanismes sous-jacents (interpolation améliorée, correction de biais, désagrégation statistique) => **importance de comprendre ces hypothèses et les limitations associées** pour l'interprétation des résultats.

Très bonne revue sur désagrégation des précipitations dans le cadre du changement climatique: Maraun et al. 2010, Reviews of Geophysics.