

Évolution des technologies des données et de calcul: opportunités et défis pour la simulation et l'analyse du système Terre

Face au changement climatique, le champ des possibles
Colloque Académie des Sciences
Institut de France, Paris

V. Balaji

Laboratoire des Sciences du Climat et de l'Environnement (LSCE)
Institut Pierre-Simon Laplace (IPSL)
NOAA/Geophysical Fluid Dynamics Laboratory (GFDL) and Princeton University

28 janvier 2020



Contours de la présentation

- 1 La science et les données du climat: synthèse des observations et simulations
 - Observations du climat terrestre, du passé lointain jusqu'à présent
 - Simulations de la planète Terre: données du futur et des terres fictives
- 2 Évolution des technologies pour la simulation et les données
 - A Vast Machine
 - Explosion dans le volume des données, démocratisation d'accès aux données
- 3 Machine Learning: nouveau enjeu dans la technologie
 - Fin de l'ère von Neumann?
 - Transition du calcul vers l'apprentissage (ML)
 - De la simulation vers l'émulation
- 4 Évolution des technologies de la simulation et des données: idées, enjeux, défis

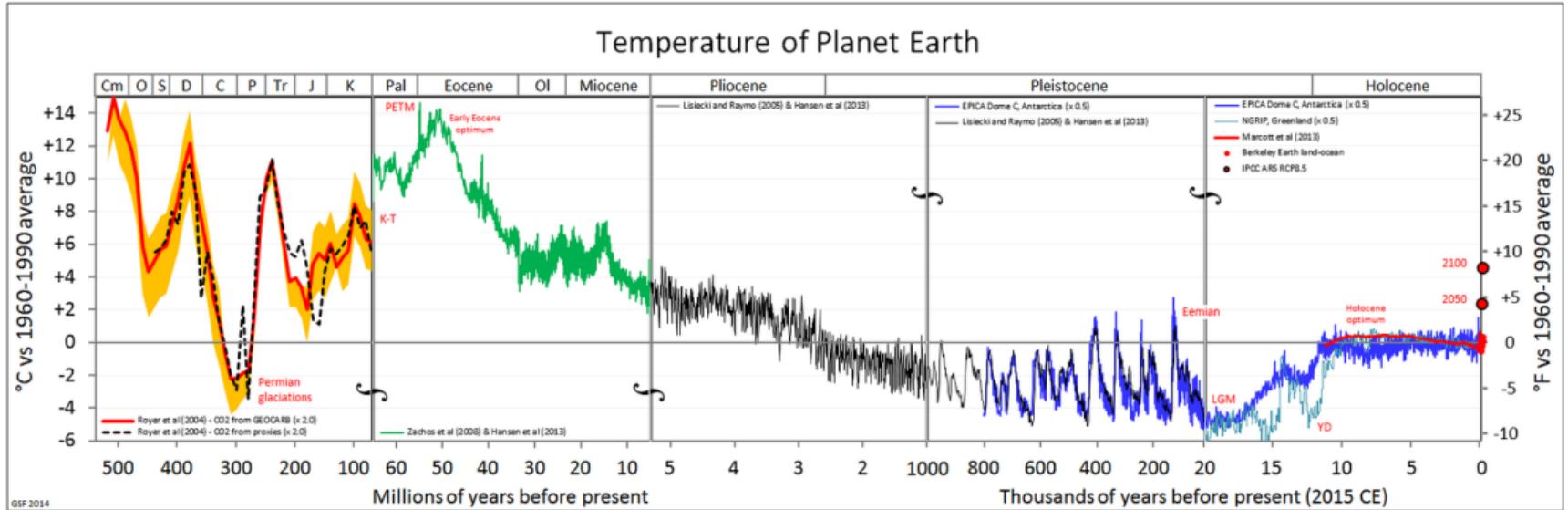


Contours de la présentation

- 1 La science et les données du climat: synthèse des observations et simulations
 - Observations du climat terrestre, du passé lointain jusqu'à présent
 - Simulations de la planète Terre: données du futur et des terres fictives
- 2 Évolution des technologies pour la simulation et les données
 - A Vast Machine
 - Explosion dans le volume des données, démocratisation d'accès aux données
- 3 Machine Learning: nouveau enjeu dans la technologie
 - Fin de l'ère von Neumann?
 - Transition du calcul vers l'apprentissage (ML)
 - De la simulation vers l'émulation
- 4 Évolution des technologies de la simulation et des données: idées, enjeux, défis



