Climate Modelling: What lies beneath?

The needs of large-scale climate modelling in the next decade

Bryan Lawrence

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

and the University of Reading

23-June, 2023

University of Reading

The European Network for Earth System modelling

ENES provides an umbrella organisation for the European climate modelling community working on “the understanding and prediction of climate variability and change”.

- Centered on the community working with GCMs and global ESMs, it also aims to support the recruitment community, with whom it shares common scientific objectives and issues.
- Established in 2001 in response to the 1998 recommendation from the Euroclivar concerted action to: “better integrate the European modelling effort with respect to human potential, hardware and software”.
- Active UK members: UREAD (NCAS), CEDA (NCAS, NCEO), Met Office.
- Supported by major EC projects since 2001, but specific EC support finished this year and we are putting in place a partnership with a central office supported by subscription.

What is Infrastructure in this context?

Infrastructure

The basic physical and organizational structures and facilities (e.g. buildings, roads, power supplies) needed for the operation of a society or enterprise.

Climate Model Infrastructure

- Software, hardware, standards, services, and the necessary individuals sustaining the systems, services, and software.
- The required software includes both science and technical components internal to the climate models, the workflow systems necessary to run them, and everything needed to analyse, document, manage, and exploit the data produced.
- The required hardware needs to support not only the simulation phase but the analysis and long-term storage.

Building on previous work

This activity follows on from previous work that has influenced a range of European activities:

- Provided high level recommendations based on analysis of the requirements that arise from the tensions between scientifc push and societal pull, where the former itself depends on aspects of both science and technology. *(Mitchell, Budich, Joussaume, Lawrence, Marotzke, 2012; Joussame, Lawrence, Guglielmo, 2017)*
- In 2023 we did a deeper dive into the requirements and generated two levels of recommendations. (*Lawrence, Joussaume, Adloff, 2023, to appear*).

The future of climate modelling *infrastructure*?

IS-ENES3 Deliverable D2.1

Infrastructure Strategy for Earth System Modelling for 2024-2033

What is needed to sustain large-scale European earth system modelling infrastructure from 2024 and beyond

Reporting period: 01/01/2022 – 31/03/2023

Authors: Bryan Lawrence (UREAD-NCAS), Fanny Adloff (DKRZ), Sylvie Joussaume (CNRS-IPSL)

Modelling Groups Interviewed (June 2022)

- French groups (IPSL, Cerfacs, Météo-France/CNRM)
- Italy (CMCC)
- EC-Earth groups (BSC, DMI, KNMI, SMHI)
- Germany (MPI-Met & DKRZ; AWI)
- UK groups (MetOffice, NCAS)
- Norwegian groups (NORCE, MetNorway)

Representative European Projects Investigated (All of which extend to the end of 2026 or beyond)

- EERIE
- OptimESM
- NextGEMS
- ESM2025
- EPOC
- OceanICE

Computing Context

Computing Hardware

- Accelerators: Data movement, arithmetic intensity.
- Heterogeneity: Very different characteristics between CPU and GPU systems. Will we take on FPGA?
- Memory and Storage: Tiering, bandwidth and latency (high bandwidth memory)?

Big consequences for programmability (portability, performance, productivity).

Computing Context

Computing Hardware

- Accelerators: Data movement, arithmetic intensity.
- Heterogeneity: Very different characteristics between CPU and GPU systems. Will we take on FPGA?
- Memory and Storage: Tiering, bandwidth and latency (high bandwidth memory)?

Big consequences for programmability (portability, performance, productivity).

New maths and algorithms?

- How do we get parallelism in the absence of strong scaling? New algorithms, new maths? Parallel in time?

Computing Context

Computing Hardware

- Accelerators: Data movement, arithmetic intensity.
- Heterogeneity: Very different characteristics between CPU and GPU systems. Will we take on FPGA?
- Memory and Storage: Tiering, bandwidth and latency (high bandwidth memory)?

Big consequences for programmability (portability, performance, productivity).

Machine Learning

- Emulation of existing components, parameterisations, learned resolution.
- Developing new models using high frequency data e.g. impact related.
- New analysis techniques

Big consequences for workflow and data handling.

New maths and algorithms?

- How do we get parallelism in the absence of strong scaling? New algorithms, new maths? Parallel in time?

Computing Context

Computing Hardware

- Accelerators: Data movement, arithmetic intensity.
- Heterogeneity: Very different characteristics between CPU and GPU systems. Will we take on FPGA?
- Memory and Storage: Tiering, bandwidth and latency (high bandwidth memory)?

Big consequences for programmability (portability, performance, productivity).

New maths and algorithms?

- How do we get parallelism in the absence of strong scaling? New algorithms, new maths? Parallel in time?

Machine Learning

- Emulation of existing components, parameterisations, learned resolution.
- Developing new models using high frequency data e.g. impact related.
- New analysis techniques

Big consequences for workflow and data handling.

Cost

- Hardware drives us to high resolution, high cost, small ensembles. Big user communities.
- We care not only about $ and £, but Joules!

