La réflexion internationale :
Programme mondial de la météorologie
Programme mondial de recherche sur le climat

Pascale Braconnot

Journées CLIMERI-France 3-4 février 2021

La HPC and AI task team de WMO

Task Teams "Exascale-Big Data and Artificial Intelligence and/or Machine Learning" lancée à l'initiative du "Research board ":,

Mark Govett	• NOAA
Bubacar Bah	• African Institute for Mathematical Sciences
Peter Bauer	• ECMWF
Dominique Berod	• WMO
Veronique Bouchet	• ECCC
Wenchao Cao	• WMO
Susanna Corti	• Institute of Atmospheric Science and Climate
Chris Davis	• NCAR
Tim Graham	• UK Met Office
Yuki Honda	• WMO
Adrian Hines	• UK Met Office
Junishi Ishida	• Japan Meteorological Agency
Bryan Lawrence	• University of Reading
Kris Rowe	• Argonne Leadership Computing Facility
Martin Schultz	• Jülich Supercomputing Centre
Martin Visbeck	• Helmholtz Centre for Ocean Research
Keith Williams	• UK Met Office
Yihong Duan	• China Meteorological Administration

WMO OMM

The Research Board convenes, organizes and motivates, in an inclusive, flexible and forward looking way, the scientific and technical community around the three research objectives of the WMO strategic plan and facilitate the integration across the WMO sponsored and **co-sponsored programmes (ex WCRP)**:
•Advance scientific knowledge of the Earth system
•Enhance the science-for-service value chain ensuring scientific and technological advances improve predictive capabilities
•Advance policy-relevant science

2 notes

- HPC and data
- IA

Auteurs = sous groupes de TT

Etat des lieux dans les pays partenaires
Analyse sous différents angles
Recommandations

Issu d'une présentation de Kris Row, Argonne Leadership Facility, et Mark Govett, NOAA Global System Laboratory

Principaux points pour différents défis identifiés

Challenge #1
Cost

Hardware

- U.S exascale systems in the range $300m-$600m
- Electricity costs for 30 MW are more than $12m USD
- Facilities and maintenance costs are also significant

Software

- Development and maintenance costs are often overlooked
- Funding a team of research software engineers can cost millions/yr.
- Example: ECP has made significant investment in this area

Environmental Impact

- Carbon footprint for 30 MW is over 100 Mt per year!
- EU Green New Deal requires data centres to be carbon neutral by 2030

Challenge #2
Data

- Estimated 0.5 PB storage for 10-day forecast with
  - 3km resolution
  - 192 vertical
  - 3-hour output interval
- Storage for climate simulation will be significantly larger
- Data-in-place strategies now are fundamental
- Data loss or corruption must be addressed at this scale
- *In situ* analysis and visualization are essential tools
- With 5G/IOT volume of data to be assimilated will continue to grow

Challenge #3
Performance

- Runtime is still the main performance metric
  - *Energy-to-solution* also key consideration
- Internode communication is still an issue for parallel scaling
  - e.g., halo exchange, global reductions
- GPUs require different data-layouts and algorithms which expose more parallelism
  - Performance tuning can be notoriously difficult
  - Subtle differences between vendor microarchitecture can be important
- Denser compute nodes require greater focus on intranode communication and optimization
- Algorithmic changes can sometimes provide the greatest benefit
- Mixed-precision techniques have significant momentum

Challenge #4
Portability

- Portable software can run on
  - Different types of hardware
  - Different vendors' hardware
- Goal is to minimize
  - Lines of source code needed to achieve portability
  - Effort to run existing code on new and future types of hardware
- Want turnkey performance
  - Otherwise with minimal (automatic) parameter tuning

Challenge #5
Productivity

- Ease with which software is developed, tested, shared, maintained, documented
- Following best practices is critical for creating high-quality scientific software
- Software which is modular, composable, and extensible retains greater value, can be more easily ported/adapted
- A co-design approach is optimal
  - Scientists and research software engineers working collaboratively, communicating effectively
  - Examples:
    - ECMWF's Scalability Programme
    - German Climate Computing Center + Max Planck Institute for Meteorology
    - US ECP Co-design Centers: CEED, AMReX, COPA