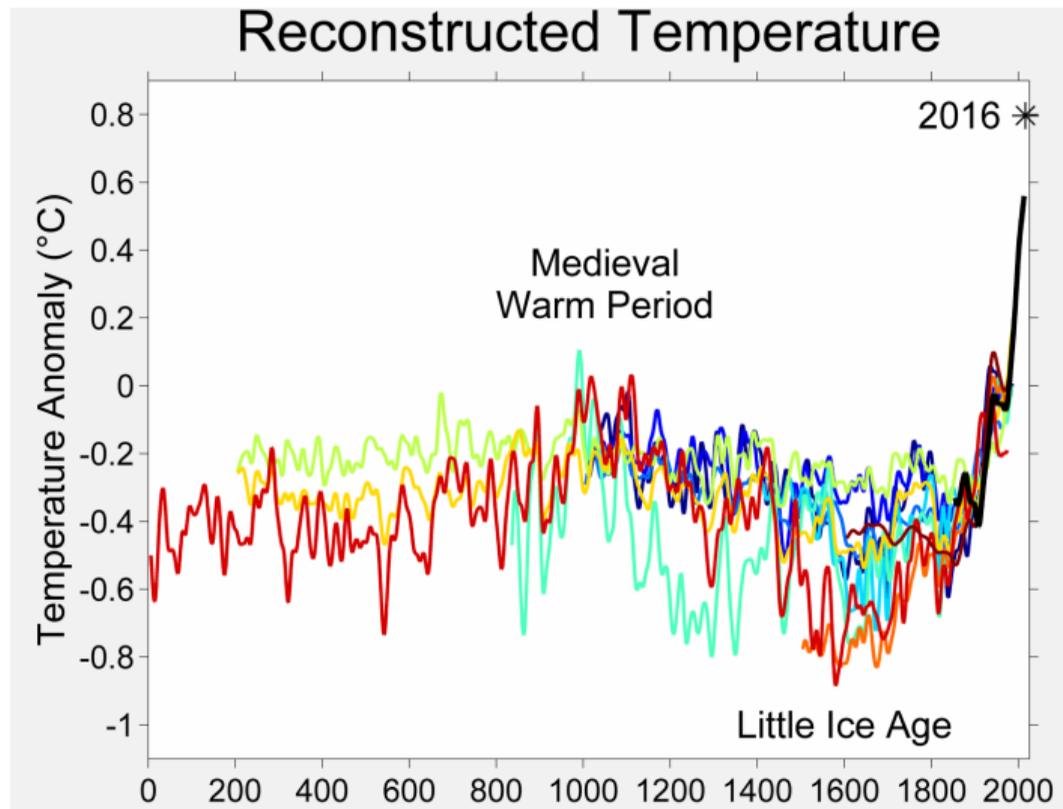
Histoire de la température terrestre



Du Wikipédia.



Histoire de la température terrestre de notre ère



- Reconstructions à partir des cernes d'arbres, carottes de glace, ...
- Le **Mediæval warm period** et le **Little Ice Age** ne sont peut-être que des signaux régionaux
- Le réchauffage du présent porte une **signature globale**
- De **Wikipédia**



