“From Robustness to Antifragility: A Philosophical Perspective on Climate and Earth System Models"

Ryan O’Loughlin
Department of Philosophy
roloughlin@qc.cuny.edu

QUEENS COLLEGE

Prologue

- Why care what a philosopher has to say: **shared goal (kind of)**
- Methodological similarities: **both look at empirical evidence**
- Methodological differences:
  - **modeling and computing vs. scientific-practice-as-evidence**
  - **differences in specialized terminology**

Key take-away points

1. *Convergence between climate models can be <u>informally</u> confirmatory* (model robustness)

2. *Failures in modeling can provide opportunities to yield new knowledge* (model antifragility)

Content

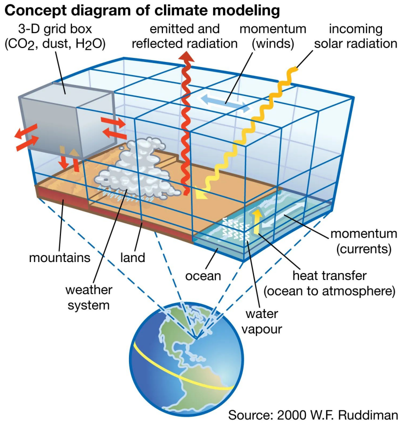
- Setting the stage – which models I'm talking about
- Convergence and model robustness
- Informal confirmation
- Failures and Antifragility
- Example: Historical Constraints
- Conclusions and Implications

Content

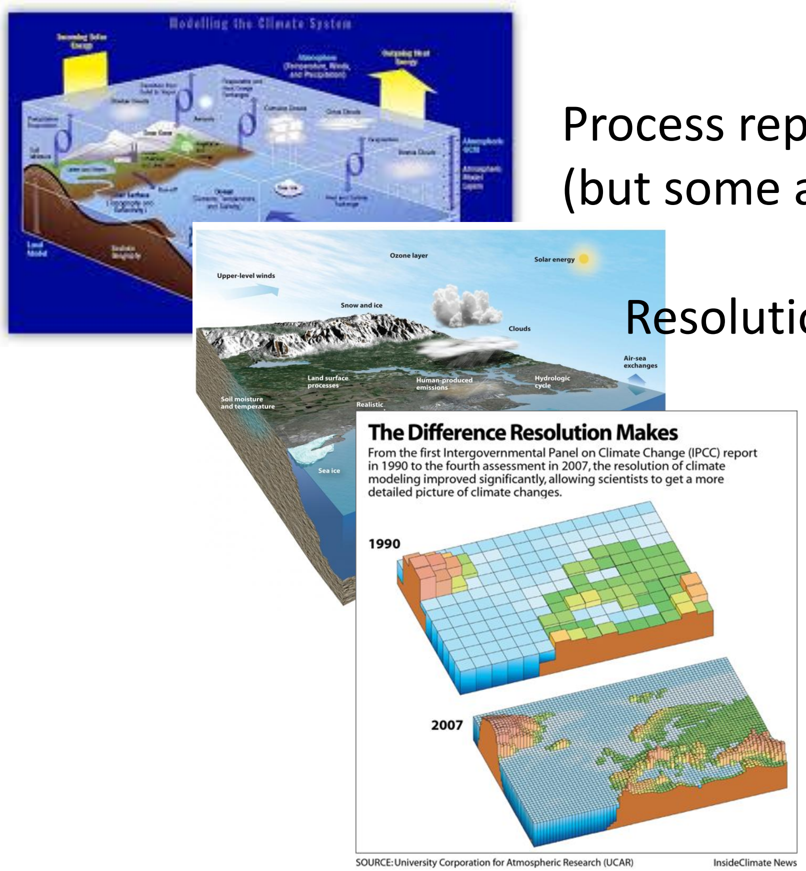
General Circulation Models (GCMs)

Same earth

Same core physics



Different models = different renderings with the same core

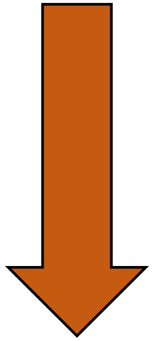


Process reps vary (but some are shared)

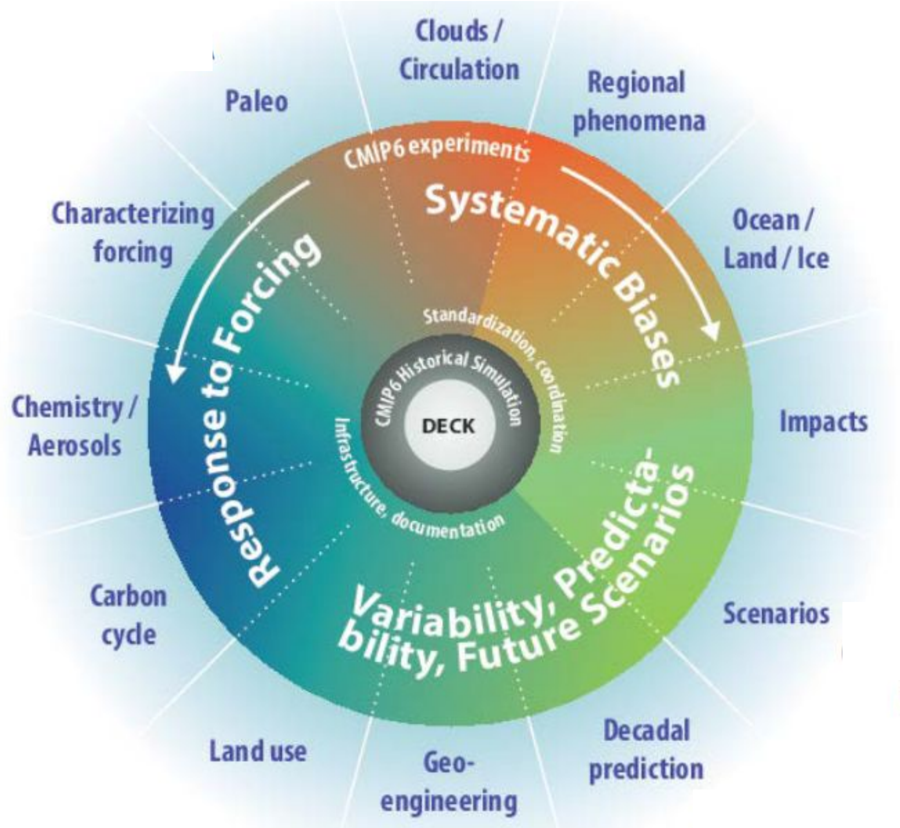
Resolutions vary

Model intercomparisons: continuity of a research strategy

- Atmospheric (1990s): 31 modeling groups, 31 distinct models



- Coupled (sixth phase; today): 49 modeling groups; 100 distinct models



National flag circles mark over 100 climate modeling groups — including the NASA Goddard Institute for Space Studies (GISS) — contributing to the Coupled Model Intercomparison Project Phase 6 (CMIP6). Google Map by CMIP6