Gap #1: Access to Sufficient HPC Resources

Gap #2: Access to Data Resources, Storage, Analysis Tools

Gap #3: Access to Specialized Knowledge & Skills

WMO OMM

Issu d'une présentation de Kris Row, Argonne Leadership Facility, et Mark Govett, NOAA Global System Laboratory

Summary of Recommendations

For the Research Board and WMO Members

We recommend urgency in dedicating efforts and attention to disruptions associated with evolving computing technologies that will be increasingly difficult to overcome, threatening continued advancements prediction capabilities.

The increasing scientific and computing complexity will require major efforts to adapt or rewrite earth system prediction models. In addition to scientific accuracy, models must be developed for performance, portability, and productivity.

The cost of computing resources, power consumption, and the related carbon footprint must be considered along with the benefit of improved predictability. Requirements to make data centers carbon neutral are already in force in a growing number of countries.

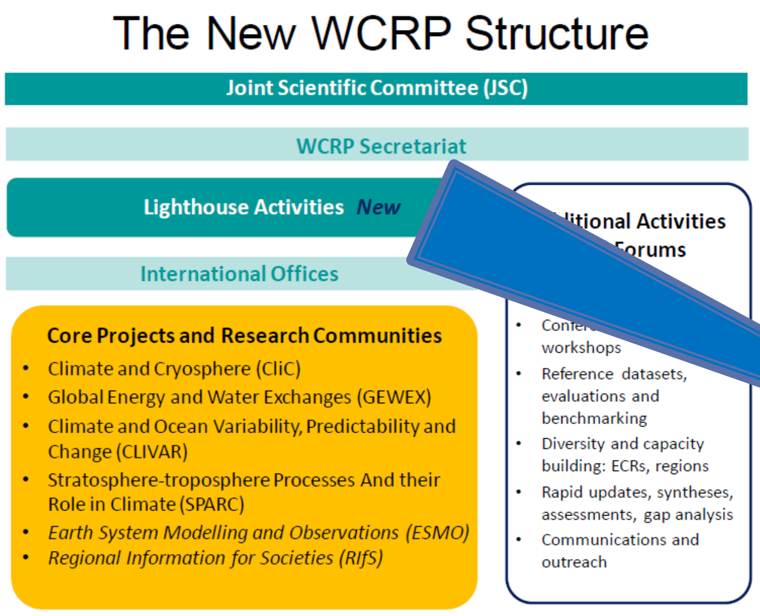
Scientists, model developers, computer scientists and software engineers need to work as equal partners on design, development, and maintenance of applications to overcome scientific, computing, and data challenges.

A data-in-place strategy is needed to support the increase in data volume from observations, model and ensemble output, and post processing. This will require co-location of HPC and data, with methods to access, extract, analyze, visualize, and store data by requesting processes & users.

WMO OMM

Issu d'une présentation de Kris Row, Argonne Leadership Facility, et Mark Govett, NOAA Global System Laboratory

Digital Earth du programme mondial de recherche sur le climat



Digital Earths Lighthouse Activity

Digital Earths = horizontal, cutting across other LHA!

Christian Jakob & Peter Bauer (co-chairs
Andreas Prein
Andrew Gettelman
Aneesh Subramanian
Bryan Lawrence
Camille Lique
Chihiro Kodama
Claas Teichmann
Dai Yamazaki
Daniel Klocke
Helene Seroussi
Mark Govett
Martin Visbeck
Michael Morgan
Pascale Braconnot
Peter Dueben
Pier Luigi Vidale
Svetlana Jebvrejeva
& Narelle van der Wel & Wenchao Cao
(soon Nico Caltabiano)

Digital Earths in WCRP should be a framework to develop new capabilities across the globe