Réponse atmosphérique au doublement CO_2

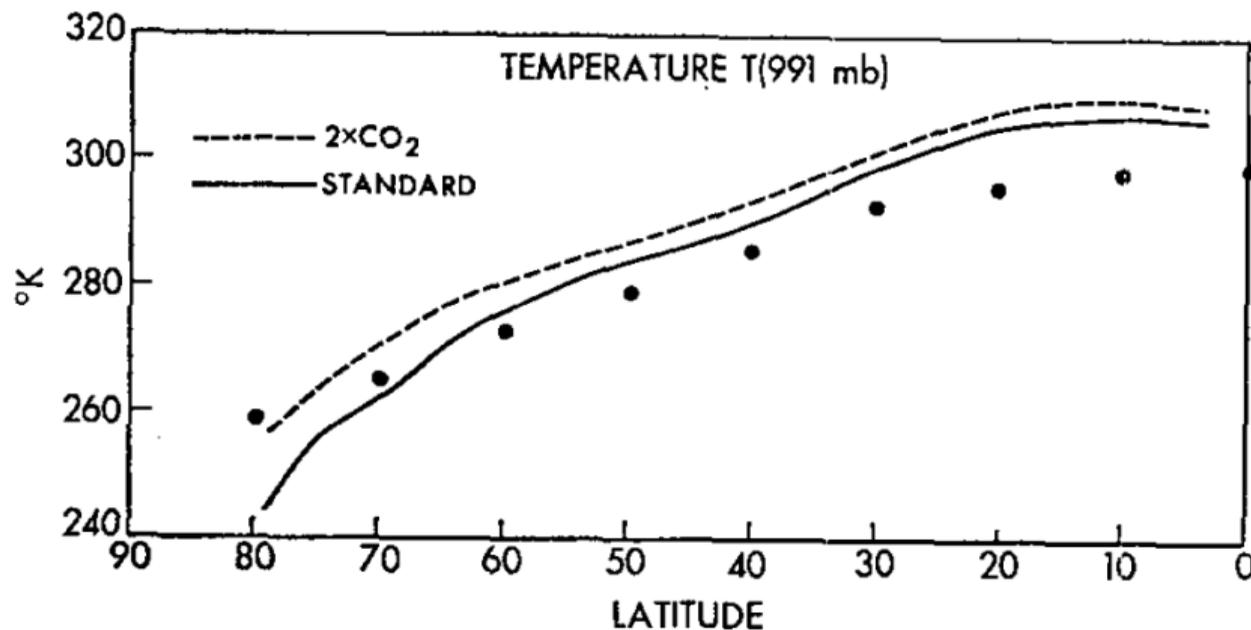


Fig 5 de [Manabe and Wetherald \(1975\)](#), réponse équilibrée à $2 \times \text{CO}_2$.

Le Rapport Charney (1979)

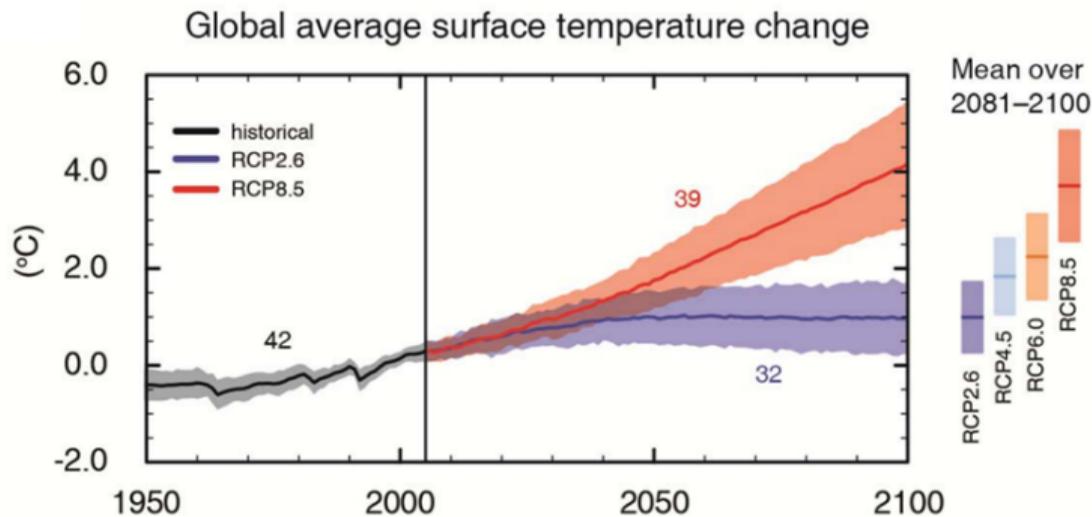
“Carbon dioxide and climate: A Scientific Assessment.”

- Précurseur aux rapports GIEC.
- Basé sur 5 simulations: 3 de Manabe (GFDL), 2 de Hansen (GISS).
- Conclusions:
 - Effets radiatifs directs dûs au doublement CO₂: $\sim 4 \text{ W/m}^2$
 - Rétroactions: vapeur d'eau (Clausius-Clapeyron), “ice albedo feedback”.
 - Rôle des nuages: “How important the cloud effects are, is, however, an extremely difficult question to answer. **The cloud distribution is a property of the entire climate system**, in which many other feedbacks are involved.”
 - “We believe, therefore, that the equilibrium surface warming will be in the range of **1.5-4.5°C**, with the most probable value near 3°C.”

Rapport Charney vu du présent: [Bony et al \(2013\)](#).