ENES Infrastructure Strategy

Computing Context Word Salad

Computing

- Acceler
  arithme
- Hetero
  charact
  system
- Memor
  bandwi
  bandwi

Big conseq
(portability

New maths

- How do
  absenc
  algorit
  time?

ng components,
, learned resolution.

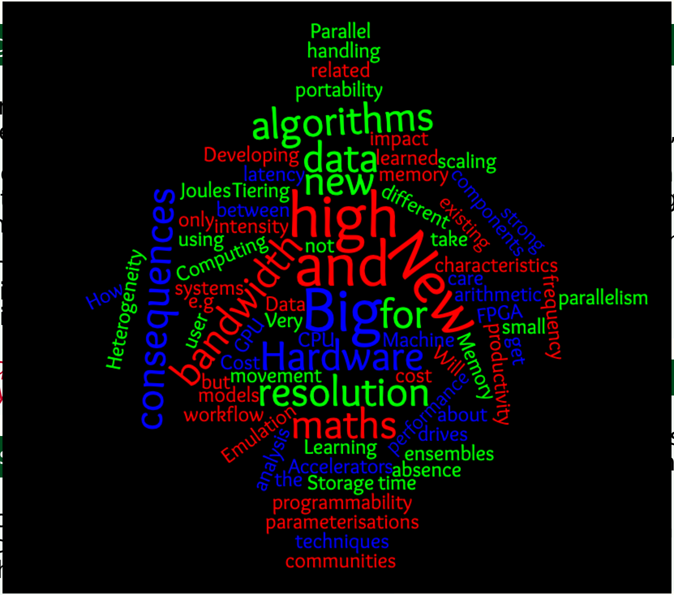
odels using high
g. impact related.

niques

workflow and data

s to high resolution,
sembles. Big user

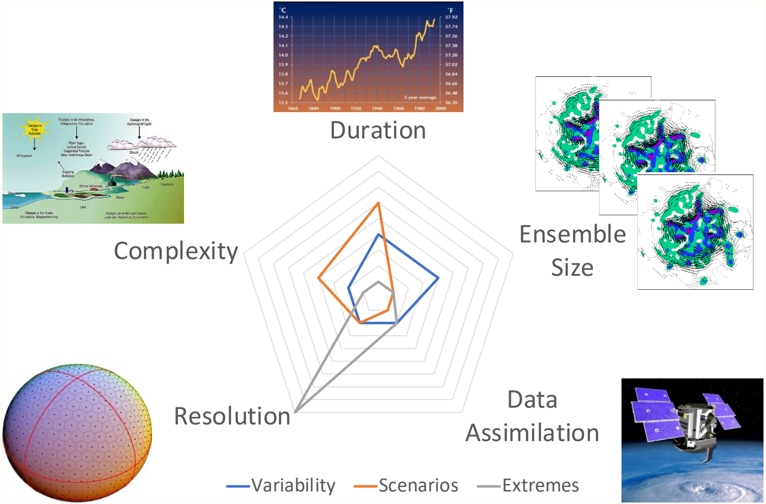
bout $ and £, but



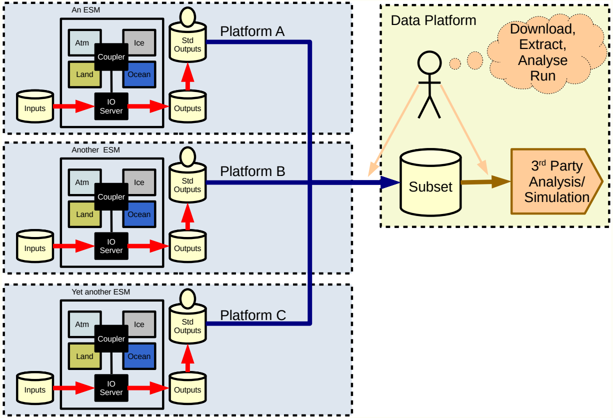
Choices

The way we think about this diagram is now wrong (if it was ever right).

These are not just choices about models, they might be choices about hardware as well!



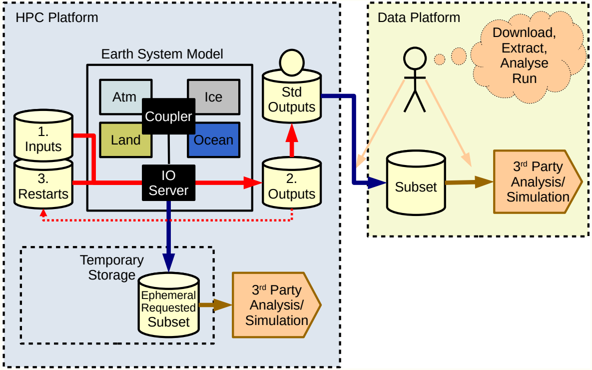
Data Workflows - CMIP6 era



- Outputs converted to standard outputs locally.
- Analysis of detailed outputs and full set of standard outputs is a local activity.
- Subsets of data brought from multiple sites to a common platform (e.g. JASMIN) for analysis.
- Considerable effort to standardise data.
- Data analysis separated from production over both space and time: issues for documentation and citation!