Content

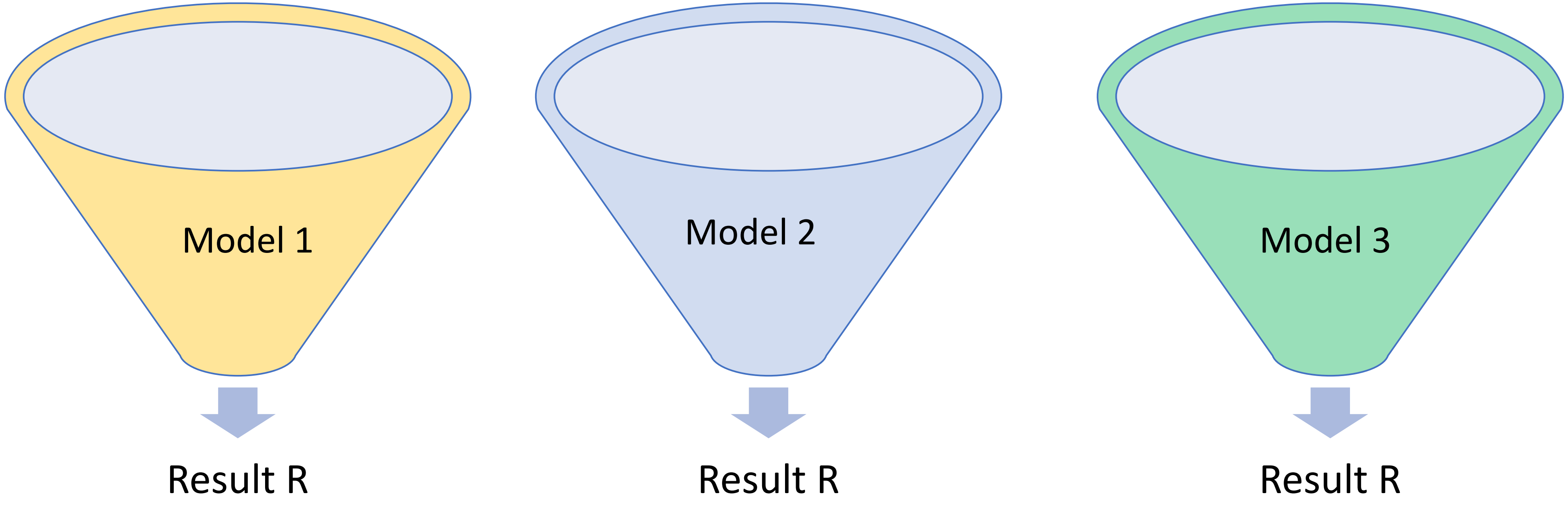
Model robustness and practice-informed philosophy of science

Lisa Lloyd
Distinguished Professor
Emerita at Indiana University

Stu Gluck
Director, Federal Railroad
Administration Office of Industry
Data and Economic Analysis @ U.S.
Department of Transportation

Typical framing…

Hypothesis of interest: under a high-emissions scenario, global mean surface temperature will be 4°C warmer in the 2060s compared to the 1960s



Question: does the (mere) fact that the models agree lend additional support to the hypothesis?

Question: does the (mere) fact that the models agree lend additional support to the hypothesis?

- **Maybe yes, because these <u>different models</u> are all getting the <u>same answer</u>**

- **But maybe no, because the models are not independent from each other:**
  - Shared code (shared lineage)
  - Shared personnel
  - Model ensembles are "ensembles of opportunity" (Tebaldi and Knutti 2007)
  - Some errors and biases are common to most/all models

- **Question: does th[illegible]at the models agree lend support to the hypothesi[illegible]**

- **New question: when models converge, is there a basis for strengthened confidence (in the models or their result)?**

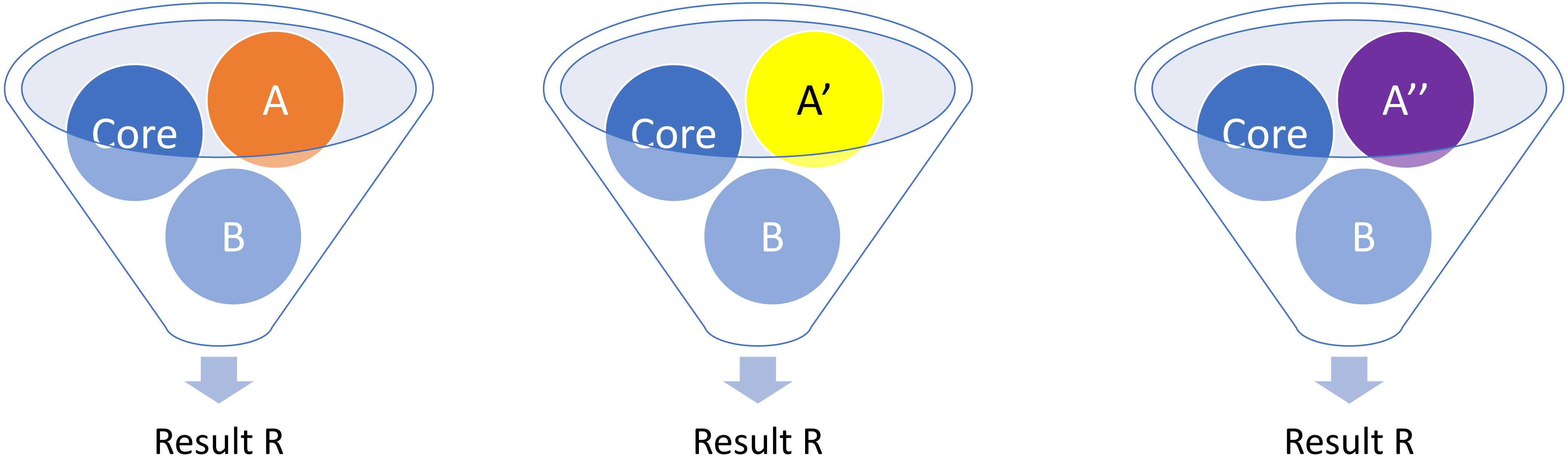
Model robustness framing:

i. Models are of a shared **model type**

ii. Models within type have some **different idealizations** with evidential support for each

iii. When models converge, there is often a **key causal process** involved, and the credibility of this causal process can be evaluated