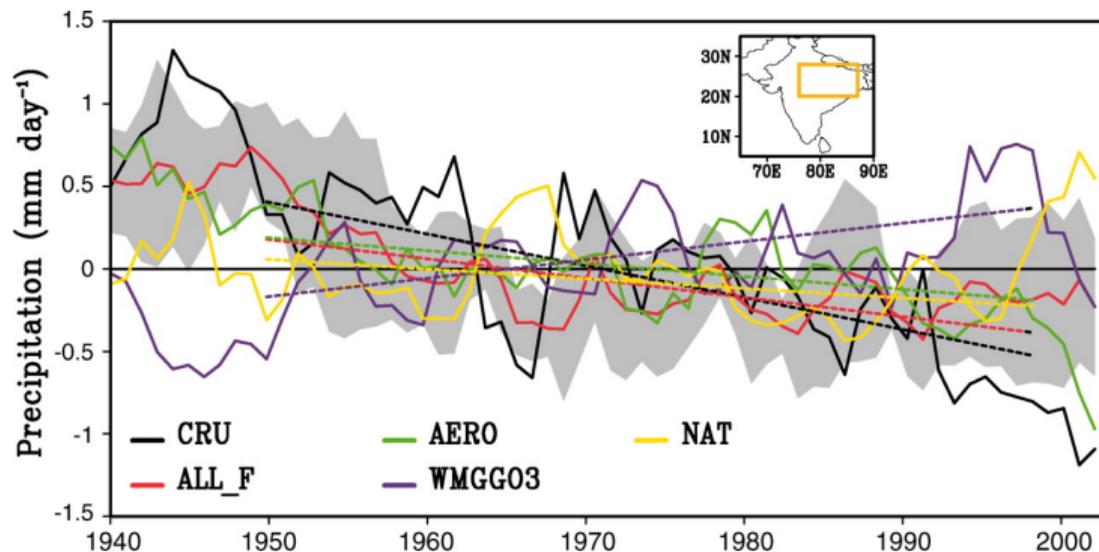


Du Charney Report à aujourd'hui...



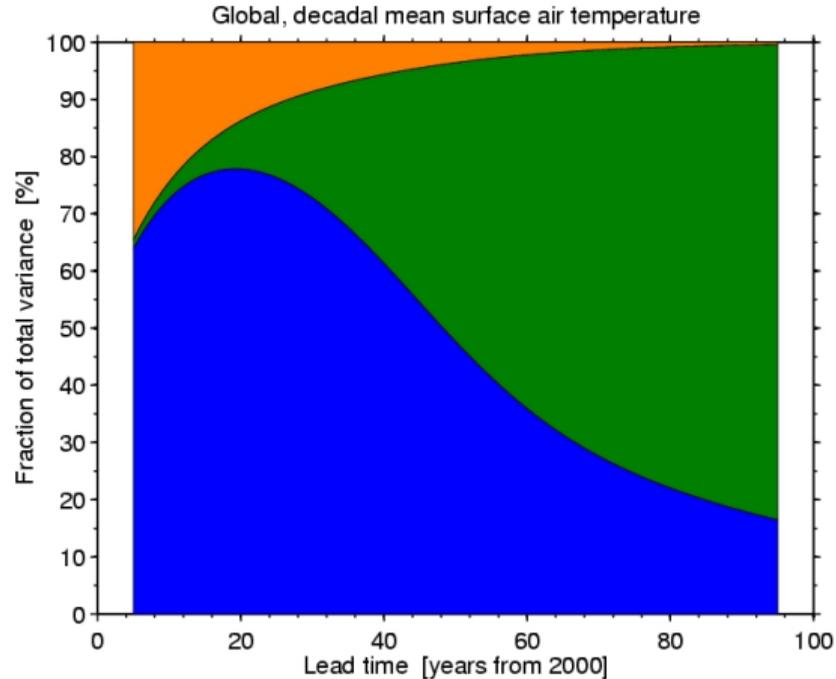
- Figure SPM.7 du rapport GIEC AR5. Température moyenne globale réduite par l'imposition des contrôles sur l'émission du CO₂.
- Plusieurs nouveaux modèles dans CMIP6 et GIEC AR6.
- Nouveaux joueurs, Inde, Chine, Brésil, Afrique du Sud, Corée, ...

Attribution: données issues des terres fictives



- Rétroaction aérosol-nuages a l'effet d'affaiblir la mousson indienne [Bollasina et al., Science 2011](#).
- L'attribution des événements individuels par la même voie.

Incertitudes: chaotique, épistémique, humaine



Hawkins and Sutton (2009).