Data Workflows - into the future

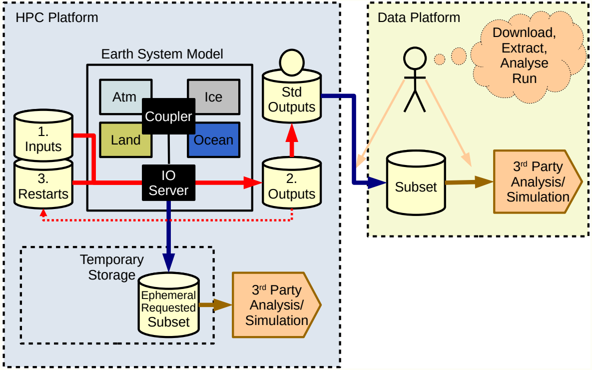
All of the issues from the CMIP era and some new ones!



- In-flight diagnostics supported by a coupler and/or IO-server (visualisations, ensemble diagnostics, downstream models using high frequency data).
- Data published to data platforms for wider sharing and analysis,
- Data platforms supporting co-located computation to support *bringing compute to the data*.

Data Workflows - into the future

All of the issues from the CMIP era and some new ones!

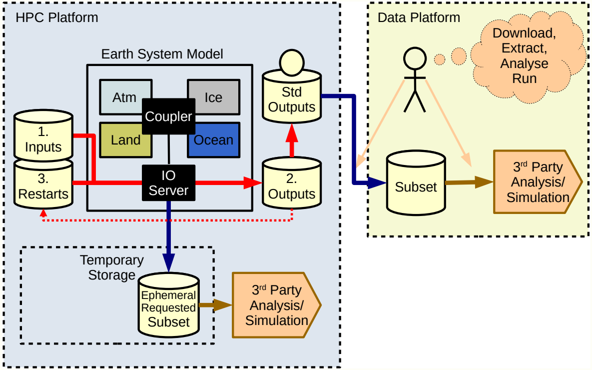


Cloud has a place in data platforms, but it will not be the only solution.

- In-fli … ed by a coup … (visu … diagr … lels using
- Data p … ms for wic
- Data p … pporting co-located computation to support *bringing compute to the data.*

Data Workflows - into the future

All of the issues from the CMIP era and some new ones!



- In-flight diagnostics supported by a coupler and/or IO-server (visualisations, ensemble diagnostics, downstream models using high frequency data).
- Data published to data platforms for wider sharing and analysis,
- Data platforms supporting co-located computation to support *bringing compute to the data*.

Our view of producers and consumers will have to change as we treat large modelling projects more like satellite missions:

- Well advertised in advance, community dicussions about what is important, well documented, etc.
- We will need to invest even more in our data systems and standards!

Scientific Landscape

WCRP Lighthouses

- Explaining and predicting climate change (EPESC) (annual to decadal)
- Safe Landing Climates (pathways to the future)
- My Climate Risk (local perspective, storylines)
- Digital Earth (integrated systems, km-scale)

Global Science

- CMIP5, CMIP6, CMIP7 (MIPs past and future)
- SM(I)LEs (More data sharing)

Technical Collaboration

- ESGF (R&D, delivery)
- Standards & Software: ESMValTool, Workflow, OASIS, XIOS, CF, Pangeo etc

European Science

- ESM2025, NextGEMS (ESM & GCMs *varying* resolution)
- EERIE, OptimESM (Oceans and variable resolution)
- DESTINE (high resolution, EO etc)
- Collaboration: NEMO, SI3, etc
- Many other projects: EPOC (medium resolution); OCEAN:ICE (observations and long duration); etc