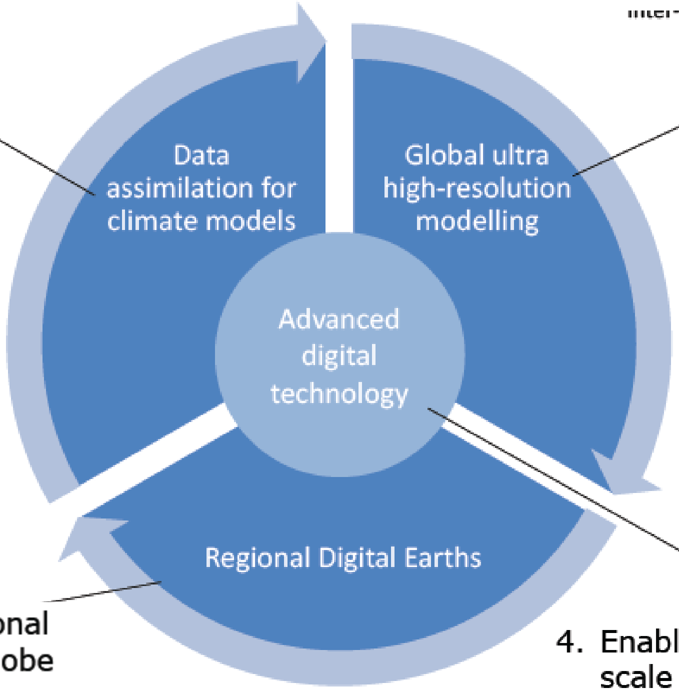
- create software infrastructures that are open and freely available, modular and interoperable
- Use cases are Global and regional Digital Twins that demonstrate benefit across WMO
- Relies on rather wide range of expertise within the WCRP family and beyond

Adapté d'une présentation de Christian Jacob et Peter Bauer, co-chairs

Digital Earth at a glance

2. Establish an active research community in data assimilation for climate that builds on the existing numerical weather prediction and re-analysis efforts and significantly expands them to fulfil the needs of Digital Earths applications

1. Establish a global research network with expertise in ultra-high-resolution (kilometer-scale or finer) of the global Earth system and its individual components



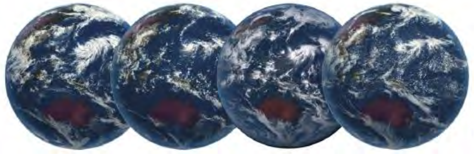
3. Support the establishment of both global and regional Digital Earths demonstration projects across the globe and provide a collaborative network for their development

4. Enable the above by optimally exploiting extreme-scale computing and data handling resources through inter-operable software infrastructures

https:// draft science plan
https://www.wcrp-climate.org/digital-earths

Adapté d'une présentation de Christian Jacob et Peter Bauer, co-chairs

Activity area 1: Global, ultra high-resolution modelling



Palmer and Stevens, PNAS, 20

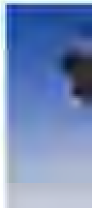
Goals:

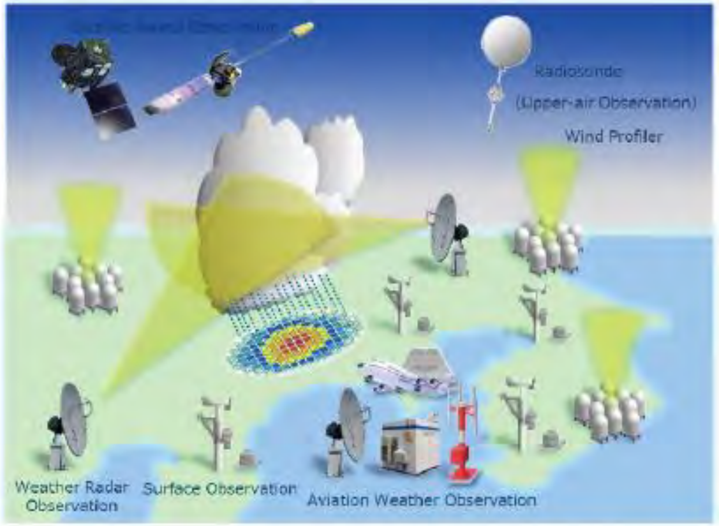
- Advance global modeling of the Earth system towards Digital Earths trelevant for human communities.
- Make Digital Earths the substrate for scientific innovation, within the WCRP and beyond

Activity area 2: Data assimilation for climate

Goals:

- Extend data assimilation capabilities to climate prediction/ projection in an operational framework.
- Provide consistent and accurate descriptions of the past and present states of the Earth System.
- Perform model testing and optimization within a unified framework.