Contours de la présentation

- 1 La science et les données du climat: synthèse des observations et simulations
 - Observations du climat terrestre, du passé lointain jusqu'à présent
 - Simulations de la planète Terre: données du futur et des terres fictives
- 2 Évolution des technologies pour la simulation et les données
 - A Vast Machine
 - Explosion dans le volume des données, démocratisation d'accès aux données
- 3 Machine Learning: nouveau enjeu dans la technologie
 - Fin de l'ère von Neumann?
 - Transition du calcul vers l'apprentissage (ML)
 - De la simulation vers l'émulation
- 4 Évolution des technologies de la simulation et des données: idées, enjeux, défis



A Vast Machine (Paul Edwards)



Model output

Observations and reanalyses



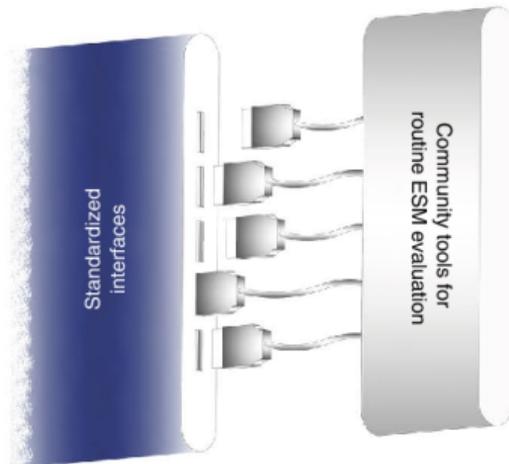
Processing capability



Data archive

Analysis computing environment
integrated with the ESGF

Well-established analysis
Sharing of diagnostic codes
Guidance and support from the CMIP panel,
WGNE/WGCM climate model diagnostics and
metrics panel, and CMIP6-Endorsed MIPs

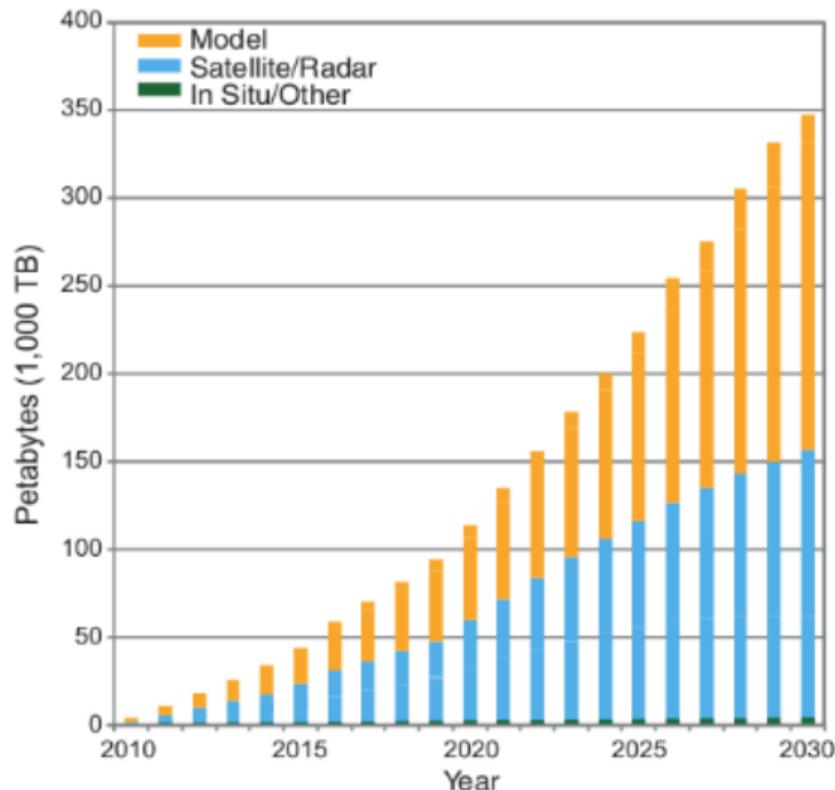


Visualization and documentation
of evaluation results
Record of provenance
Scientific interpretation
Additional in-depth analysis

State evaluation of ECVs
(climatology, trends and so on)
Process and phenomena evaluation
Link to projections
(MMM analysis and emergent constraints)
Performance metrics