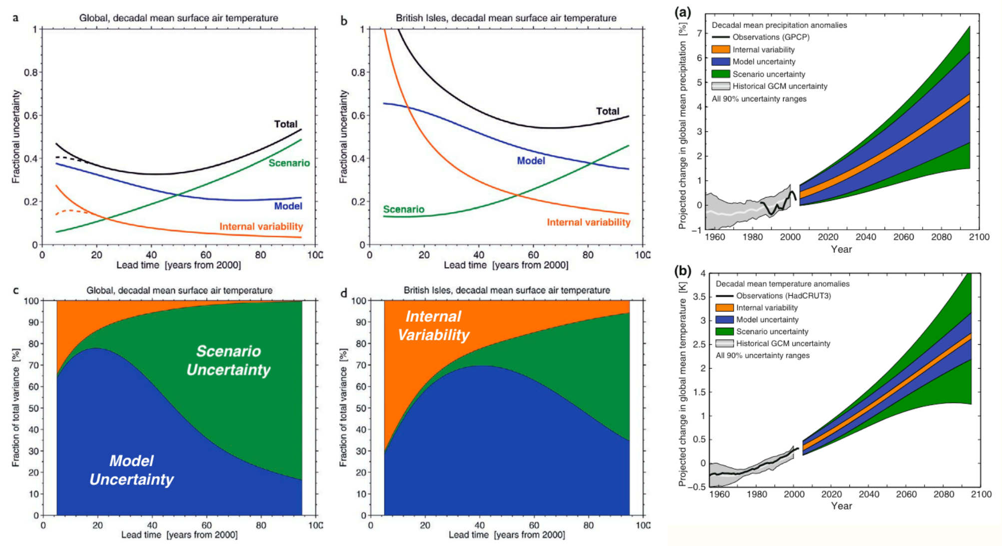
Technical Integration

- EuroHPC, Copernicus, ESA, EOSC etc

Climate Services

- Portals, data products, projections, etc

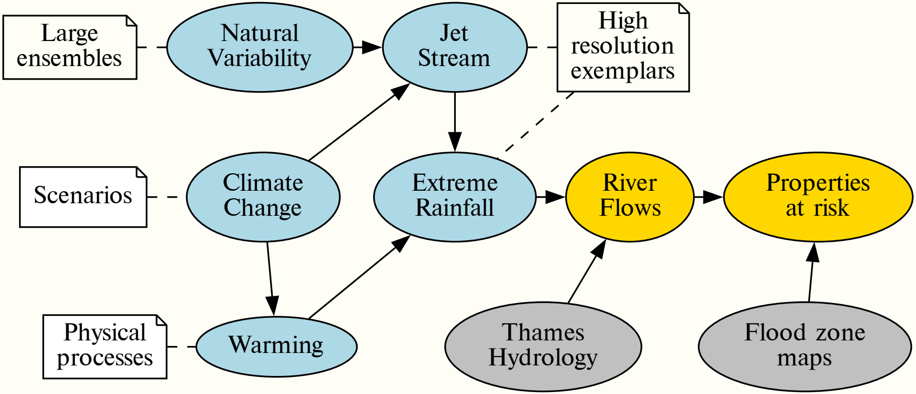
Uncertainty



Hawkins & Sutton (2009) https://doi.org/10.1175/2009BAMS2607.1
Hawkins & Sutton (2011) https://doi.org/10.1007/s00382-010-0810-6

Storylines, Climate Modelling, and Vulnerabilty

(Lloyd & Shepherd (2020). https://doi.org/10.1111/nyas.14308)

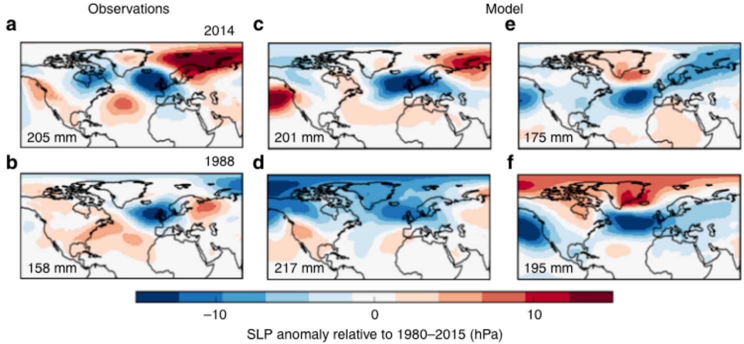


Multiple different modelling approaches, multiple different communities, within the physical climate modelling community and beyond.

Interesting issues around information and data flow!

The seen and unseen climate

Exploiting global ensemble prediction systems to investigate seasons and extremes that COULD have occurred. (Thompson et al, 2017, doi:10.1038/s41467-017-00275-3).

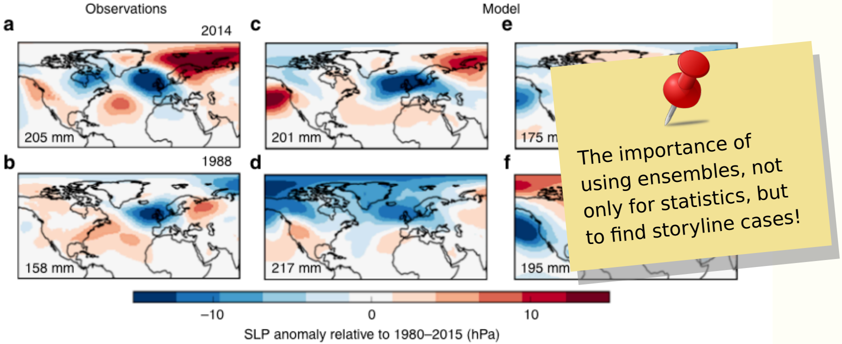


Observed and not yet realised climate states:

- a, b: The sea level pressure anomaly fields (in hPa, relative to the January MSLP field) of the two observed ERA Januarys with highest rainfall totals: 2014 and 1988.
- cf: The sea level pressure anomalies of four extreme simulated Januarys, one of which — d — presents a potential new record rainfall scenario
- (Interesting question as to importance of UPSCALE effects. Could these have been seen in RCM simulations?)

