Perspectives on the Future of Climate Modelling

(heavily based on the nascent infrastructure strategy for ENES)

Bryan Lawrence

National Centre for Atmospheric Science
NATURAL ENVIRONMENT RESEARCH COUNCIL

National Centre for Atmopheric Science
and the *University of Reading*

15-May, 2023

The future of climate modelling?

The future of climate modelling?

PHILOSOPHICAL TRANSACTIONS A

royalsocietypublishing.org/journal/rsta

Discussion

Check for updates

Cite this article: Balaji V. 2021 Climbing down Charney's ladder: machine learning and the post-Dennard era of computational climate science. *Phil. Trans. R. Soc. A* **379**: 20200085. https://doi.org/10.1098/rsta.2020.0085

Climbing down Charney's ladder: machine learning and the post-Dennard era of computational climate science

V. Balaji[1,2]

[1]Princeton University and NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA

[2]Institute Pierre-Simon Laplace, Paris, France

VB, 0000-0001-7561-5438

The future of climate modelling?

PHILOSOPHICAL TRANSACTIONS A

royalsocietypublishing.org/journal/rsta

Discussion

Check for updates

Cite this article: Balaji V. 2021 Climbing down Charney's ladder: machine learning and the post-Dennard era of computational climate science. *Phil. Trans. R. Soc. A* **379**: 20200085. https://doi.org/10.1098/rsta.2020.0085

Climbing down Charney's ladder: machine learning and the post-Dennard era of computational climate science

V. Balaji[1,2]

[1]Princeton University and NOAA/Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA

[2]Institute Pierre-Simon Laplace, Paris, France

VB, 0000-0001-7561-5438

(I am in no way implying Balaji is a witch!)

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!

Computing Context Word Salad

Computing

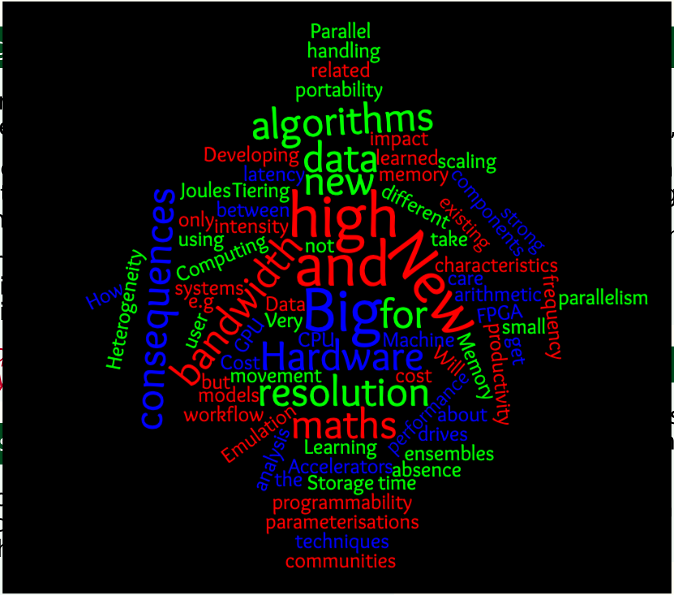
- Acceler ... ng components, arithme ... , learned resolution.
- Hetero ... odels using high charact ... g. impact related. system
- Memor ... niques bandwi bandwi

Big conse ... workflow and data (portability

New maths

- How d ... s to high resolution, absen ... nsembles. Big user algorit time?