Model robustness



Case study

LETTERS
PUBLISHED ONLINE: 11 JULY 2016 | DOI: 10.1038/NGEO2761
nature geoscience

Enhanced summer convective rainfall at Alpine high elevations in response to climate warming

Filippo Giorgi[1]*, Csaba Torma[1], Erika Coppola[1], Nikolina Ban[2], Christoph Schär[2] and Samuel Somot[3]

Model-type: regional climate models, all simulating a high-emissions scenario

Different idealizations: models used different convective parameterizations (evidential support = aircraft data, satellite data, higher resolution simulations)

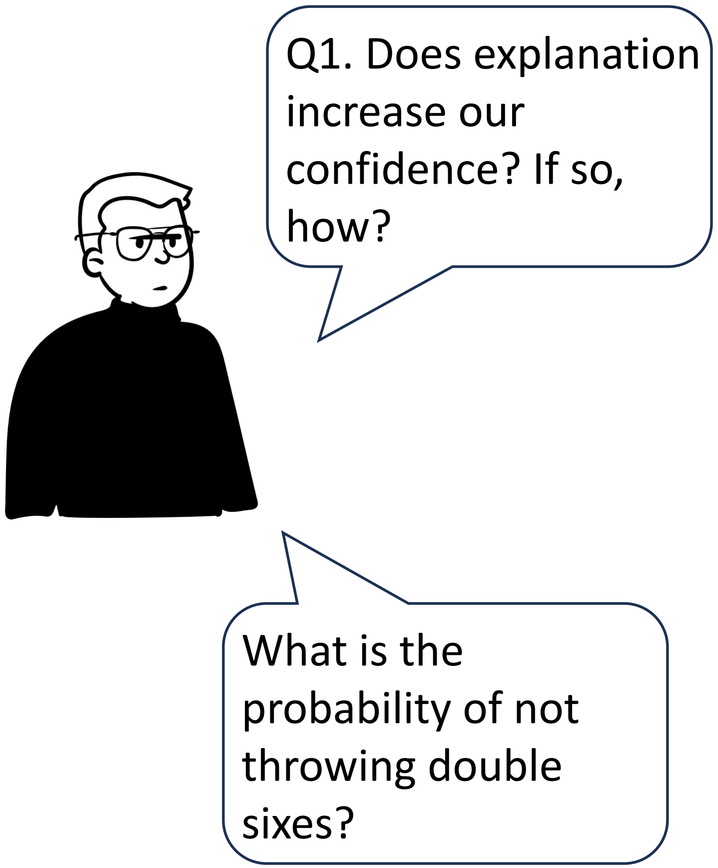
Key causal process: reduction in 500 hPa and 850hPa potential temperatures differences (plus Clausius Clapeyron), and evapotranspiration, leading to more convective precipitation; deemed credible

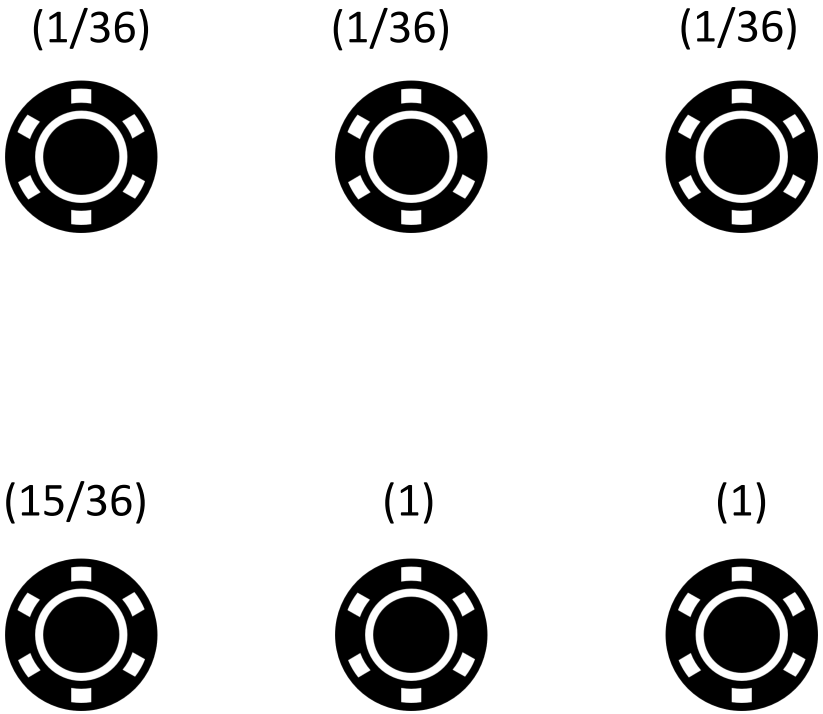
- New question: When models converge, is there a basis for strengthened confidence (in the models or their result)?

- In Giorgi et al. case: Yes, finding that each model shared the same core causal mechanism (a mechanism that is credible) **informally confirms the model-type.** This is especially so given the evidential support for the models' idealizations etc. Thus, our confidence in these models is strengthened.