The seen and unseen climate

Exploiting global ensemble prediction systems to investigate seasons and extremes that COULD have occurred. (Thompson et al, 2017, doi:10.1038/s41467-017-00275-3).



Observed and not yet realised climate states:

- a, b: The sea level pressure anomaly fields (in hPa, relative to the January MSLP field) of the two observed ERA Januarys with highest rainfall totals: 2014 and 1988.
- cf: The sea level pressure anomalies of four extreme simulated Januarys, one of which — d — presents a potential new record rainfall scenario
- (Interesting question as to importance of UPSCALE effects. Could these have been seen in RCM simulations?)

On Verifying and Validating Models

Definitions

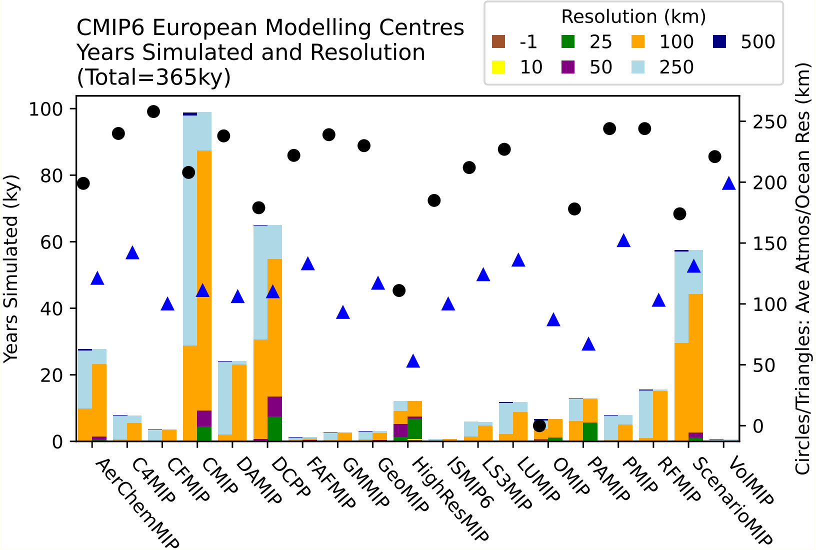
- *Verification*: "To establish the truth", but we can never do that — a model is never a closed system, there are always explicit, implicit, and 4th Rumsfeldian assumptions.
- *Validation*: " To establish legitimacy", but of necessity this is "*legitimacy in context*" which is not necessarily the same as the *quality* of the *representation* of the real world as embodied in the model and evaluated by comparison with observations!

See Oreskes et al (1994) and Parker (2020).

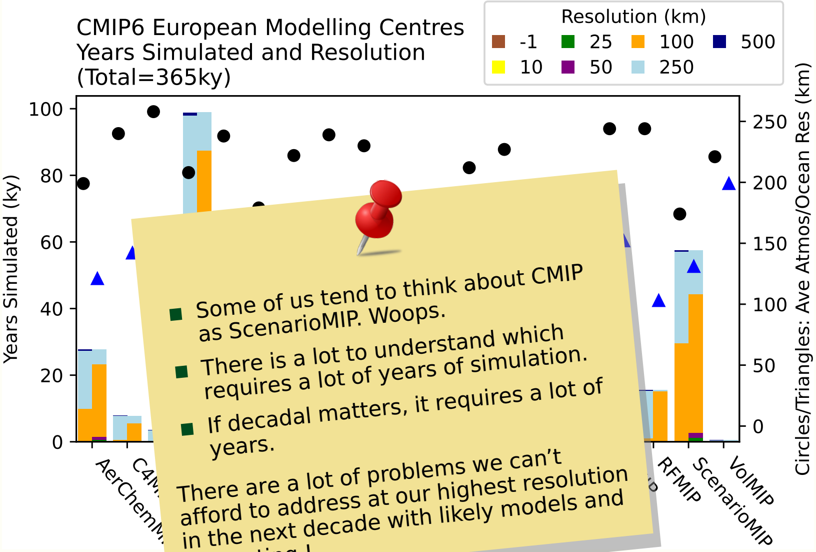
Fitness and Adequacy for purpose

- We need to avoid the "fallacy of affirming the consequent". If *A* implies *B*, and we observe *B*, in a non-closed system, we cannot be sure that there is not some other cause *C* that also implies *B*. We're probably more familiar with thinking about this in the context of compensating errors where we know our model might be getting "better" for the wrong reasons.
- Adequacy and fitness also encompass practicality. I might think this model is better, but another cheaper model may be adequate. Indeed a model we presume to be fitter may be so impractical that it cannot be used . . . which is of course where we are with our kiviat diagram!

European CMIP6 effort in years as of 2022



European CMIP6 effort in years as of 2022



- Some of us tend to think about CMIP as ScenarioMIP. Woops.
- There is a lot to understand which requires a lot of years of simulation.
- If decadal matters, it requires a lot of years.