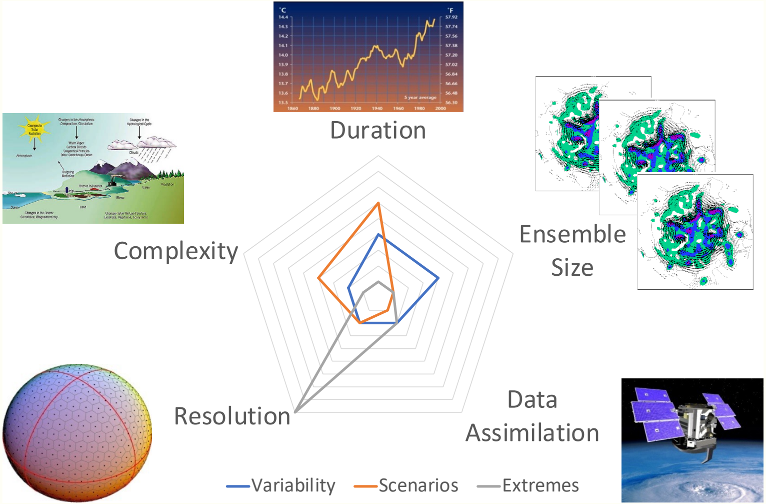
bout $ and £, but



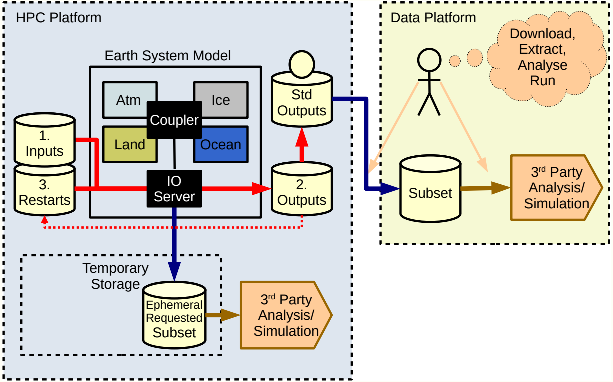
Choices

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!

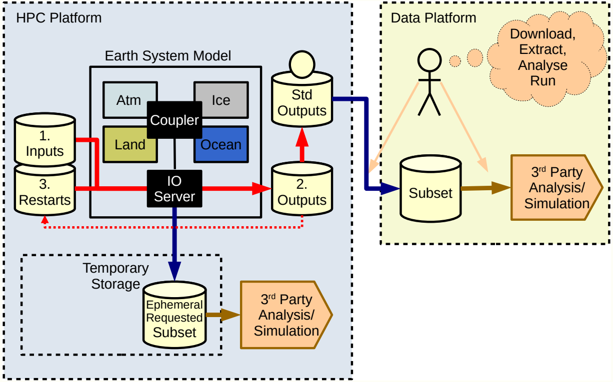


Data Workflows



- 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*.

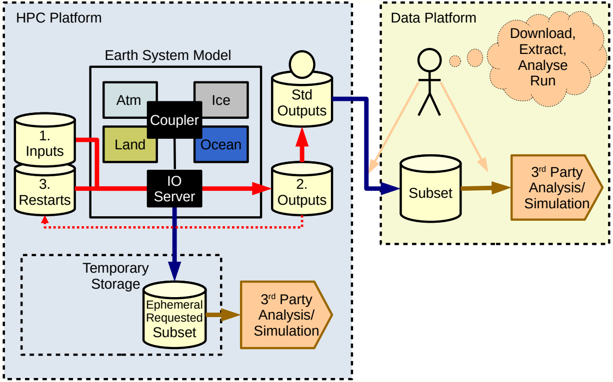
Data Workflows



- In- ... rted by a c... (vis... dia... odels usi...
- Dat... orms for v... is,
- Data ... co-lo... computation to support *bringing compute to the data*.

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

Data Workflows



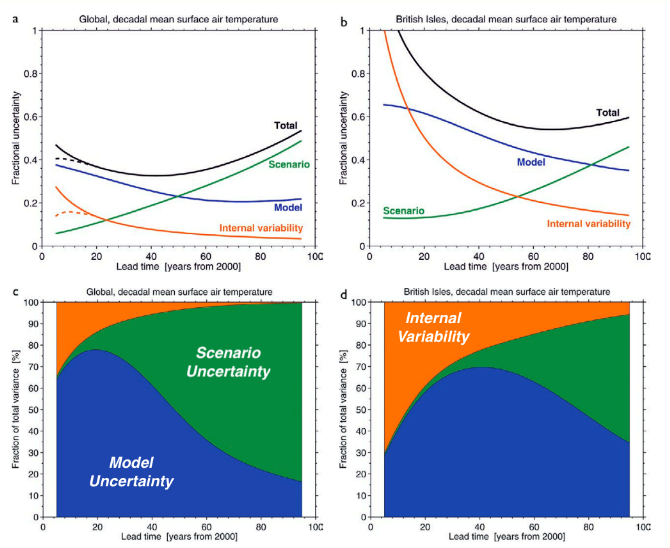
- In-flight diagnostics supported by a coupler and/or IO-server (visualisations, ensemble diagnostics, downstream models using high frequency data).
- Data published to data platforms for wider sharing and analysis,
- Data platforms supporting co-located computation to support *bringing compute to the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

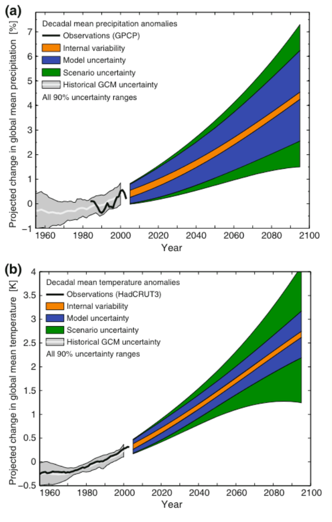
Our view of producers and consumers will have to change as we treat large modelling projects more like satellite missions:

- 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!