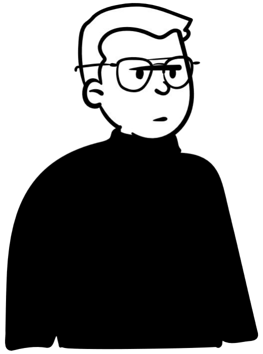
Content

- Setting the stage – which models I'm talking about
- Convergence and model robustness
- **Informal confirmation**
- Failures and Antifragility
- Example: Historical Constraints
- Conclusions and Implications



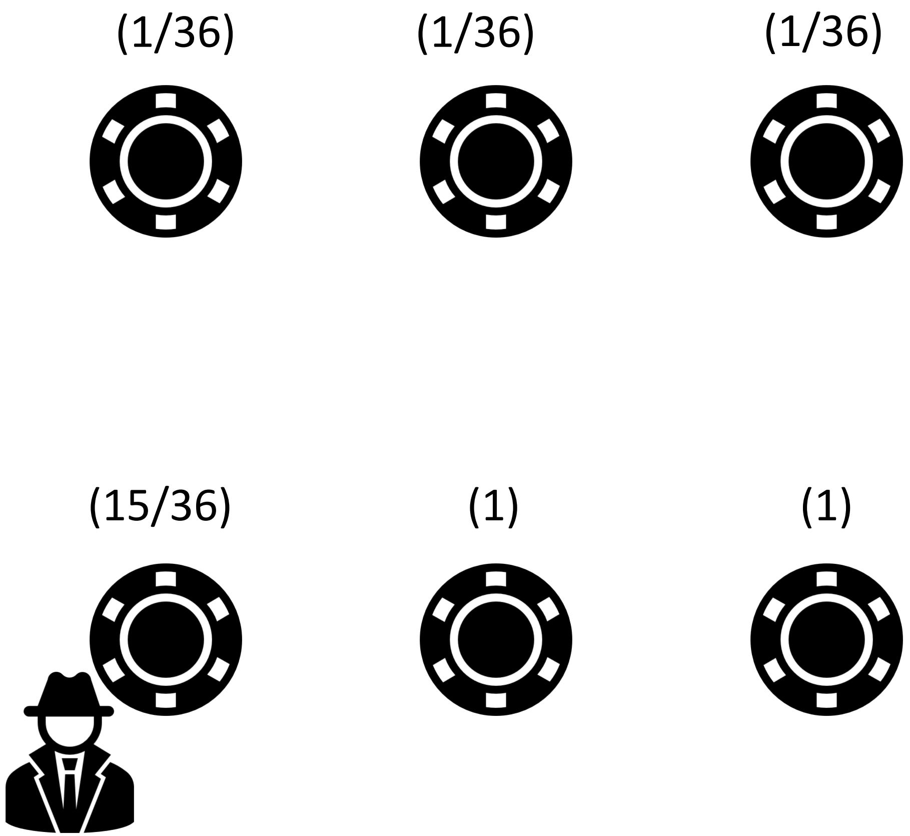


$$1-\left(\frac{1}{36}\times\frac{3}{6}+\frac{15}{36}\times\frac{1}{6}+1\times\frac{2}{6}\right)=1-\frac{15}{36}=\frac{21}{36}$$



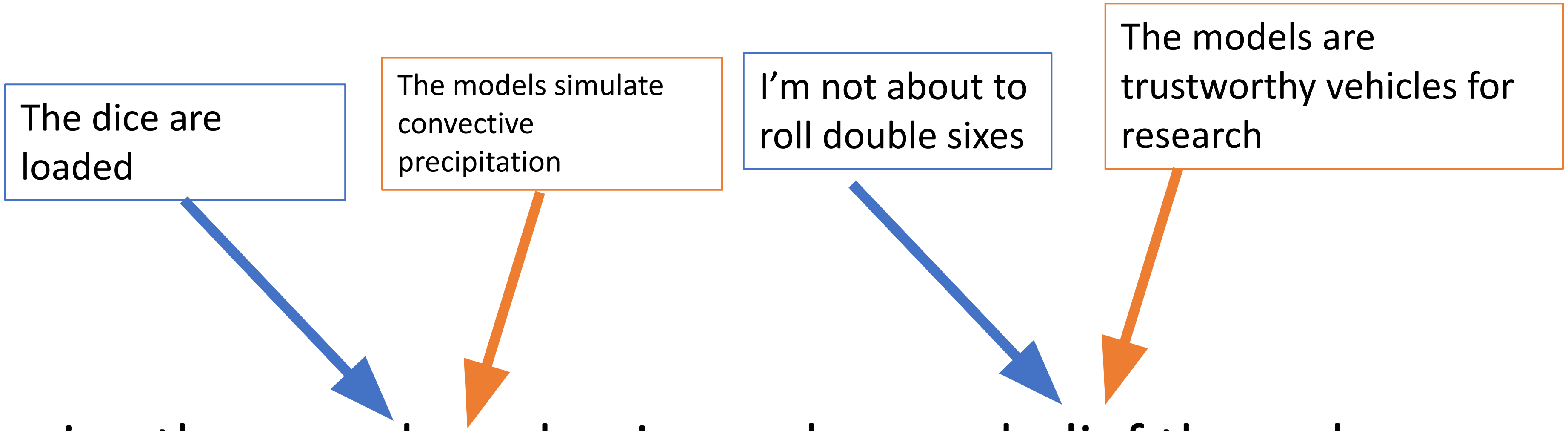
Q1. Does explanation increase our confidence? If so, how?

What is the probability of not throwing double sixes?



$$1 - \frac{15}{36} = \frac{21}{36}$$

Credit for thought experiment goes to: Gimbel, S., 1999. Peirce snatching: Towards a more pragmatic view of evidence. Erkenntnis, 51(2), pp.207-231.