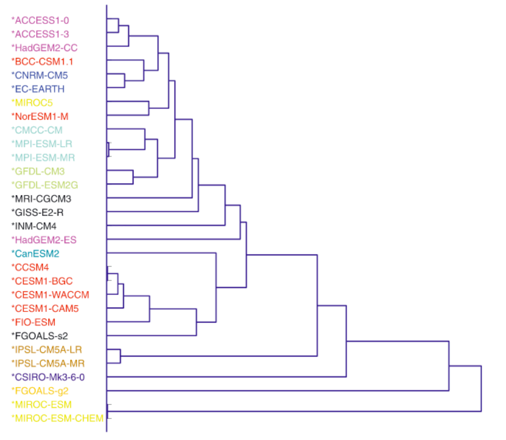
There are a lot of problems we can't afford to address at our highest resolution in the next decade with likely models and computing !

A spectrum of climate modelling

		10K Time	Spinup, Control & Millenium	Scenarios (Past, Future, Sensitivity)	Single Model Large Ensembles (SMLEs)	Decadal Prediction (Hindcasts, Forecasts)	Equilibrium /Transient Timeslices	Process Studies	Storyline Timeslices
Duration		> 1000y	o(1000)y	o(100)y		o(10)y		o(10-100)d	
Domain	ESM	x	x	x	x				
	GCM		x	x	x	x	x	x	x
	RCM			x			x	x	x
Speed (SYPD)		> 50	> 4	> 4	> 1	> 1	> 0.5	> 0.1	> 0.1
Ensemble Size		1-1000	o(1)	o(10)	10-100	o(10) x NH*	o(1)	o(1)	o(1)
Atmos XY Res (km)		o(500)	100-500	25-500	50-300	25-500	1-100	1-10	1-10
Ocean XY Res (km)		300+	100-250	25-250	25-250	25-250	10+	2.5+	2.5+
Exascale Status		Currently impossible		Not currently usable			Currently some capability		

(NH*: Evaluating a decadal prediction system might involve NH hindcasts, each with o(10) ensemble members)

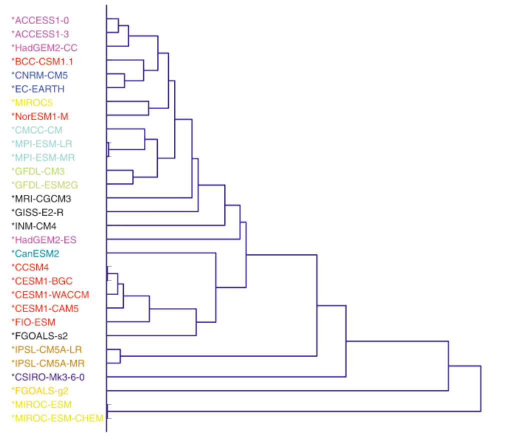
Model Diversity and model uncertainty



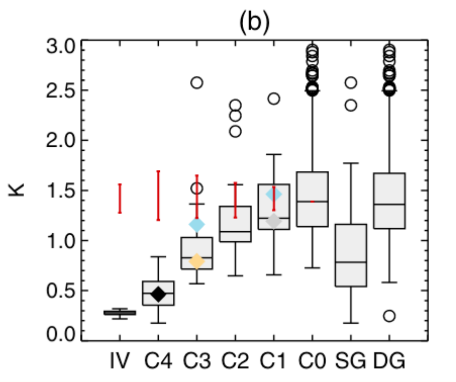
Knutti, et.al (2013) https://doi.org/10.1002/grl.50256

Model Diversity and model uncertainty

ENES Infrastructure Strategy

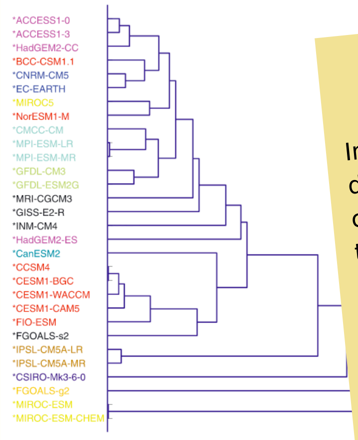


Knutti, et.al (2013) https://doi.org/10.1002/grl.50256



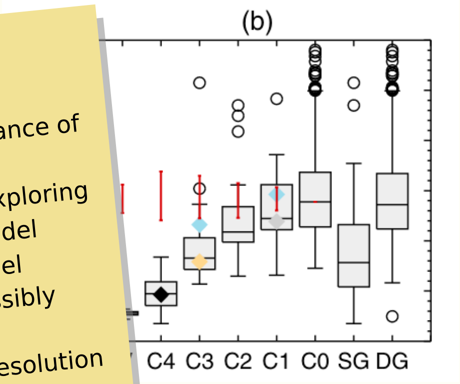
Pairwise RMSE for simulations with different numbers of shared components: C1 share only one component, C2 two components, C4 differ only in resolution. (Components have to be different, not just parameters).
Boé (2018) https://doi.org/10.1002/2017GL076829

Model Diversity and model uncertainty



Knutti, et.al (2013) https://doi.org/10.1002/grl.502

Incredible importance of differing model components in exploring the impact of model diversity on model uncertainty! Possibly even exceeds importance of resolution *in some contexts?*