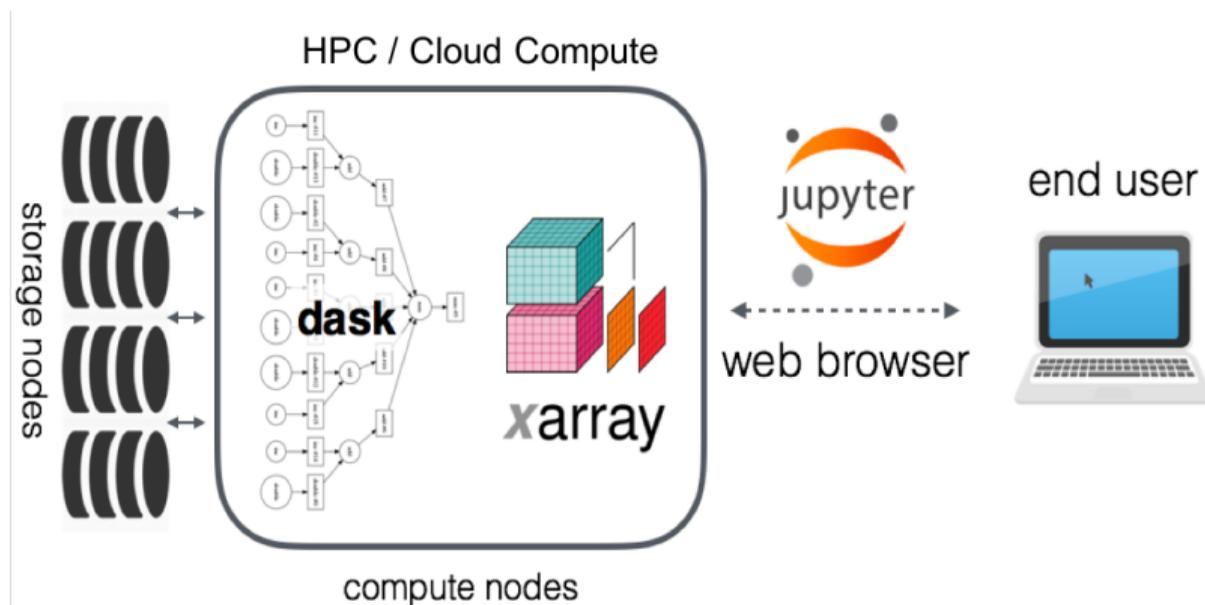
Explosions des données des observations et simulations



- Données issues des simulations dépassent en ampleur celles des observations!
- [Overpeck et al., Science \(2011\)](#) ont prévu une croissance de facteur 100
- [Balaji et al \(2016\)](#) prévoient 18-24 PB for CMIP6



Pangeo: amener l'analyse chez les données



- Si on n'est capable d'apporter les données chez nous, peut-on apporter nos analyses chez les données?
- **Pangeo**, plateforme "big data" pour les géosciences.

Contours de la présentation

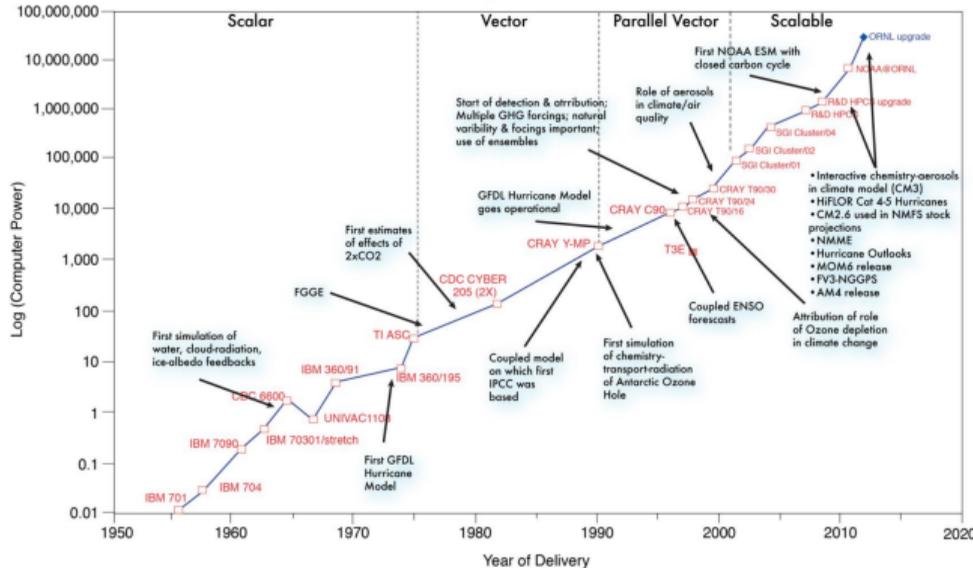
- 1 La science et les données du climat: synthèse des observations et simulations
 - Observations du climat terrestre, du passé lointain jusqu'à présent
 - Simulations de la planète Terre: données du futur et des terres fictives
- 2 Évolution des technologies pour la simulation et les données
 - A Vast Machine
 - Explosion dans le volume des données, démocratisation d'accès aux données
- 3 **Machine Learning: nouveau enjeu dans la technologie**
 - **Fin de l'ère von Neumann?**
 - **Transition du calcul vers l'apprentissage (ML)**
 - **De la simulation vers l'émulation**
- 4 Évolution des technologies de la simulation et des données: idées, enjeux, défis