Uncertainty

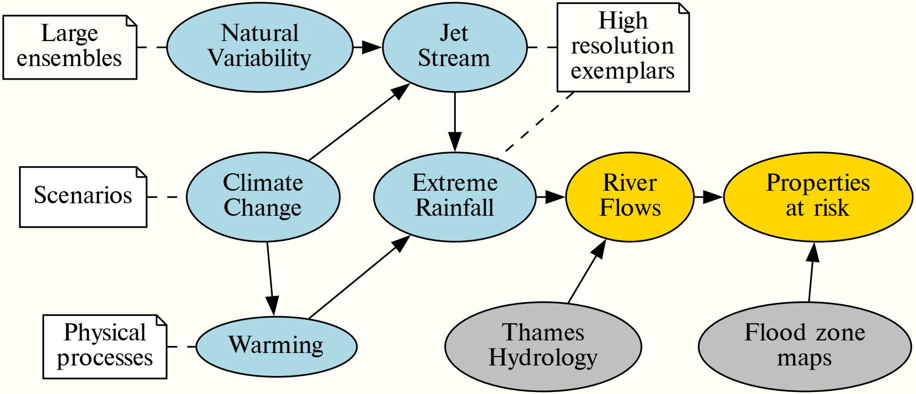




Hawkins & Sutton (2009) https://doi.org/10.1175/2009BAMS2607.1
Hawkins & Sutton (2011) https://doi.org/10.1007/s00382-010-0810-6

Storylines, Climate Modelling, and Vulnerabilty

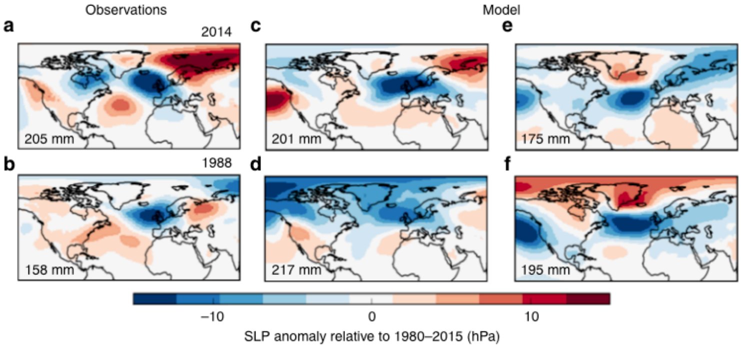
(Lloyd & Shepherd (2020). https://doi.org/10.1111/nyas.14308)



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.

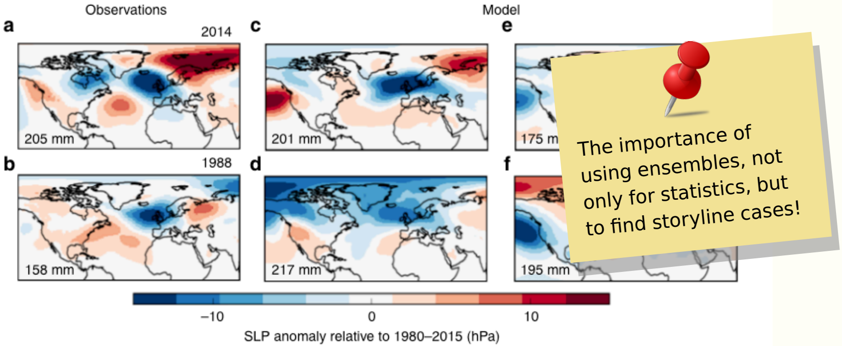


Observed and not yet realised climate states:

- 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

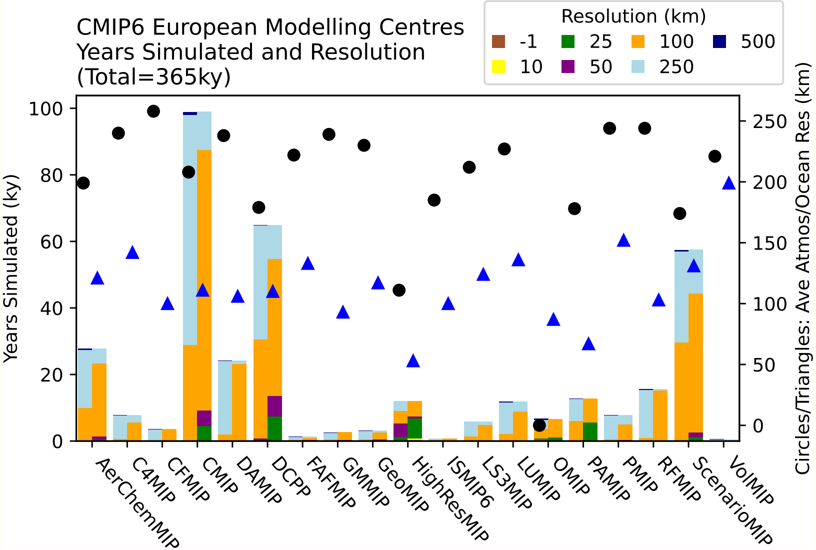
Exploiting global ensemble prediction systems to investigate seasons and extremes that COULD have occurred.



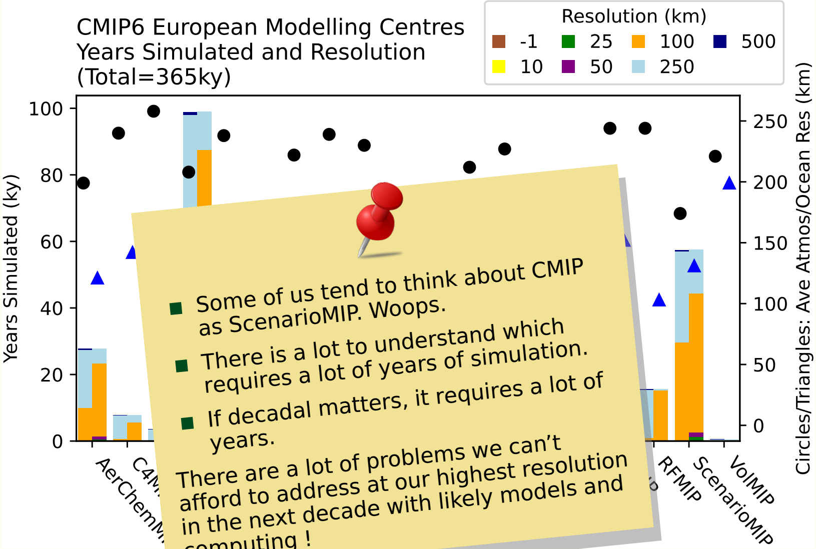
Observed and not yet realised climate states:

- 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?)

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 Predic-tion (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)

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!