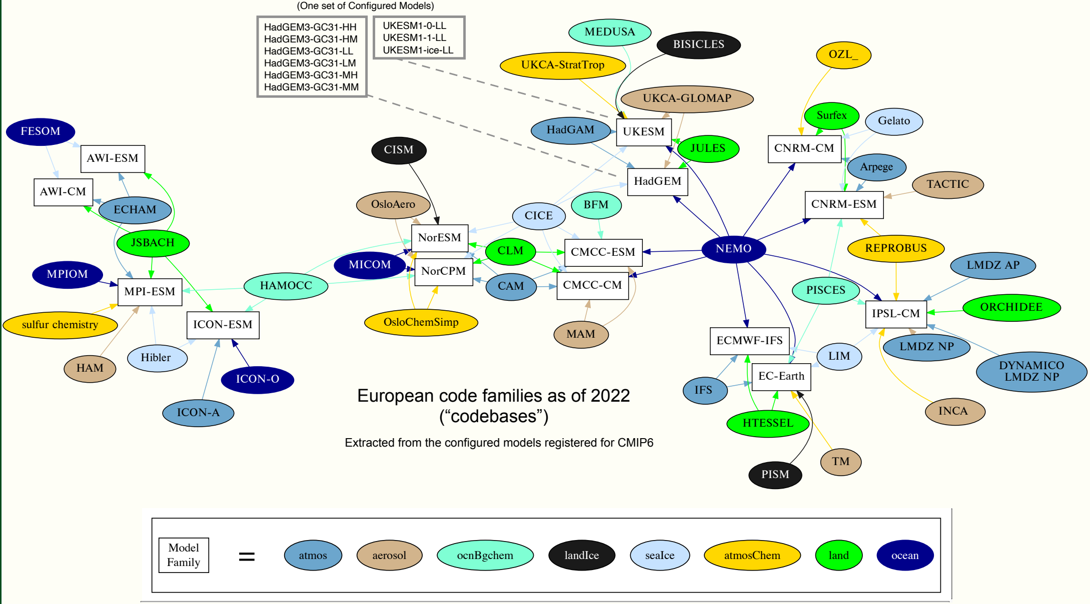


...r simulations with different ...ed components: C1 share only one component, C2 two components, C4 differ only in resolution. (Components have to be different, not just parameters).

Boé (2018) https://doi.org/10.1002/2017GL076829

ENES Infrastructure Strategy

European Model Diversity (2022)



European code families as of 2022 ("codebases")

Extracted from the configured models registered for CMIP6

Model Diversity Consequences

- We know that investing in building some classes of model requires a huge effort from large communities.
- There is a finite number of communities that can make these sorts of investment.

Model Diversity Consequences

- We know that investing in building some classes of model requires a huge effort from large communities.
- There is a finite number of communities that can make these sorts of investment.
- But we know that model diversity matters, and *making good use of it* will remain a key part of understanding climate uncertainty for the foreseeable future.

Model Diversity Consequences

- We know that investing in building some classes of model requires a huge effort from large communities.
- There is a finite number of communities that can make these sorts of investment.
- But we know that model diversity matters, and *making good use of it* will remain a key part of understanding climate uncertainty for the foreseeable future.
- Are there risks in too much interdependency on shared components?

Model Diversity Consequences

- We know that investing in building some classes of model requires a huge effort from large communities.
- There is a finite number of communities that can make these sorts of investment.
- But we know that model diversity matters, and *making good use of it* will remain a key part of understanding climate uncertainty for the foreseeable future.
- Are there risks in too much interdependency on shared components?
- So what is the right amount of diversity, and how can we ensure we fully exploit it?

Model Diversity Consequences

- We know that investing in building some classes of model requires a huge effort from large communities.
- There is a finite number of communities that can make these sorts of investment.
- But we know that model diversity matters, and *making good use of it* will remain a key part of understanding climate uncertainty for the foreseeable future.
- Are there risks in too much interdependency on shared components?
- So what is the right amount of diversity, and how can we ensure we fully exploit it?

I think we will need to spend more time as a community *planning* when and where we can best use diversity!

Foresight Recommendations - Headlines

There are seven headline recommendations, nearly all of which are relevant for NCAS, either directly, or for our service providers (such as national HPC):

1. The HPC community must continue to provide both CPU and GPU machines, a lot of climate codes will not be sensibly deployed on GPU machine in the near future (even if they can be made to run on them).
2. There is a need for a more operational aspect to some aspects of climate science.
3. The community should continue to invest in managing and sustaining shared insfrastructure.
4. *Model development takes a long time and is resource intensive*. Modellers will have to pay attention to the choices between *Performance, Portability, and (scientific) Productivity*; in the new world, we can only have two!
5. Large expensive modelling projects need to be treated like satellite missions, *well publicised and documented*.
6. The community should continue to *invest in the necessary underpinning diagnostic tools and libraries*.
7. Storage and data systems need to support a variety of use-cases.