Histoire du calcul à GFDL

HISTORY OF GFDL COMPUTING

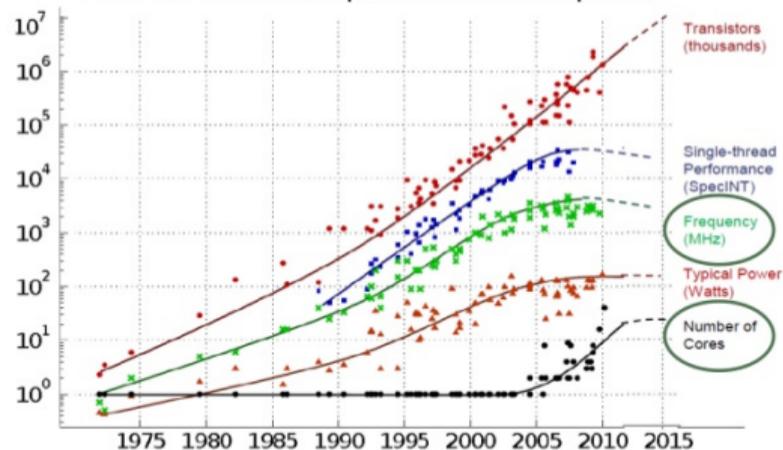
Growth of Computational Power with Time



- John von Neumann a fait naître dans les 1950s des programmes pour la prévision numérique météorologique et pour le climat (“infinite forecast”).
- Courtesy V. Ramaswamy, NOAA/GFDL.

Moore's Law et Dennard scaling

Power and Heat Problems Led to Multiple Cores and Prevent Further Improvements in Speed



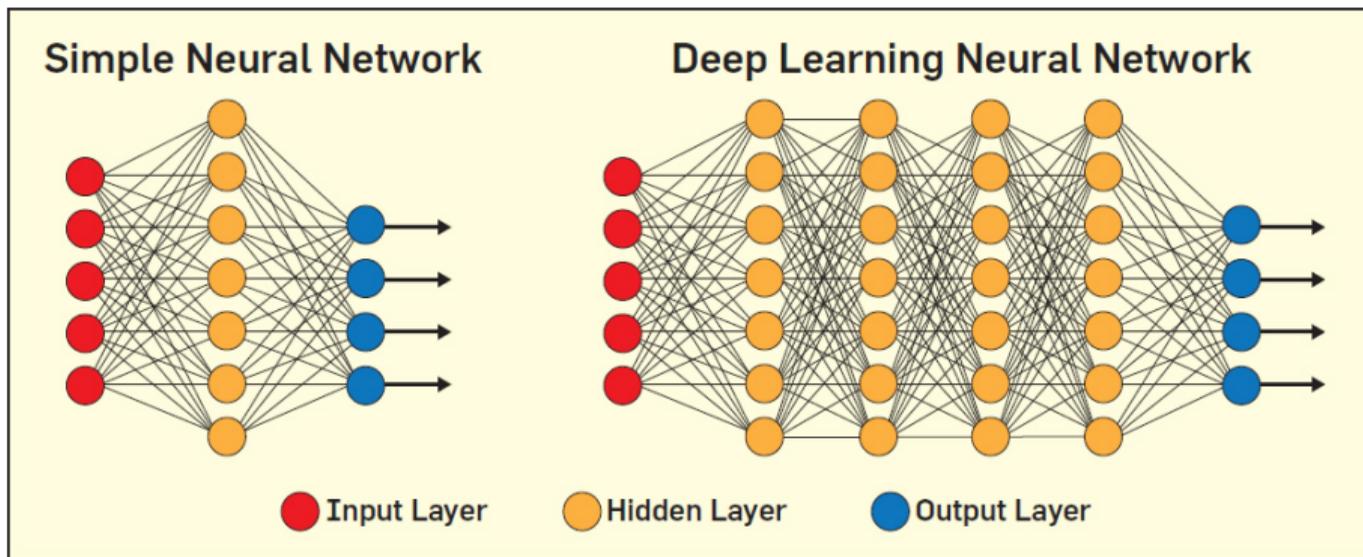
Original data collected and plotted by M. Horowitz, F. Labonte, O. Shacham, K. Olukotun, L. Hammond and C. Batten
Dotted line extrapolations by C. Moore