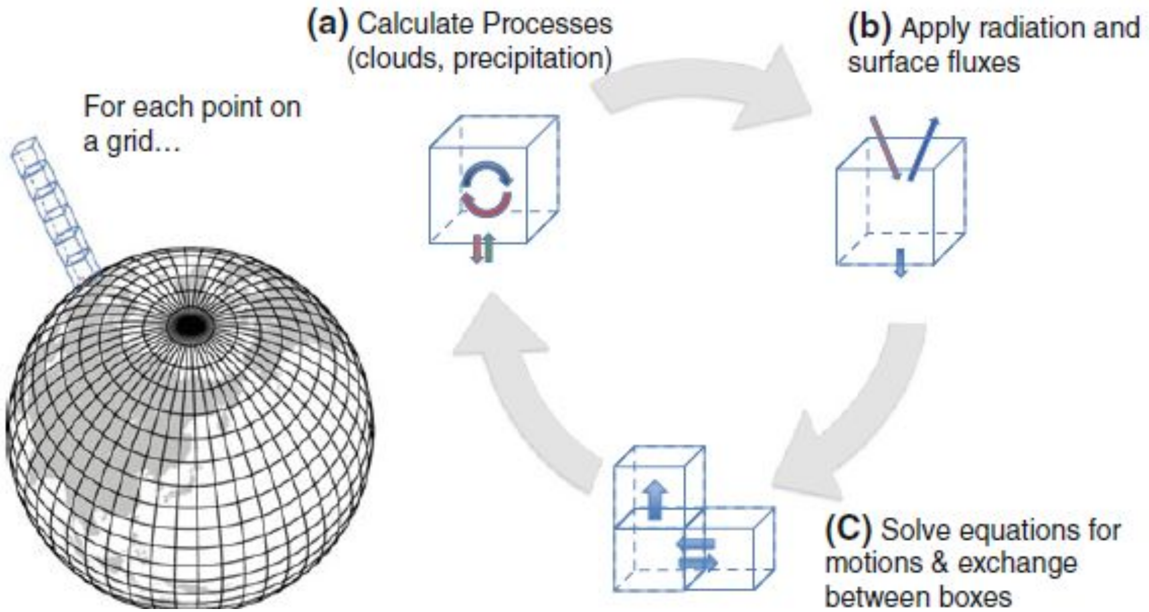


Knowing the causal mechanism enhances belief through explanation rather than probability adjustment.

Content

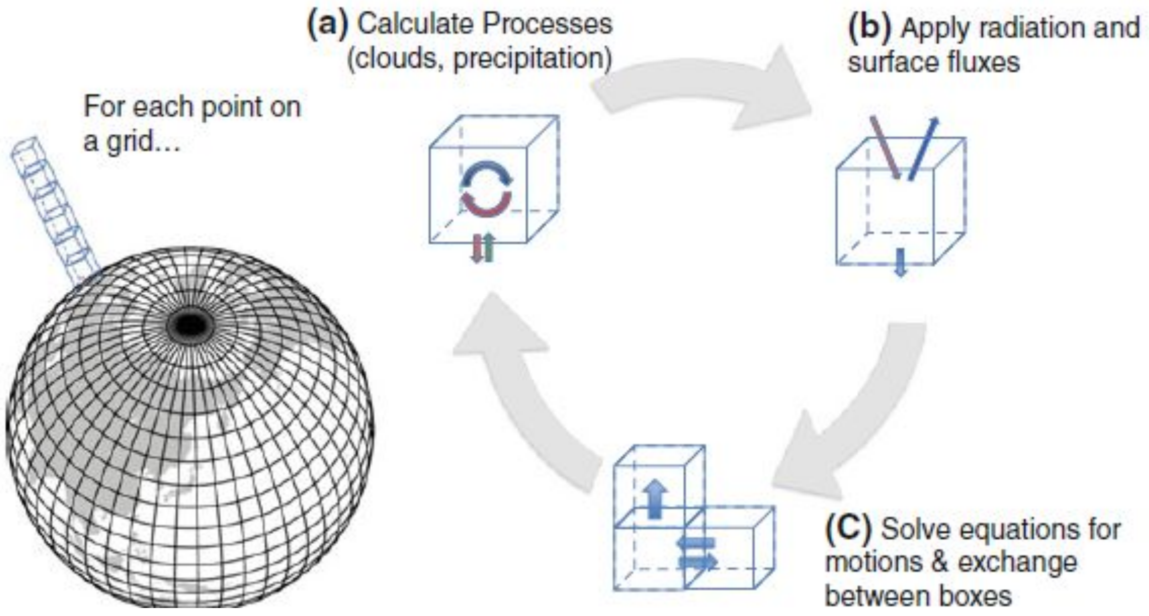
All models are wrong (George Box 1976)

- One interpretation:
  - Models *inevitably* idealize and abstract
    - Parameterizations
    - Any difference b/w model and real world
      - E.g., discrete (15 min) time-steps in simulation

