

# Evaluation des modèles/simulations pour les services climatiques Introduction et activités dans TRACCS

**Jérôme Servonnat (LSCE-IPSL), Aurore Voltaire (CNRM)**

Masa Kageyama, Julie Deshayes, Cécile Caillaud, Benoît Hingray, Juliette Mignot, Etienne Perez,  
Aurélien Ribes, Samuel Somot, Benjamin Sultan, Mathieu Vrac, et Samuel Morin

Une information utile à la prise de décision est (notamment)

**Pertinente    Crédible**

Cash et al. (2002): Saliency, Credibility, Legitimacy and Boundaries: Linking  
Research, Assessment and Decision Making

Une information utile à la prise de décision est (notamment)

## Pertinente (saliens)

Que l'on peut traduire en expérience concrète (physique, émotionnelle, impact financier), et déclencher une action

→ Service climatique: Information spécifique, nouvelle

Une information utile à la prise de décision est (notamment)

**Crédible**

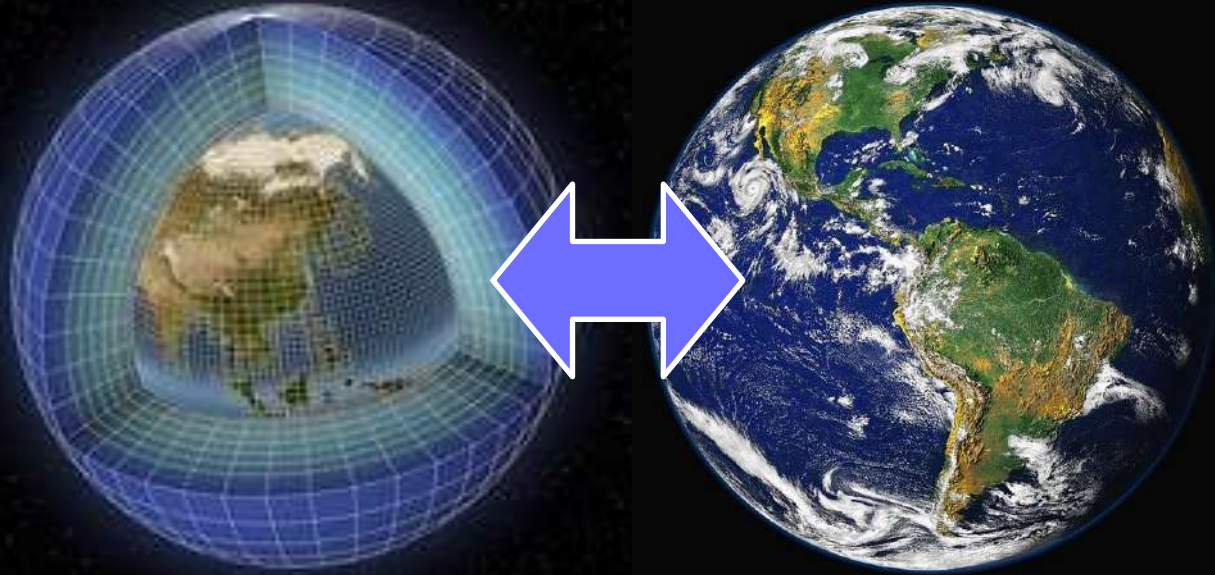
Confiance: On considère que ce qui a été annoncé  
peut effectivement se produire

# La confiance dans nos modèles/projections

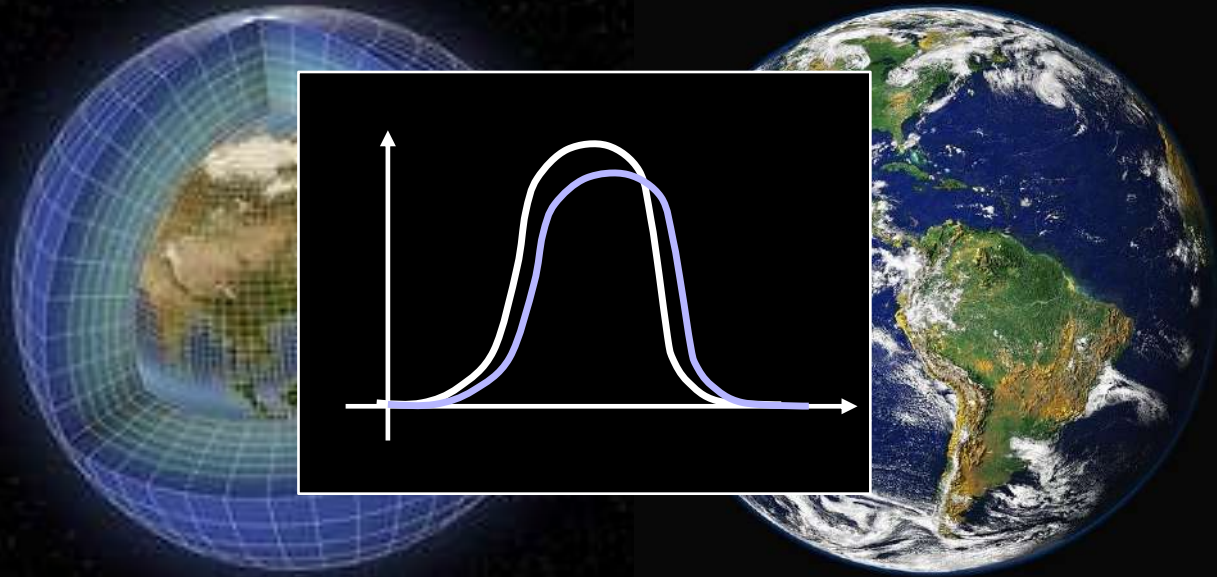
(Un sujet qu'on ne traitera pas exhaustivement aujourd'hui)

## Evaluation

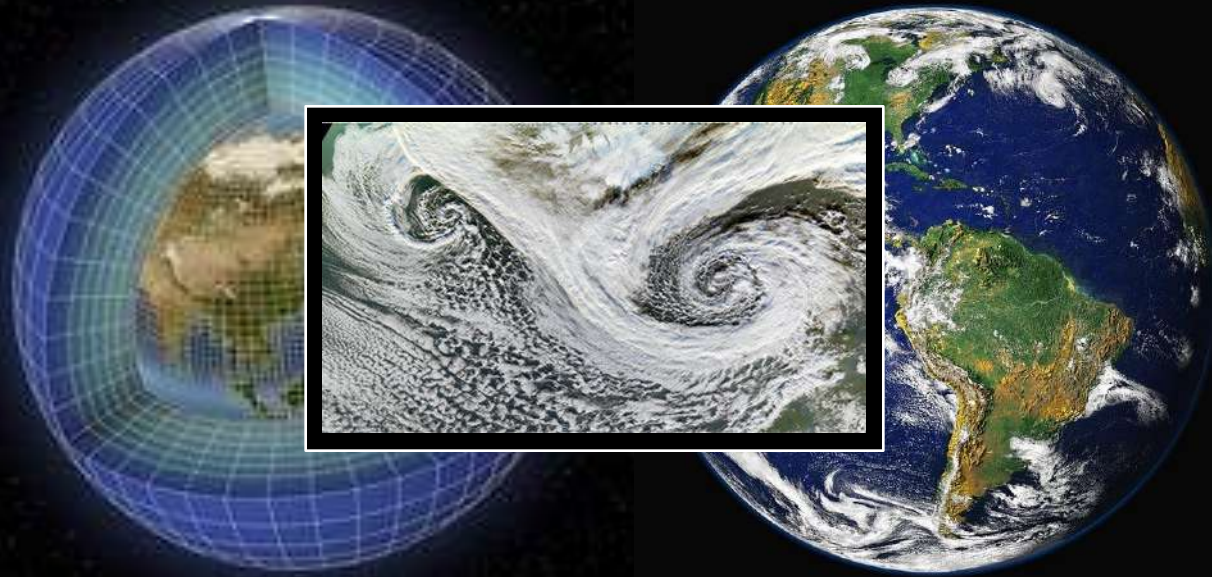
**Evaluer un modèle (pour cette session):  
comparer le climat simulé avec le climat observé**