Source: Chuck Moore, Data Processing in Exascale-Class Systems, April 27, 2011: Salishan Conference on High Speed Computing

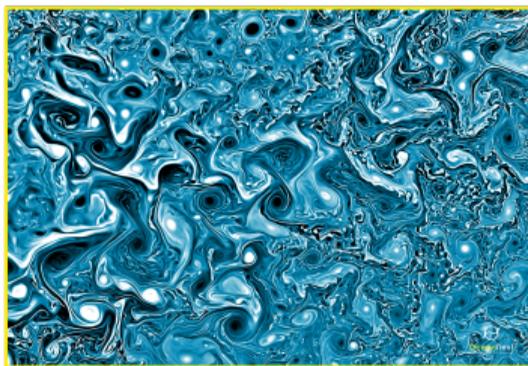
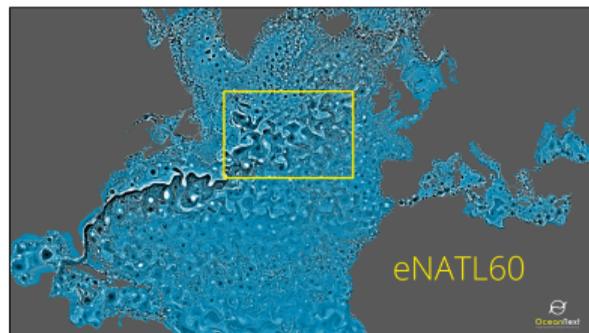
- L'arithmétique numérique n'accélère plus. mais on peut en faire davantage en parallèle.
- Moore 2011: *Data processing in exascale-class systems.*



Deep Learning



De [Edwards \(2018\)](#), ACM.



- **MOPGA Projet Hermès: High Resolution Modeling of the Earth System, 2018-2022.**
- **Fin du loi Moore** signale que la résolution de nos modèles ne peut pas continuer à croître à l'infini.
- **Apprendre** les comportements des nuages et de la turbulence à très fine échelle pour l'appliquer à basse résolution.
- Simulations eNATL60 de Julien le Sommer (Uni Grenoble), modèle NEMO.

Contours de la présentation

- 1 La science et les données du climat: synthèse des observations et simulations
 - Observations du climat terrestre, du passé lointain jusqu'à présent
 - Simulations de la planète Terre: données du futur et des terres fictives
- 2 Évolution des technologies pour la simulation et les données
 - A Vast Machine
 - Explosion dans le volume des données, démocratisation d'accès aux données
- 3 Machine Learning: nouveau enjeu dans la technologie
 - Fin de l'ère von Neumann?
 - Transition du calcul vers l'apprentissage (ML)
 - De la simulation vers l'émulation
- 4 Évolution des technologies de la simulation et des données: idées, enjeux, défis



Récapitulatif

- Les **données** et le **calcul** sont inséparables: notre connaissance du Système Terre repose sur une base de **simulations** du climat passé, climat futur, d'autres climats possibles.
- **Démocratisation** et internationalisation de la science du climat
- Le calcul évolue: du mode impérative vers l'apprentissage.
- ... ce que nous mène de la **simulation** vers **l'émulation**
- **"Interpretable AI"**: les méthodes dites **"model-free"** posent un problème pour la connaissance!
- **Coût en carbon** des données et calcul! – le projet CPMIP ([Balaji et al 2017](#)) proposent de le mesurer pour CMIP6.
- **Grand mercis**: Alistair Adcroft, Tom Bolton, Noah Brenowitz, Chris Bretherton, Hannah Christensen, Julie Deshayes, Peter Düben, Marie-Alice Foujols, Frédéric Hourdin, Julien le Sommer, Tim Palmer, Tapio Schneider, Anna Sommer (MOPGA postdoc), Laure Zanna, maintes autres!
- ... et au LSCE et IPSL, et le programme MOPGA.