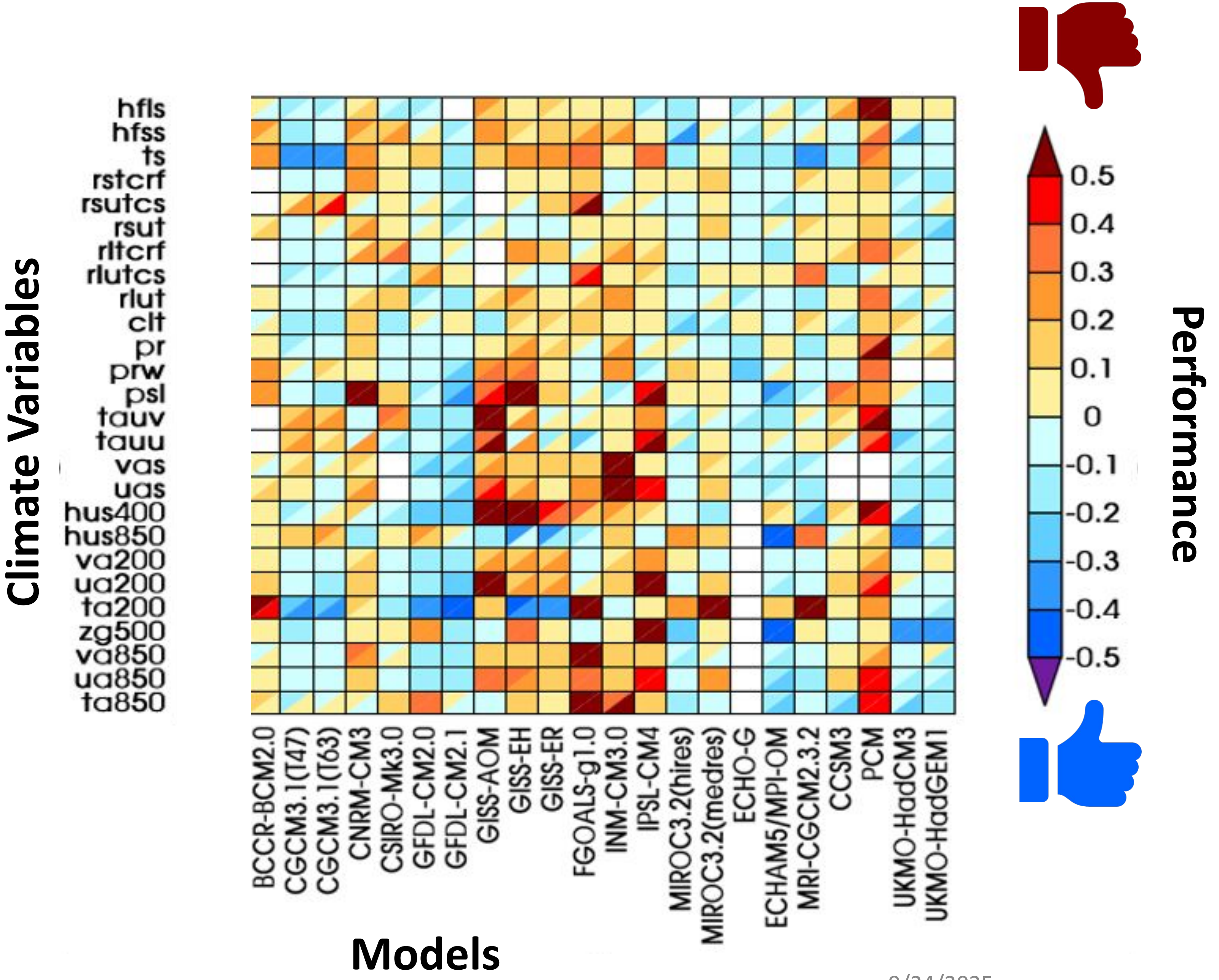


All models are wrong (George Box 1976)

- One interpretation:
  - Models *inevitably* idealize and abstract
    - Parameterizations
    - Any difference b/w model and real world
      - E.g., time-steps in simulation
- *Another* interpretation:
  - All models produce output that is inaccurate
    - Or at least: all models produces at least some output that is inaccurate