Future of Climate Modelling

Methodology

Technical Context

Scientific Context

CMIP

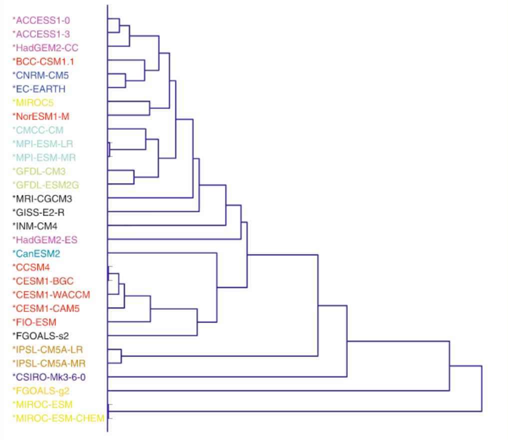
Modelling

Fitness

Diversity & Uncertainty

Summary

Model Diversity and model uncertainty



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

University of Reading

Model Diversity and model uncertainty

Future of Climate Modelling

Methodology

Technical Context

Scientific Context

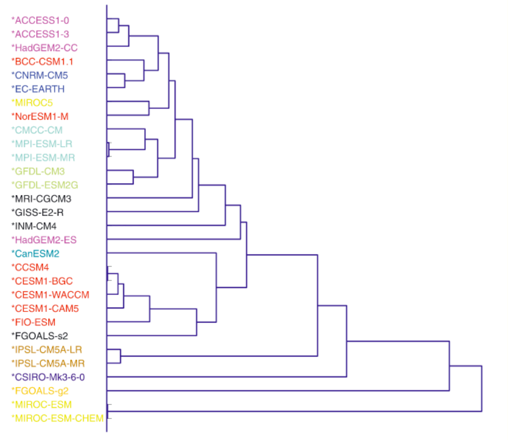
CMIP

Modelling

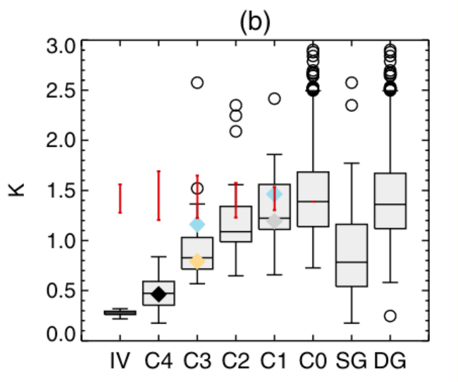
Fitness

Diversity & Uncertainty

Summary



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

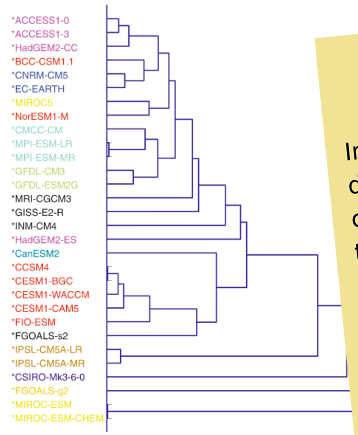


Pairwise RMSE for simulations with different numbers of shared components: C1 share only one component, C2 two components, C4 differ only in resolution. (Components have to be different, not just parameters).

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

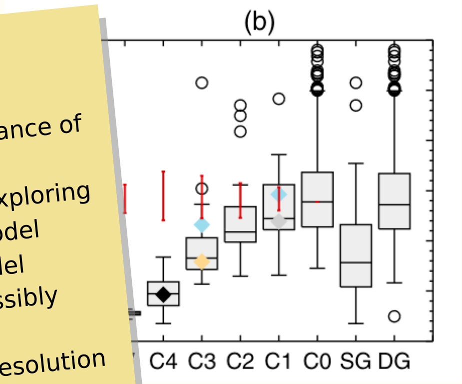
University of Reading

Model Diversity and model uncertainty



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

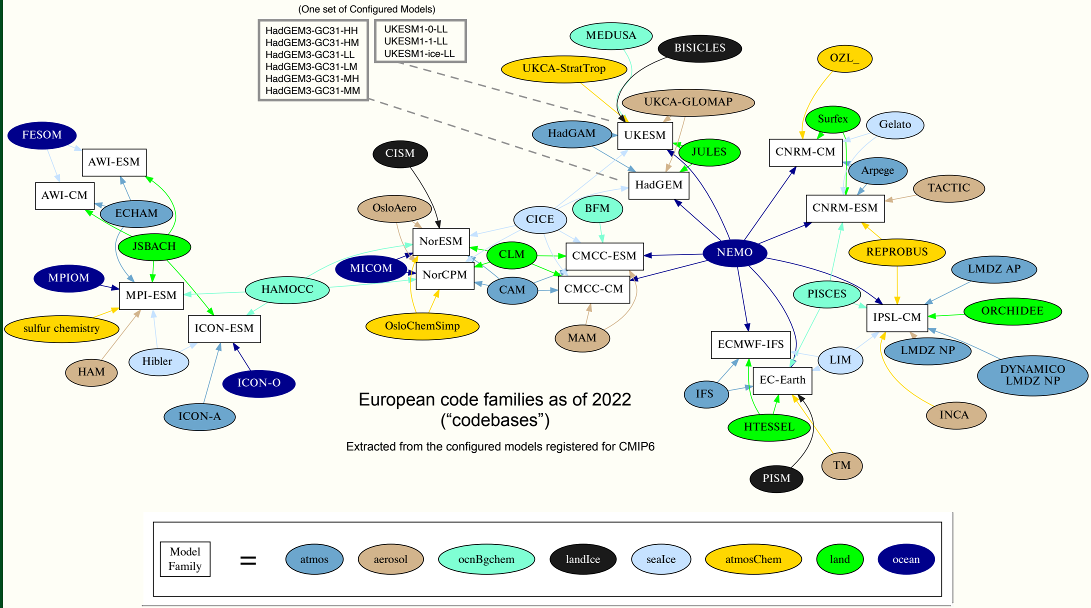
Incredible importance of differing model components in exploring the impact of model diversity on model uncertainty! Possibly even exceeds importance of resolution *in some contexts?*



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

European Model Diversity (2020)



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!

The future of climate modelling?

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.
- 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!