JMA website

Activity area 3: Regional Digital Earths

Goals:

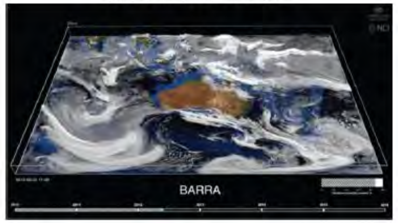
- Establish a global-regional modeling alliance that includes the all regions in the design and application of global and regional Digital Earths Systems

- Define several demonstration projects covering diversity of regional foci, and regional-global dependencies.

WRF Hydro - NOAA National Water Model

NCAR website

Australian re-analysis



Su et al., GMD 2019

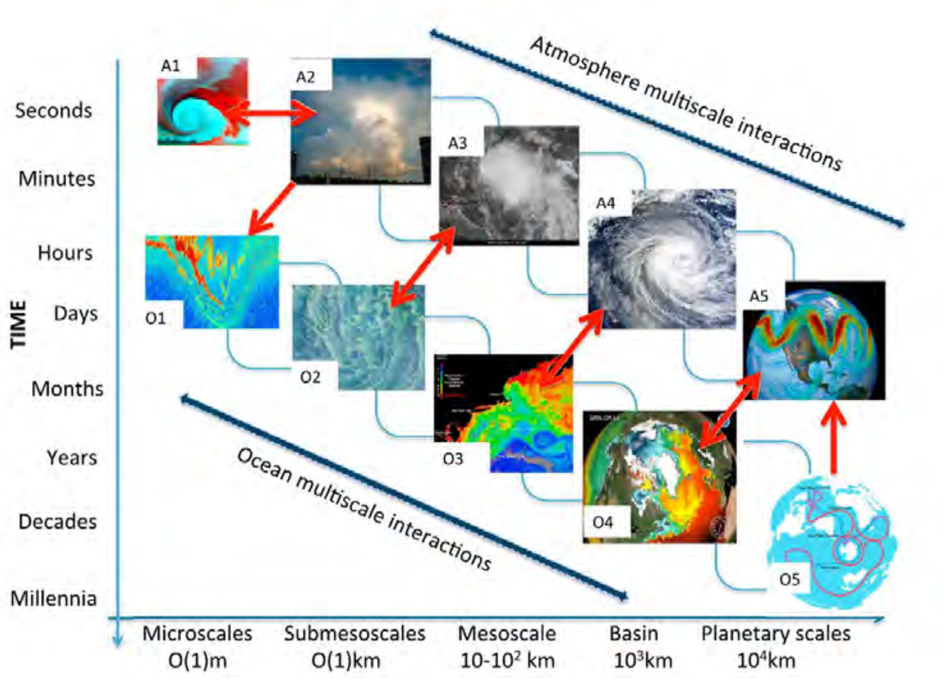
Activity area 4: Digital technology

Goals:

- Host WCRP relevant coordination activities proposed by the Research Board scoping papers on extreme-scale computing, data handling and machine learning.

Besoin de nombreux experts avec des points de vue variés!

- Soyez attentif aux annonces « workshops » etc.. sur ces sujets
- N'hésitez pas à candidater pour faire partie des groupes de travail :le processus doit rester ouvert et participatif



From Stammer et al. 2018

Message perso :

S'assurer à chaque étape que tous les milieux (atm, ocean, glace, calottes) sont bien présents et échangent sur questions, solutions, etc…

Points de vue : simulations frontières
Bien entendu, ultra haute res …. mais pas que… sur le chemin il y a des transferts d'échelles, passages entre des modèles d'échelles différentes, l'évolution des concepts etc…. qui sont aussi à accrocher dès le départ.

Suivre aussi le travail du Groupe WGNE sur le développement des modèles http://wgne.meteoinfo.ru/