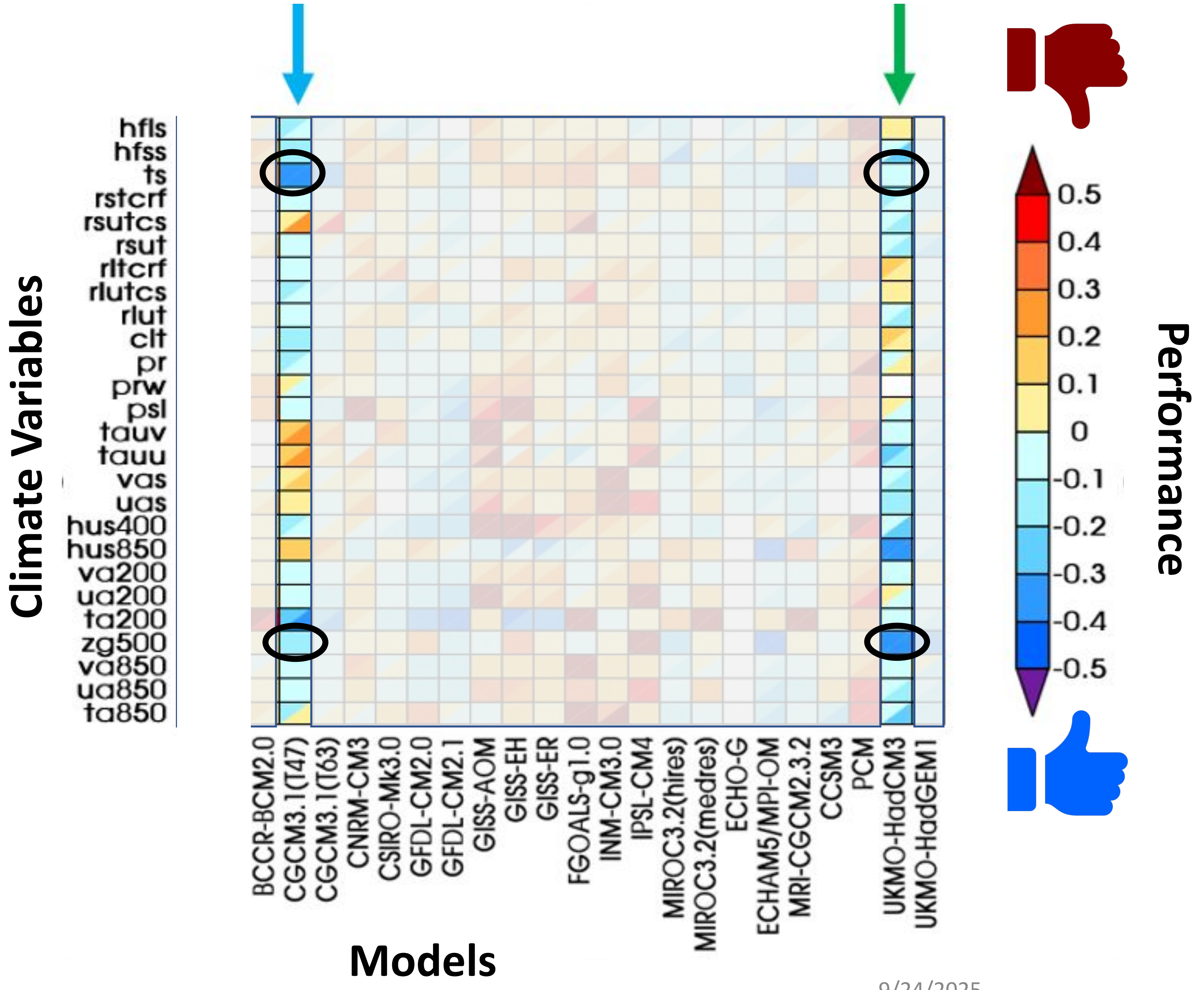


Comparing model performance

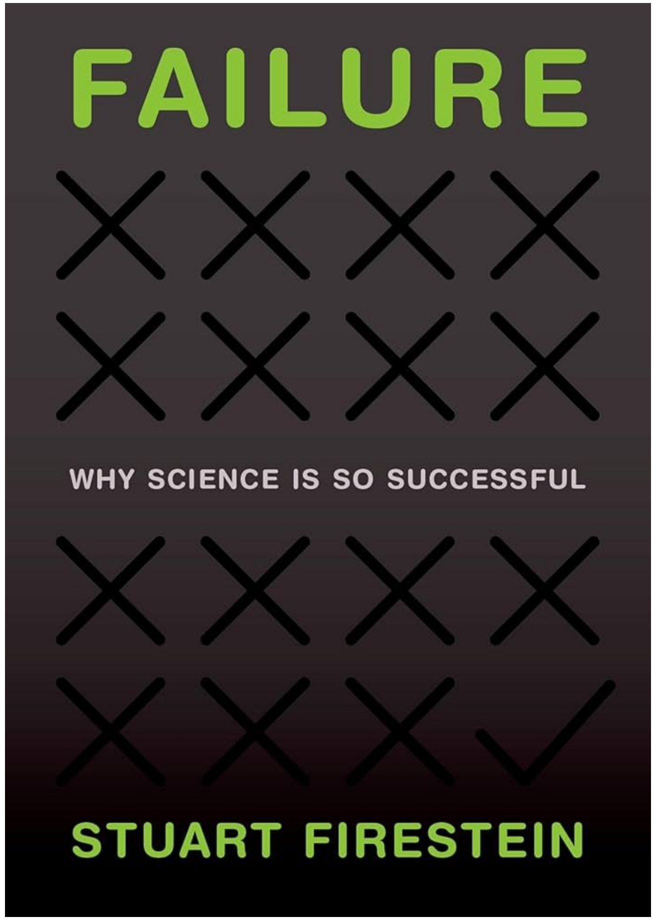


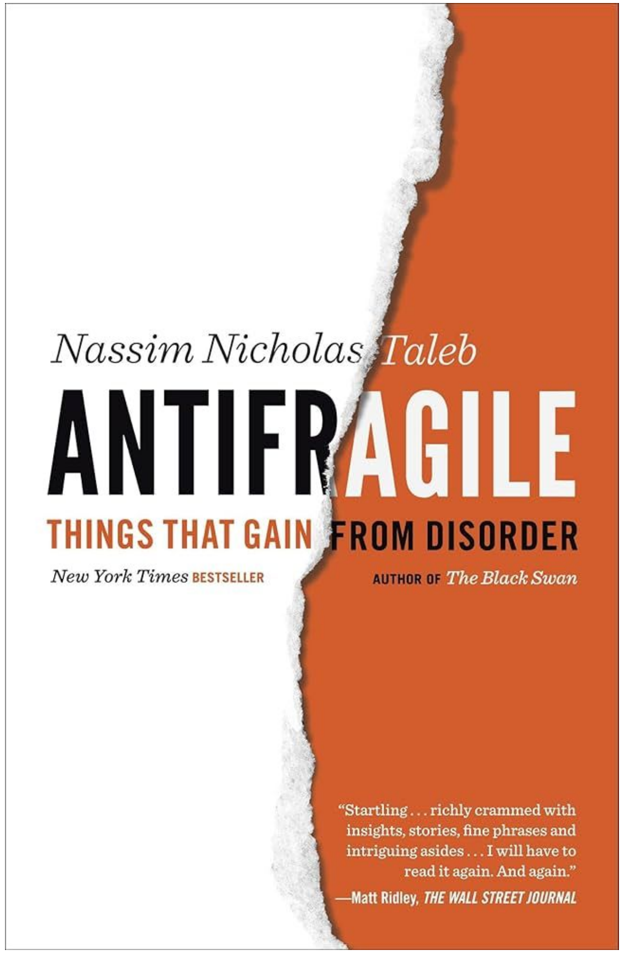
Gleckler et al. 2008

Comparing model performance



Gleckler et al. 2008

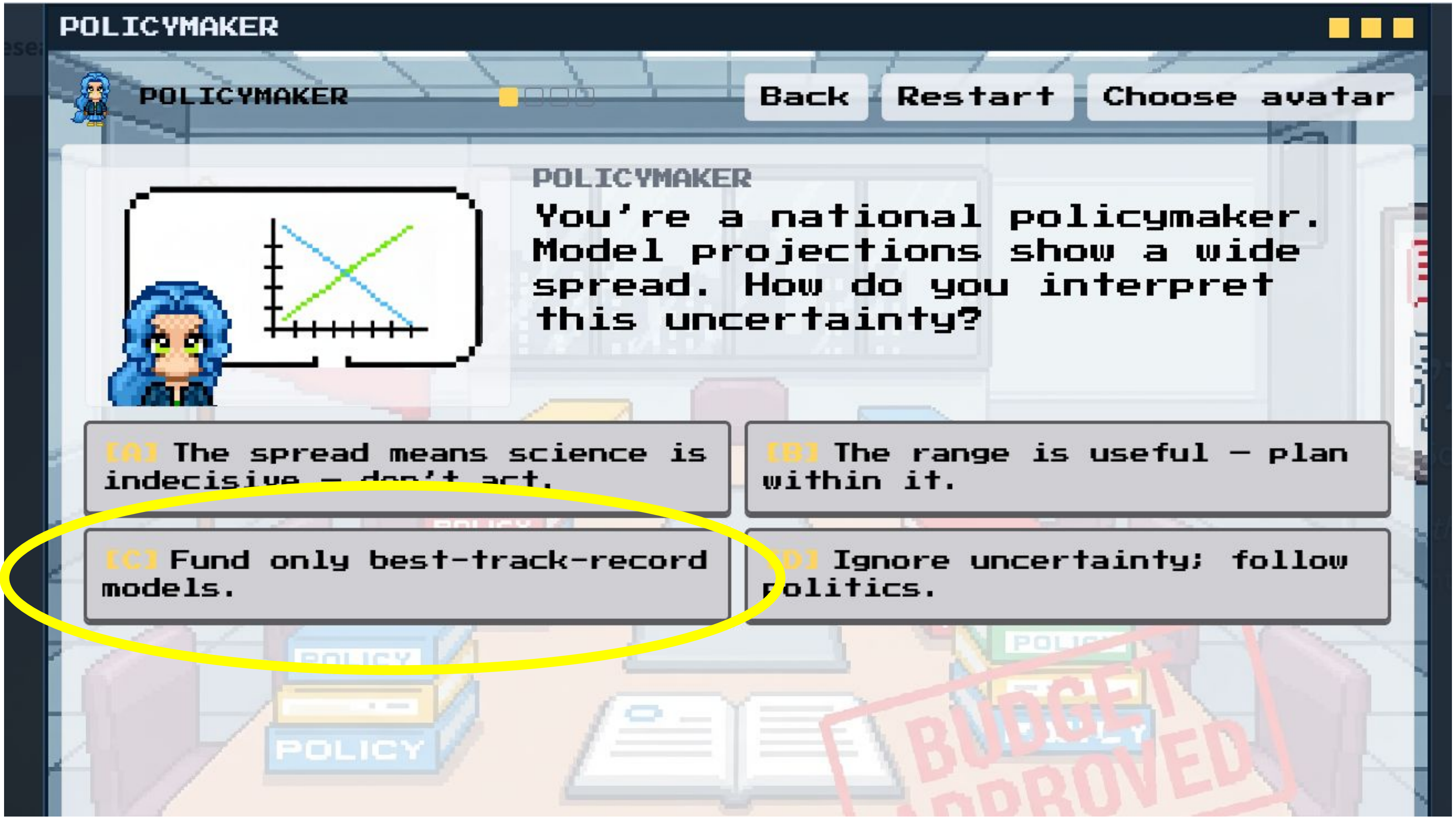

FAILURE
WHY SCIENCE IS SO SUCCESSFUL
STUART FIRESTEIN


Nassim Nicholas Taleb
ANTIFRAGILE
THINGS THAT GAIN FROM DISORDER
New York Times BESTSELLER
AUTHOR OF The Black Swan
"Startling . . . richly crammed with insights, stories, fine phrases and intriguing asides . . . I will have to read it again. And again."
—Matt Ridley, THE WALL STREET JOURNAL