Foresight Recommendations - Headlines

There are seven headline recommendations, nearly
NCAS, either directly, or for our service providers (su

1. The HPC community must continue to provide both C
climate codes will not be sensibly deployed on GPU m ... if
they can be made to run on them).
2. There is a need for a more operational aspect to some
3. The community should continue to invest in managing ... sustaining shared insfrastructure.
4. *Model development takes a long time and is resource intensive*. Modellers will have to pay attention to the choices between *Performance, Portability, and (scientific) Productivity*; in the new world, we can only have two!
5. Large expensive modelling projects need to be treated like satellite missions, *well publicised and documented*.
6. The community should continue to *invest in the necessary underpinning diagnostic tools and libraries*.
7. Storage and data systems need to support a variety of use-cases.

These strategic requirements underpin a lot of what CMS has done, and will do.

Some selected recommendations - JASMIN related

4. National and International HPC providers should enhance provision for large-scale storage systems with co-located analysis compute, both CPU and GPU, suitable for machine learning and distributed analytic workflows.
5. Storage systems should include tiered storage which supports both high-performance data-analysis and the storage of exabytes of data, with minimal carbon cost and relatively (with respect to commercial) high turnover between tiers.
6. Systems providers should recognise the need to support a range of different access profiles for different users, ranging from full batch access to analysis compute access, storage only, and remote usage via gateway service software (e.g., DASK Gateway, WPS and other community standard protocols as they become prevalent).
7. Storage and analysis systems should have interfaces and user management systems that support membership of multiple different federations with differing authentication and authorisation regimes.

Selected modelling recommendations

11 It will be important to seek out new methods to increase speed and decrease the cost of simulations. Key opportunities exist in the use of *variable resolution, mixed precision, and machine learning*, but it is also likely that completely new algorithmic paradigms will also be necessary.

12 Modelling systems will need to accommodate *more flexibility in on-line diagnostics and interfaces* with downstream applications.

13 Model output should use *state-of-the-art compression techniques*, and clearly document the impact of such compression on the inherent information content.

17 Large expensive modelling projects need to be treated like satellite missions: *well publicised, and documented*, engaging with user communities to support the maximum efficient use of data products, both ephemeral and persistent.

32 *Data users* will also continue to need *appropriate documentation* as to how and why data were produced, and to be able to discover and report issues with the data after *simulation* have concluded. *Systems* to streamline the production and use of such information will *need to be improved and maintained*.

Diagnostics and Data Recommendations

22 Shared diagnostic tools and libraries provide efficiencies into the exploitation of model data. Ongoing investment will be needed . . . as *new data formats, new methods of compression, and new meshes* are introduced.

23 New investments will be needed to develop and/or maintain tools which can *facilitate high volume data analysis*, such as those which exploit parallelism, or which expedite data exploration and selection.

27 Data volumes will continue to grow, and simulations will continue to be carried out on multiple platforms. It will be necessary to maintain *distributed catalogue systems* and methods to replicate data to national and international archives with co-located analysis compute.

29 Even with large cache archives, not all data will be collocated for all workflows, and so software systems to support *distributed analytics* will need to be developed and integrated into standard tools.

30 Aspirations to share data systems, catalogues and data analytics will continue to demand *common standards for data storage and metadata*. Modellers should continue to use and extend the *Climate Forecast conventions* to maximise data re-use in accordance with FAIR principles.

Workforce Recommendations

33 In order to make the best use of new technologies it is necessary to *enhance training programmes*, both for new entrants into climate simulation, and for mid-career scientists. It is also necessary to maximise information sharing between European modelling groups and those carrying out data analysis.

34 Alongside training the existing workforce, there is a need to *grow the size of that workforce*, to support the software and scientific developments needed to sustain the delivery of climate science, while retooling models and analytic workflows for next generation computing and to make the best use of AI/ML.

35 The need for a growing workforce may involve growing by collaboration rather than by recruitment alone. Finding ways of engaging and funding already trained individuals from other disciplines to deliver precursor work for environmental science is necessary – e.g. *encouraging computer science funders to support programmes which enable practical climate modelling outcomes*.

36 To support the full spectrum of activities, from science to software engineering, it may be necessary to facilitate, encourage, and *reward individuals who transition between science and software engineering and vice-versa*.

An infrastructure Strategy?

Diversity!

- There is no one clean definition of what a climate model is and when it is fit for purpose. We will need to consider *adequacy given validation for purpose when constrained by cost*.
  - The George Box quote " All models are wrong, some are useful", might be better adjusted to "All models are wrong, different ones are useful for different purposes".
- We are going to have to get used to a compromise between performance, portability and (scientific) productivity.
- At scale, we are going to have to invest even more into data systems and standards.
- We need to worry (even) more about our workforce.
- At scale, as we invest more into particular platforms, we will need to think about which platforms, and why, and *manage* scientific diversity.
- At scale, we are going to have to treat our investments more like satellite missions, and justify our experiments in the court of our peers!