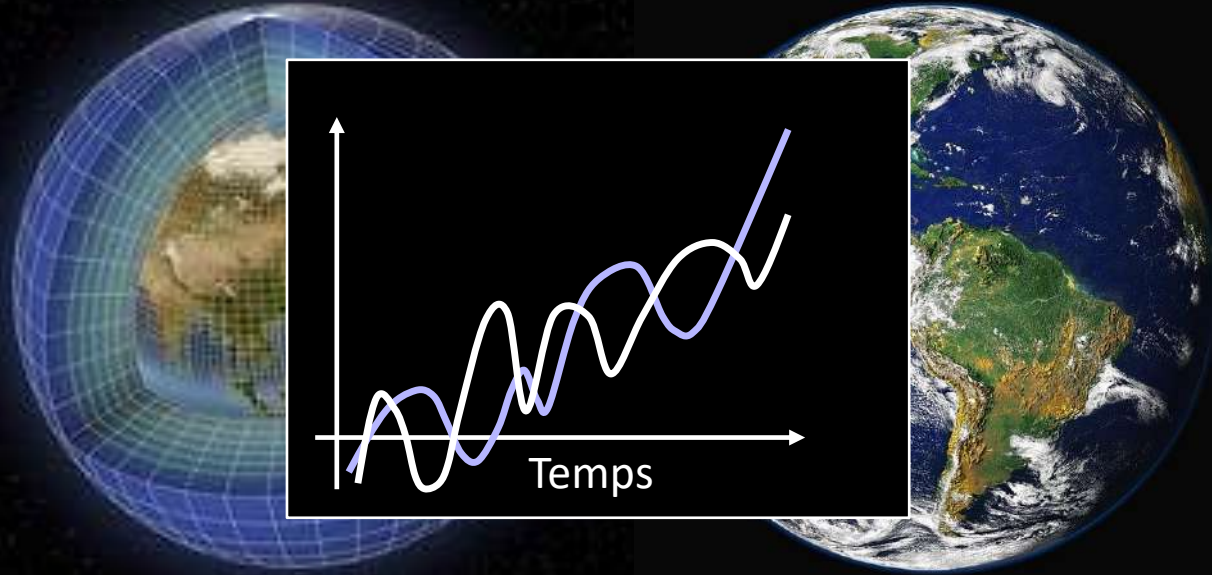
# Est-ce qu'ils se ressemblent? Quantification statistique



**Est-ce qu'ils ont les mêmes objets?**



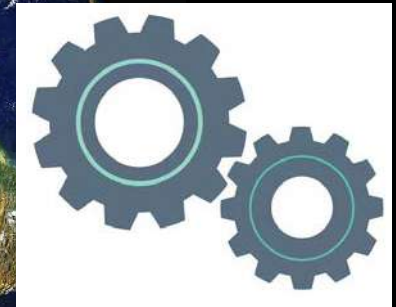
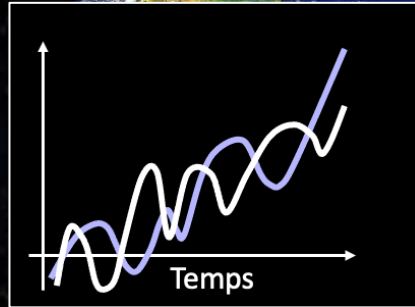
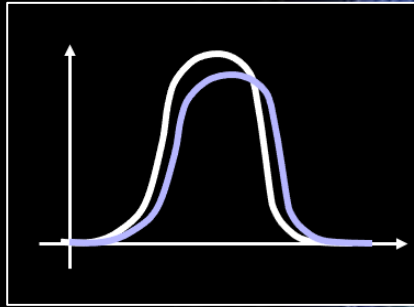
**Est-ce qu'ils évoluent de la même manière?**



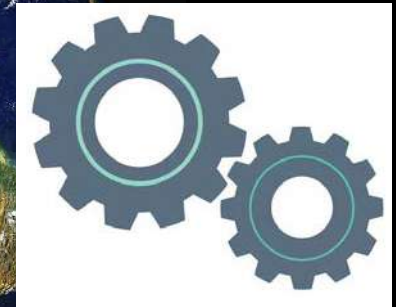
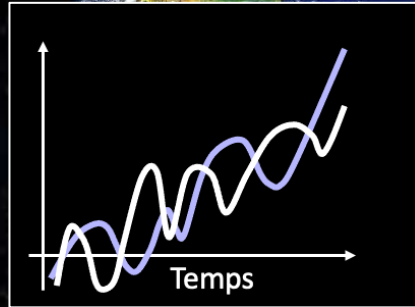
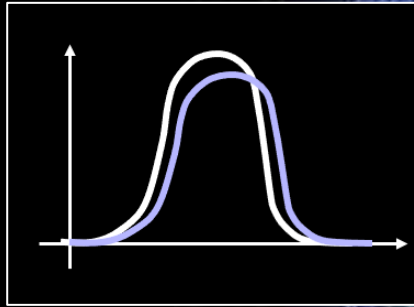
**Est-ce pour les mêmes raisons?**



**A quel point ces différences affectent  
la capacité du modèle à répondre à la question?**

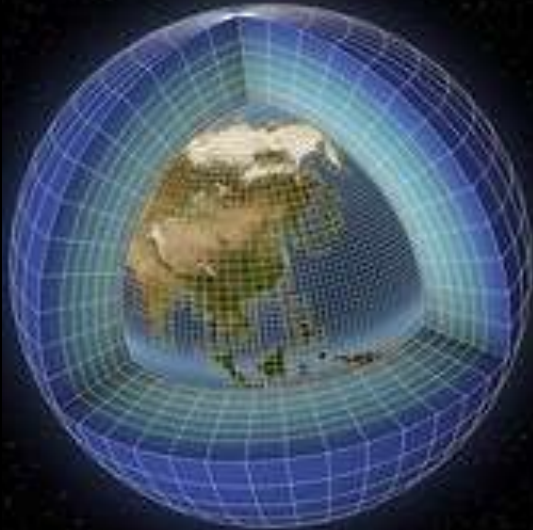


**A quel point ces différences affectent  
la capacité du modèle à répondre à la question?  
Est-ce qu'on peut les rattraper avec la Correction de biais?**



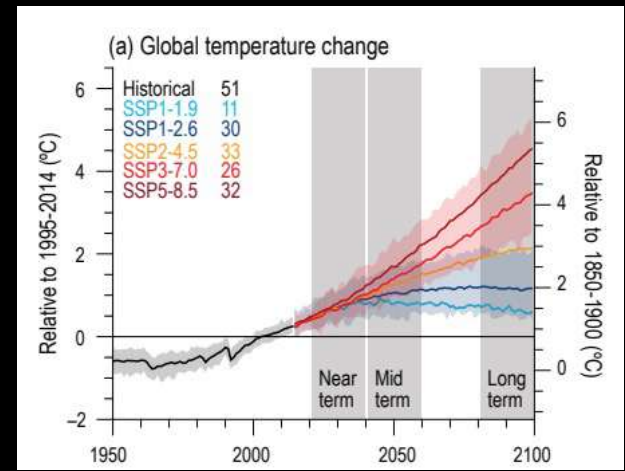
# AMONT:

développement des modèles

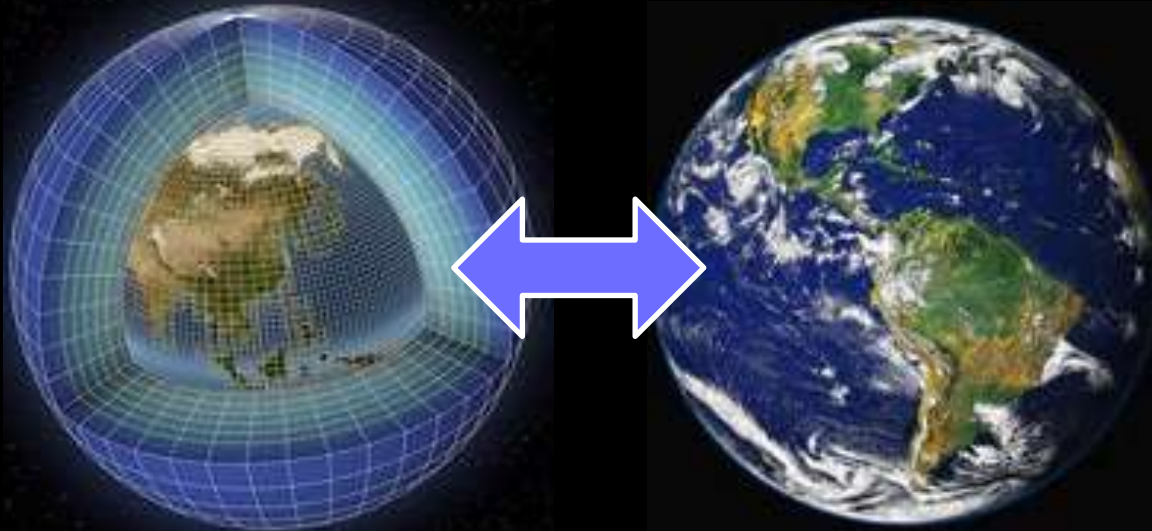


# AVAL:

analyse des simulations

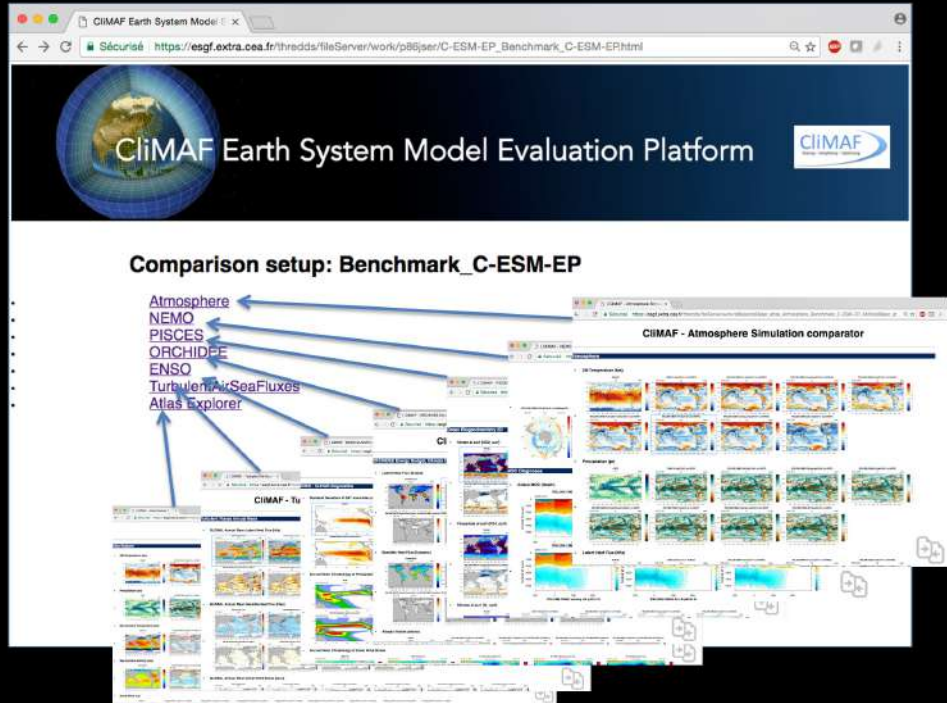


L'évaluation AMONT accompagne le développement d'un modèle, pour satisfaire le maximum d'usages: sciences académiques et applications pour la société



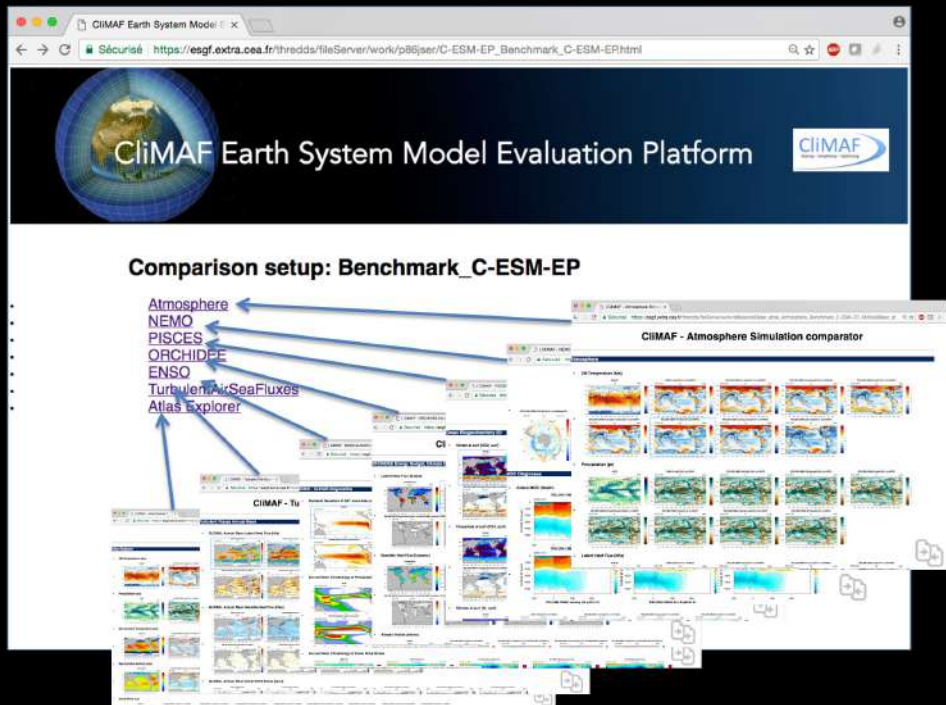
**ESGF**  
Earth System Grid Federation

Le développement des modèles repose sur un ensemble de diagnostics d'évaluation le plus exhaustif possible et « humainement gérable » pour prendre des décisions



**ESGF**  
Earth System Grid Federation

Le modèle sera bon là où on l'a regardé dans les développements, et ne sera pas obligatoirement aussi bon ailleurs

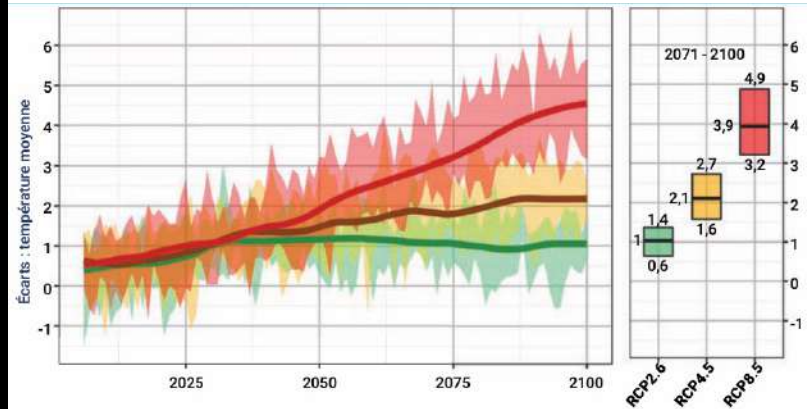


L'évaluation aval, c'est évaluer plein de modèles/simulations pour ma question scientifique, sans pouvoir modifier le modèle: Est-ce qu'ils peuvent répondre à ma question?

**ESGF**  
Earth System Grid Federation

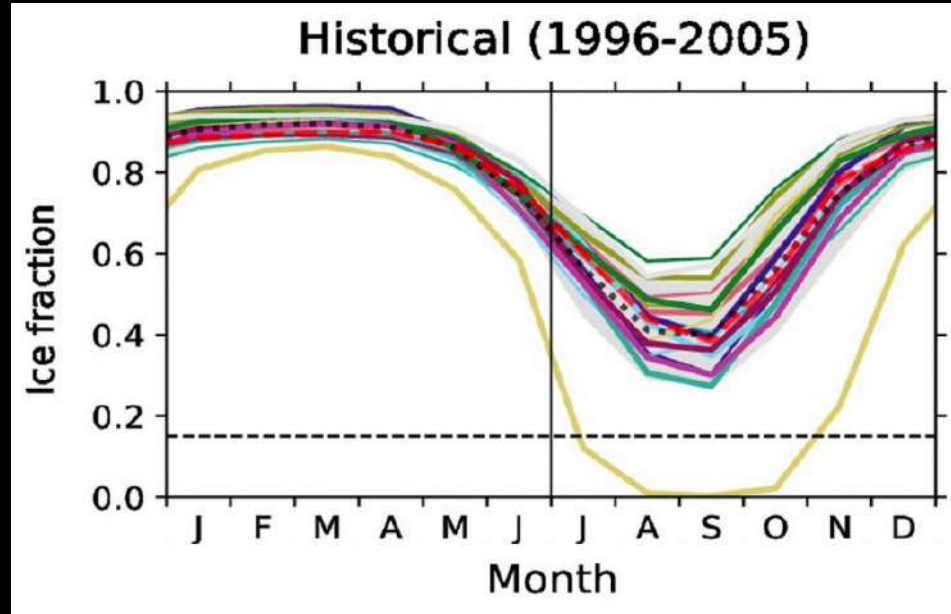


## DRIAS les futurs du climat



**Figure 12 :** Évolution de l'écart de température moyenne annuelle au cours du XXI<sup>e</sup> siècle (par rapport à la référence 1976-2005) pour les trois scénarios RCP2.6 en vert, RCP4.5 en orange et RCP8.5 en rouge. Le trait continu correspond à la médiane lissée (moyenne glissante sur 20 ans). L'enveloppe de couleur illustre l'intervalle entre les centiles 5 et 95 de la distribution de l'ensemble des simulations. La boxplot de droite représente les écarts de température sur le dernier horizon temporel (2071-2100).

Dans le cadre des services climatiques, il me faut une évaluation spécifique pour mon information pertinente (qui n'est pas forcément faite en amont)



Orr et al. 2022

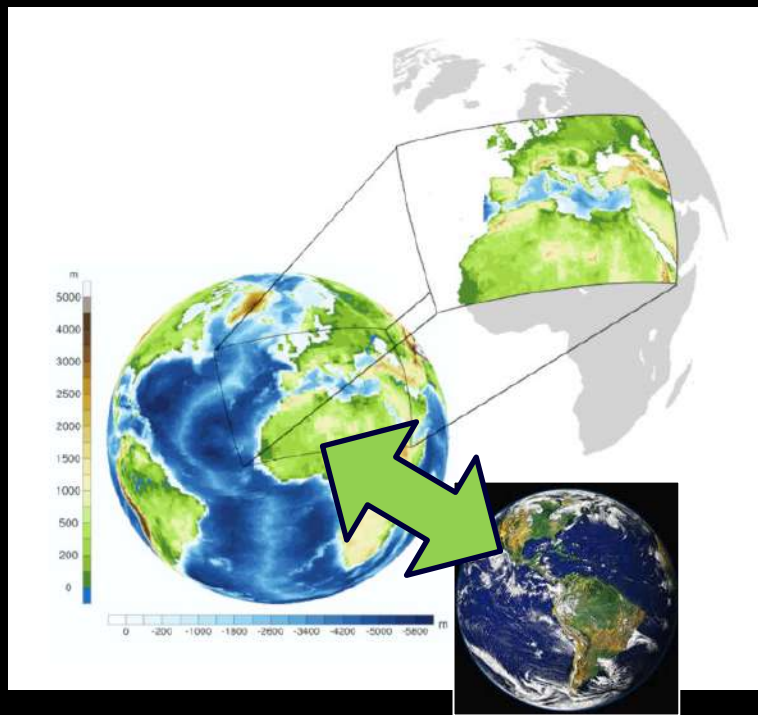
A quel point ces différences affectent la capacité du modèle à répondre à la question?

Pour EUROCORDEX – CMIP6, la sélection des simulations globales pour forcer les modèles régionaux est conditionnée (notamment) par l'évaluation de la circulation atmosphérique

Validation → GO / No GO

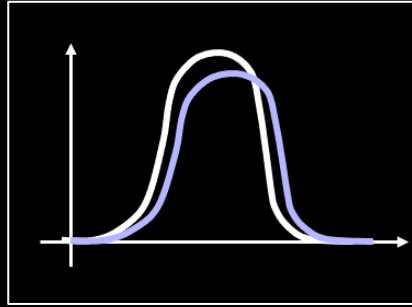
## GCM Selection and Ensemble Design: Best Practices and Recommendations from the EURO-CORDEX Community

Stefan Sobolowski<sup>a,b</sup>, Samuel Somot,<sup>c</sup> Jesus Fernandez,<sup>d</sup> Guillaume Evin,<sup>e</sup> Swen Brands,<sup>d</sup> Douglas Maraun,<sup>f</sup> Sven Kotlarski,<sup>g</sup> Martin Jury,<sup>f</sup> Rasmus E. Benestad,<sup>h</sup> Claas Teichmann,<sup>i</sup> Ole B. Christensen,<sup>j</sup> Katharina Bülow,<sup>i</sup> Erasmo Buonomo,<sup>k</sup> Eleni Katragkou,<sup>l</sup> Christian Steger,<sup>m</sup> Silje Sørland,<sup>n</sup> Grigory Nikulin,<sup>o</sup> Carol McSweeney,<sup>k</sup> Andreas Dobler,<sup>h</sup> Tamzin Palmer,<sup>k</sup> Renate Wilcke,<sup>o</sup> Julien Boé,<sup>p</sup> Lukas Brunner,<sup>q</sup> Aurélien Ribes,<sup>c</sup> Said Qasmi,<sup>c</sup> Pierre Nabat,<sup>c</sup> Florence Sevault,<sup>c</sup> and Thomas Oudar<sup>r</sup>

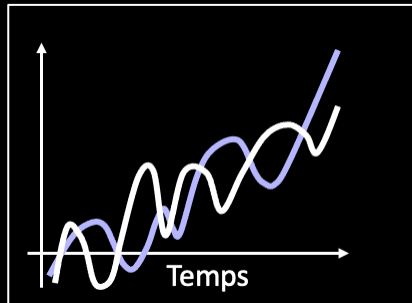


Est-ce que les simulations les plus proches des observations projettent un futur plus fiable?

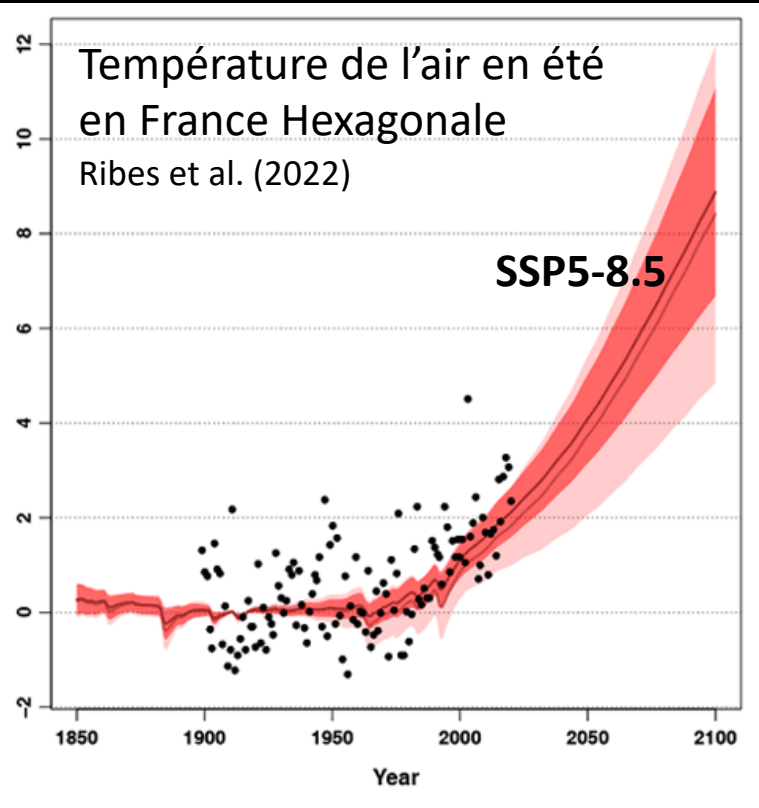
BIAIS: PAS TROP



TENDANCE: DÉJÀ PLUS



**ESGF**  
Earth System Grid Federation

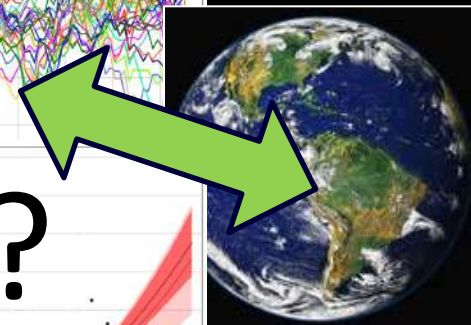
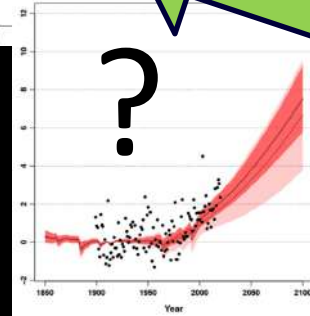
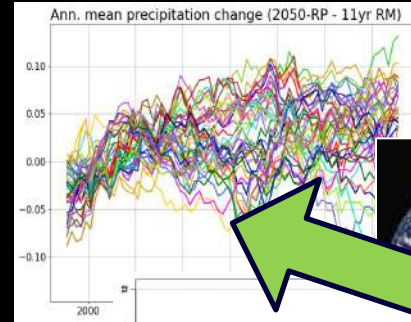
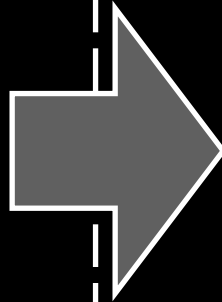


# AMONT



# AVAL

Partager sur le sujet de l'évaluation spécifique pour les services climatiques



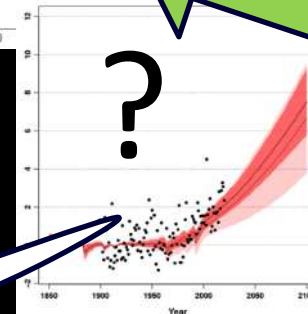
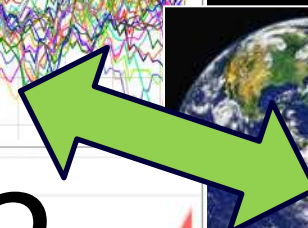
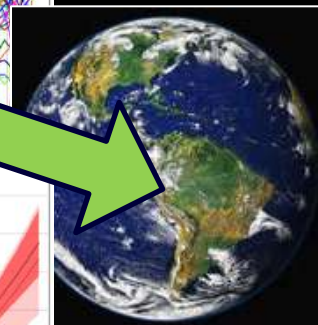
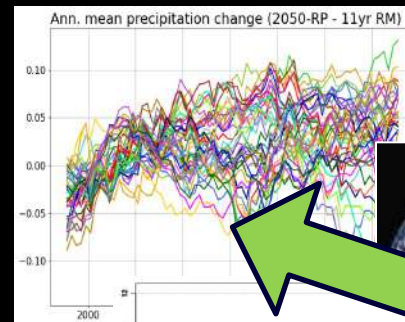
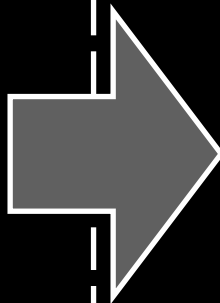
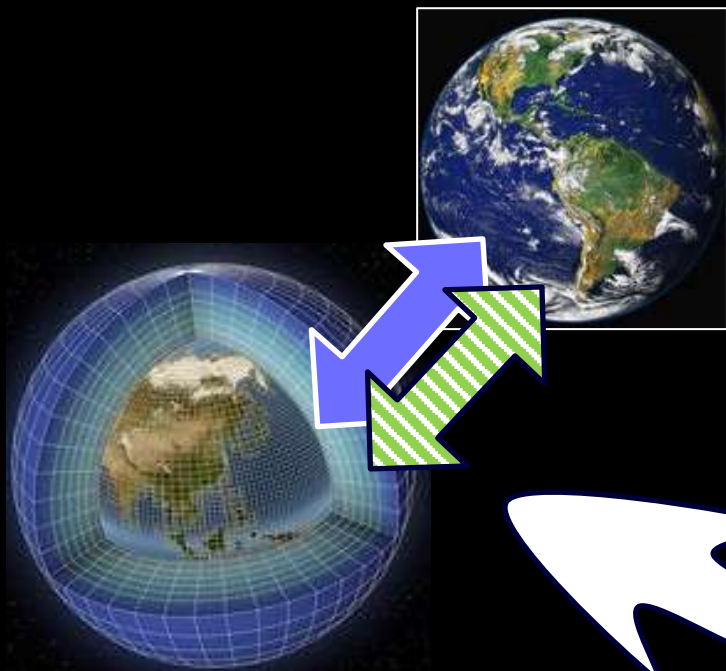
# AMONT

Inclure des diagnostics dans le développement des modèles?



# AVAL

Partager sur le sujet de l'évaluation spécifique pour les services climatiques

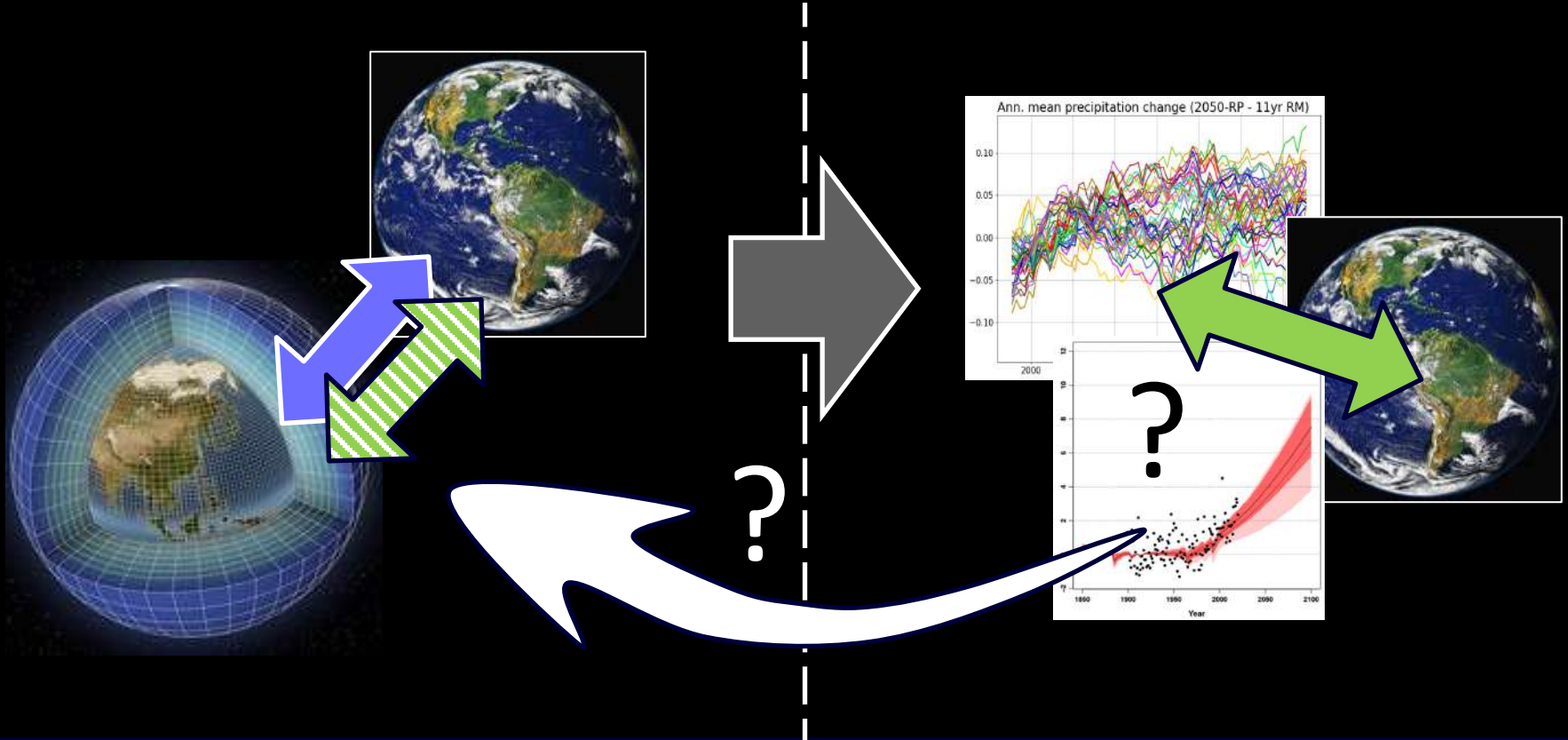


Consultation "Evaluation des modèles pour les services climatiques" (en amont de la session éponyme de l'AG TRACCS 2026)

14 répondants → 12 propositions



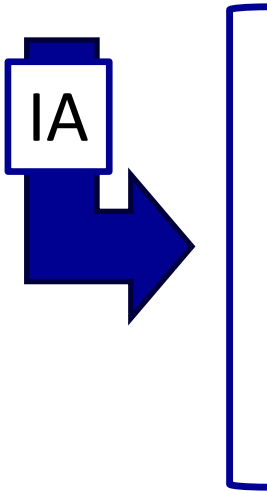
Inclure des diagnostics issus des services climatiques dans le développement des modèles comporte des enjeux **techniques**, de **moyens humains**, des questions **scientifiques** et **épistémologiques**



## Consultation "Evaluation des modèles pour les services climatiques" (en amont de la session éponyme de l'AG TRACCS 2026)

### 14 répondants:

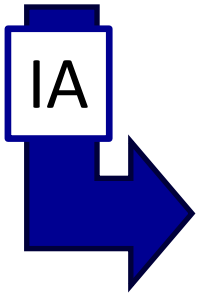
- Du Nord au Sud
- Global au régional
- De la montagne à la mer
- Annuel à horaire



## Consultation "Evaluation des modèles pour les services climatiques" (en amont de la session éponyme de l'AG TRACCS 2026)

### 14 répondants:

- Du Nord au Sud
- Global au régional
- De la montagne à la mer
- Annuel à horaire

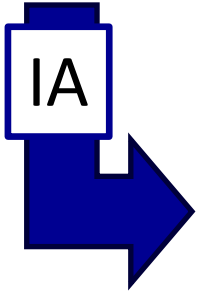


« Évaluer un modèle n'a de sens que par rapport à son usage. »  
👉 Pas de métrique universelle, pas de "meilleur modèle" absolu.

## Consultation "Evaluation des modèles pour les services climatiques" (en amont de la session éponyme de l'AG TRACCS 2026)

### 14 répondants:

- Du Nord au Sud
- Global au régional
- De la montagne à la mer
- Annuel à horaire



« Évaluer un modèle n'a de sens que par rapport à son usage. »

👉 Pas de métrique universelle, pas de "meilleur modèle" absolu.

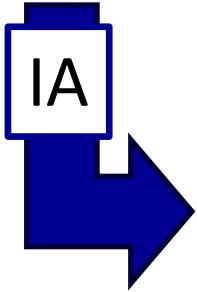
« La sélection des modèles doit être physiquement justifiée, pas statistiquement aveugle. »

👉 Transparence, diagnostics orientés processus, pas de boîte noire.

## Consultation "Evaluation des modèles pour les services climatiques" (en amont de la session éponyme de l'AG TRACCS 2026)

### 14 répondants:

- Du Nord au Sud
- Global au régional
- De la montagne à la mer
- Annuel à horaire



« Évaluer un modèle n'a de sens que par rapport à son usage. »

👉 Pas de métrique universelle, pas de "meilleur modèle" absolu.

« La sélection des modèles doit être physiquement justifiée, pas statistiquement aveugle. »

👉 Transparence, diagnostics orientés processus, pas de boîte noire.

« L'évaluation doit servir autant les utilisateurs que le développement des modèles. »

👉 Un outil structurant, pas un exercice académique isolé.

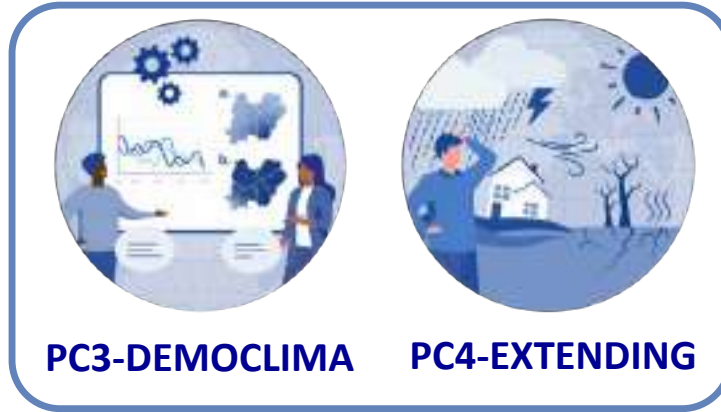
# Les PC3, PC4 et PC10 contiennent des activités en lien avec l'évaluation des modèles/simulations pour les services climatiques



PC1-DIALOG



PC2-INVEST



PC3-DEMOCLIMA

PC4-EXTENDING



PC5-COMPACT



PC6-QUINTET



PC7-IMPRESSIION-ESM



PC8-CYCL-ESM



PC9-ISCLIM



PC10-LOCALISING

# Le PC6 est concerné par l'inclusion des diagnostics d'évaluation pour les services climatiques via le tuning des modèles



**PC1-DIALOG**



**PC2-INVEST**



**PC3-DEMOCLIMA**



**PC4-EXTENDING**



**PC5-COMPACT**



**PC6-QUINTET**



**PC7-IMPRESSIION-ESM**



**PC8-CYCL-ESM**



**PC9-ISCLIM**



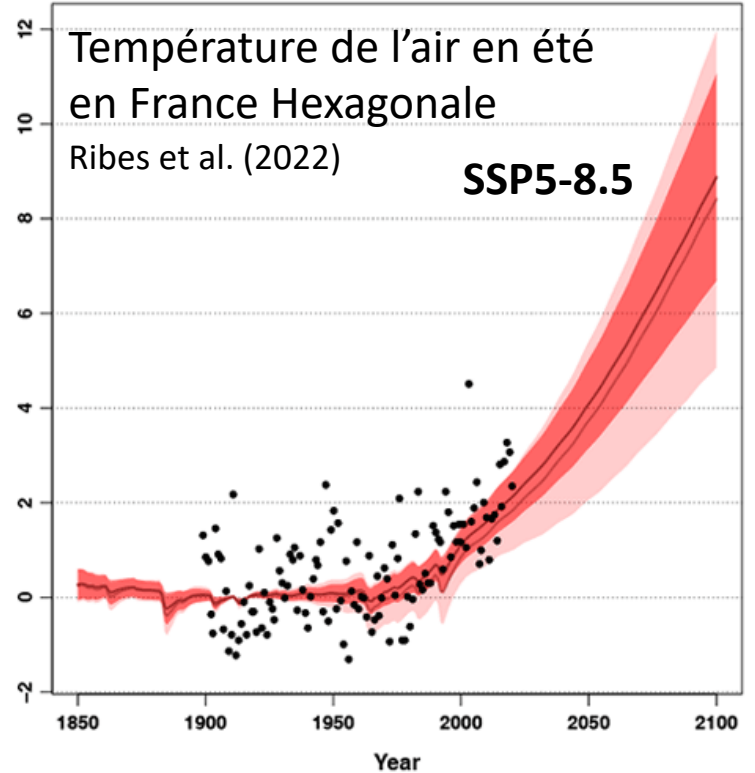
**PC10-LOCALISING**



## PC4 EXTENDING

Extreme events

**Application des contraintes observationnelles  
(Kriging for Climate Change) aux valeurs  
extrêmes de température**  
(+questions liées à la correction de biais)





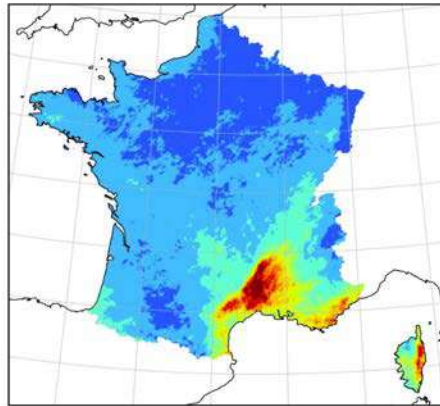
## PC10 LOCALISING

Km-scale climate information

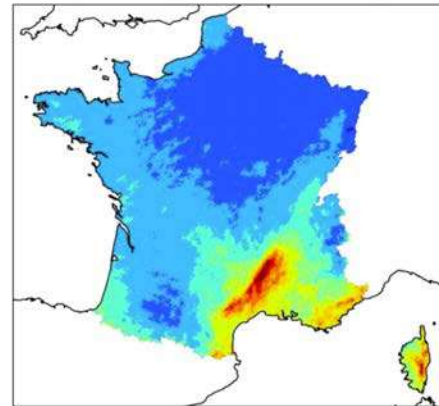
## Evaluation des Evénements extrêmes horaires de précipitation

Préparation du nouveau DRIAS, incluant des modèles km

COMEPHORE



AROME





# PC10 LOCALISING

## Km-scale climate information

## Approche objet, France hexagonale et Outre-Mer:

Outils de détection et suivi des systèmes  
précipitants

### 1. Détection ⇒ Cellules

A partir des cumuls horaires de  
précipitation (seuil=10mm/h)



### 2. Suivi ⇒ Trajectoires - - -

Conditions de recouvrement et de  
corrélation entre cellules



### 3. Caractéristiques

Pour chaque système précipitant  
détecté et suivi

- Durée
- Intensité
- Surface
- Volume
- Vitesse
- Sévérité

Un ensemble de  
cellules (c)



Extrait du  
webinaire TRACCS  
de Cécile Caillaud

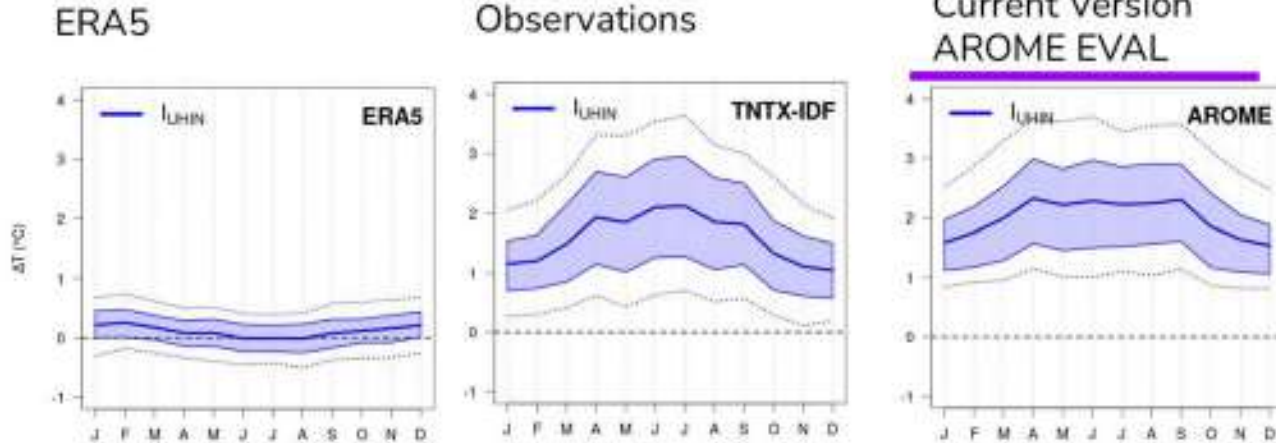


# PC10 LOCALISING

## Km-scale climate information

## Îlots de Chaleur Urbain (Aude Lemonsu)

Evaluation du cycle saisonnier de l'îlot de chaleur urbain nocturne sur Paris (différence zones urbaines – rurales)



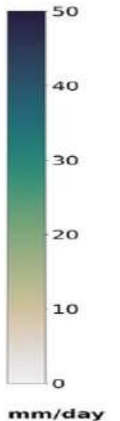
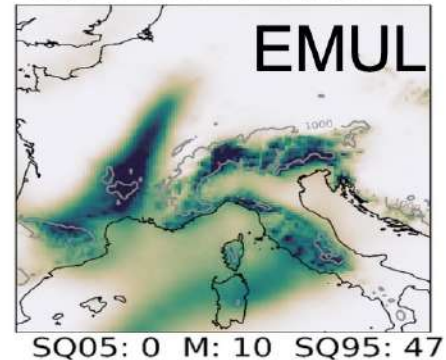
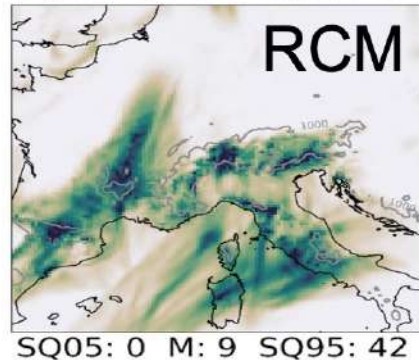
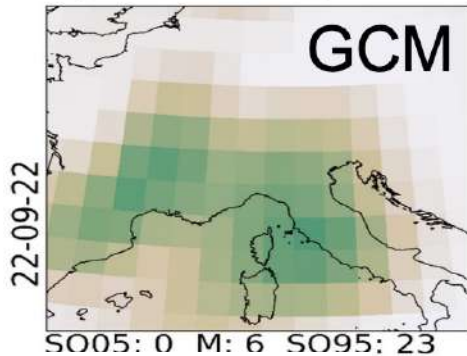


# PC10 LOCALISING

Km-scale climate information

## Evaluation des émulateurs de modèles régionaux basés sur l'IA

Événement de précipitation (mm/j, 1 jour)



*Doury et al. 2024*



## PC3 DEMOCLIMA

Démonstrateurs de services climatiques

### DEM1-DESTIAG (agriculture):

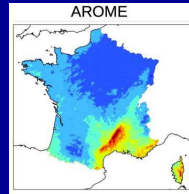
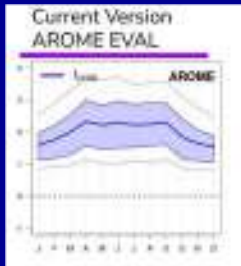
- évaluation des indicateurs agroclimatiques
- variabilité inter-annuelle du climat

Les autres: déjà fait ou pas encore à ce stade

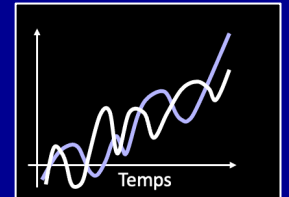
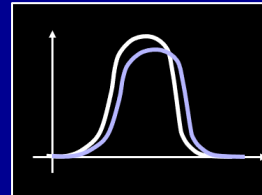
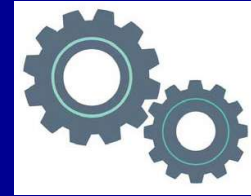


Une information utile à la prise de décision est (notamment)

**Pertinente**



**Crédible**



Cash et al. (2002): Saliency, Credibility, Legitimacy and Boundaries: Linking Research, Assessment and Decision Making

9h00-10h30

**Session « L'évaluation des modèles pour les services climatiques »**

Modération et animation : Jérôme Servonnat (LSCE-IPSL), Aurore Voldoire (CNRM), Etienne Perez (LSCE-IPSL)

**Introduction à la session et activités recensées dans TRACCS**

Jérôme Servonnat (LSCE-IPSL)

**Model evaluation of EURO-CORDEX CMIP5 for Climate CH2025 and model subselection for Switzerland**

Anna Merrifield (ETH Zurich)

**L'évaluation des modèles dans le projet Socle Outre-mer à Météo-France**

Philippe Peyrillé (CNRM)

**L'évaluation des modèles dans ARCHIVES, projet lauréat de l'appel à projets TRACCS**

Myriam Khodri (LOCEAN-IPSL), Pierre Sepulchre (LSCE-IPSL)

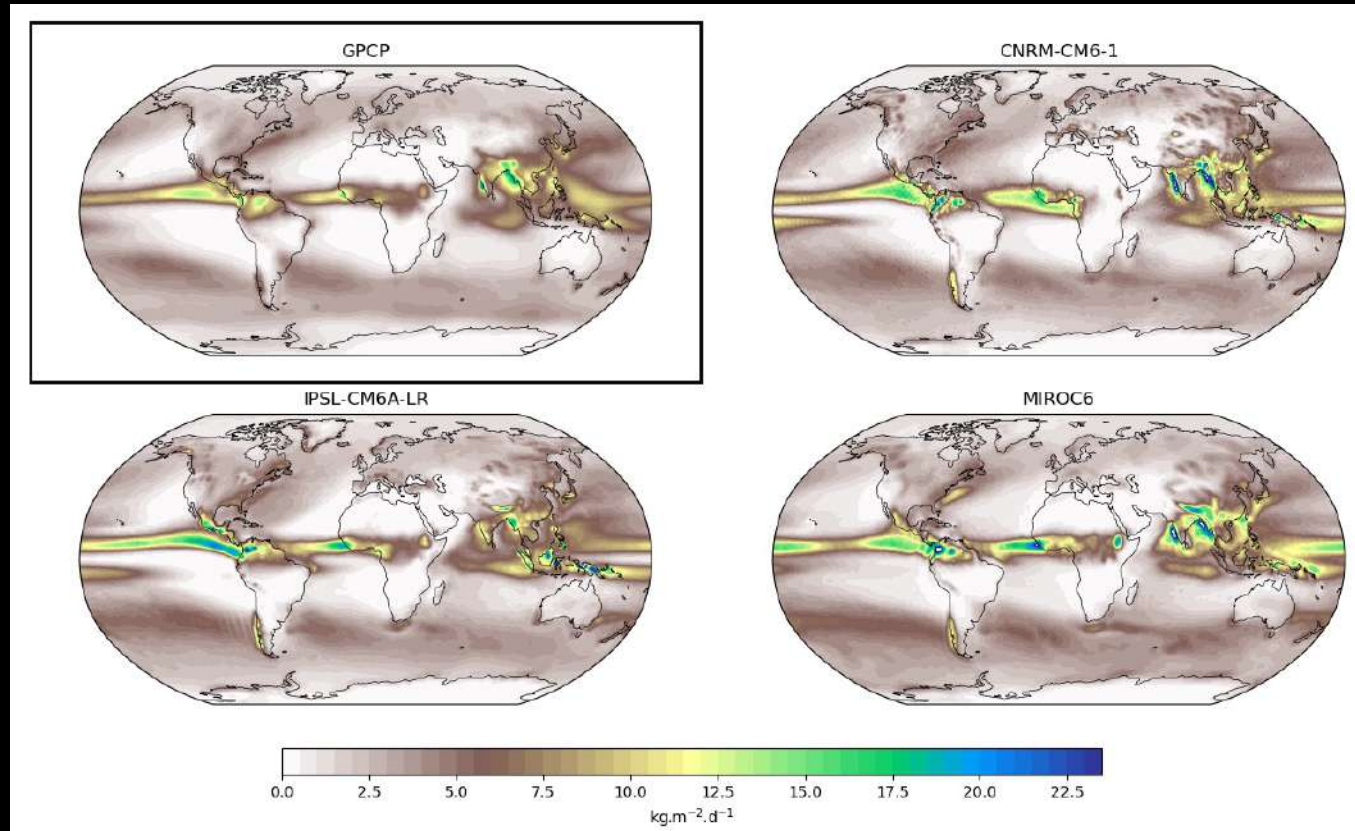
**Discussion générale**

*Amphithéâtre Farabeuf*



Les modélisateurs montrent tout le temps les défauts des modèles, n'oublions pas qu'ils font des choses formidables!

Climatologie de précipitations



## Messages principaux de la consultation

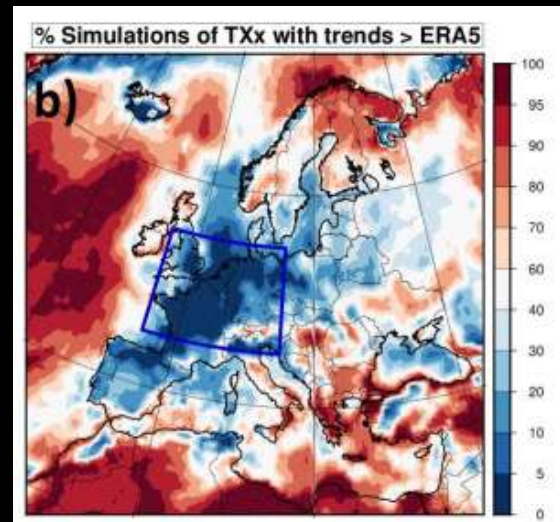
- Température, précips, niveau de la mer, neige, vent, rayonnement
- Annuel/saisonnier à horaire
- Moyennes et extrêmes
  
- Seulement 4 utilisent des méthodes pour donner plus d'importance aux modèles jugés les plus probables
- 11 utilisent pour disqualifier
- 10 pour sélectionner les meilleurs
- 12 pensent que certains de leurs diagnostics pourraient avoir vocation à remonter dans l'amont

Dans le cadre des services climatiques, il me faut une évaluation spécifique pour mon information pertinente (qui n'est pas forcément faite en amont)



## Heat extremes in Western Europe increasing faster than simulated due to atmospheric circulation trends

Robert Vautard<sup>1</sup>✉, Julien Cattiaux<sup>2</sup>, Tamara Happé<sup>3</sup>, Jitendra Singh<sup>4</sup>, Rémy Bonnet<sup>1</sup>, Christophe Cassou<sup>5</sup>, Dim Coumou<sup>1,3,6</sup>, Fabio D'Andrea<sup>7</sup>, Davide Faranda<sup>8</sup>, Erich Fischer<sup>4</sup>, Aurélien Ribes<sup>2</sup>, Sebastian Sippel<sup>4</sup> & Pascal Yiou<sup>8</sup>



Des efforts à l'international pour un accès à une évaluation plus homogène pour tous les modèles, mais un risque de standardisation



## Rapid Evaluation Framework for the CMIP7 Assessment Fast Track

Forrest M. Hoffman [✉](#), Birgit Hassler [✉](#), Ranjini Swaminathan [✉](#), Jared Lewis, Bouwe Andela, Nathaniel Collier, Dóra Hegedűs, Jiwoo Lee, Charlotte Pascoe, Mika Pflüger, Martina Stockhause, Paul Ullrich, Min Xu, Lisa Bock, Felicity Chun, Bettina K. Gier, Douglas I. Kelley, Axel Lauer, Julien Lenhardt, Manuel Schlund, Mohanan G. Sreeush, Katja Weigel, Ed Blockley, Rebecca Beadling, Romain Beucher, Demiso D. Dugassa, Valerio Lembo, Jianhua Lu, Swen Brands, Jerry Tjiputra, Elizaveta Malinina, Brian Mederios, Enrico Scoccimarro, Jeremy Walton, Philip Kershaw, André L. Marquez, Malcolm J. Roberts, Eleanor O'Rourke, Elisabeth Dingley, Briony Turner, Helene Hewitt, and John P. Dunne