- A modeling endeavor is fragile sensitive to assumptions, models unproductively disagree, conclusions are undermined
- A modeling endeavor is robust insensitive to assumptions, models with same causal core agree on many conclusions
- A modeling endeavor is antifragile sensitive to assumptions, models productively disagree, new knowledge can be produced through apparent “failures”

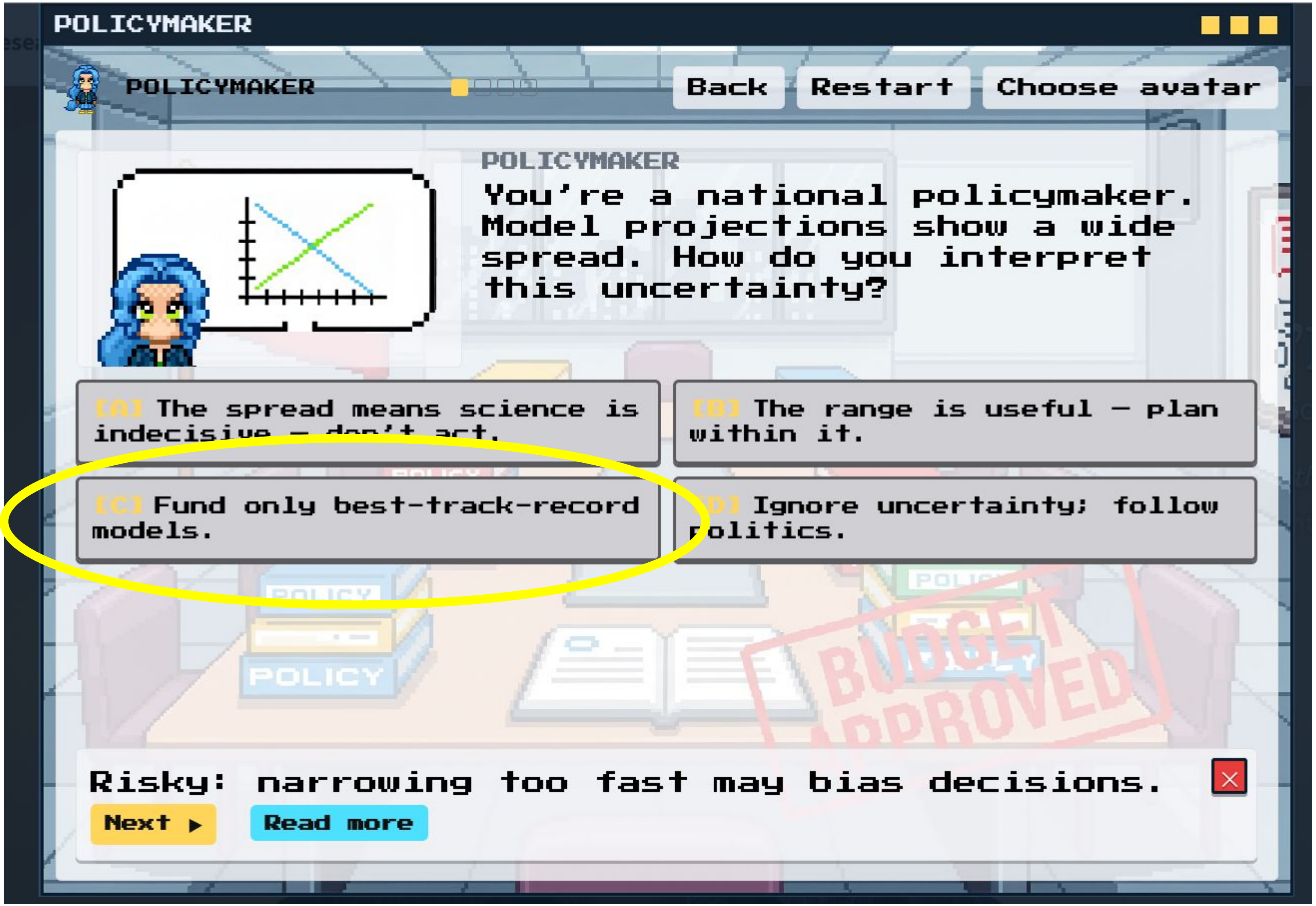
Content

- Setting the stage – which models I'm talking about
- Convergence and model robustness
- Informal confirmation
- Failures and Antifragility
- Example: Historical Constraints
- Conclusions and Implications



ryanoloughlin.org

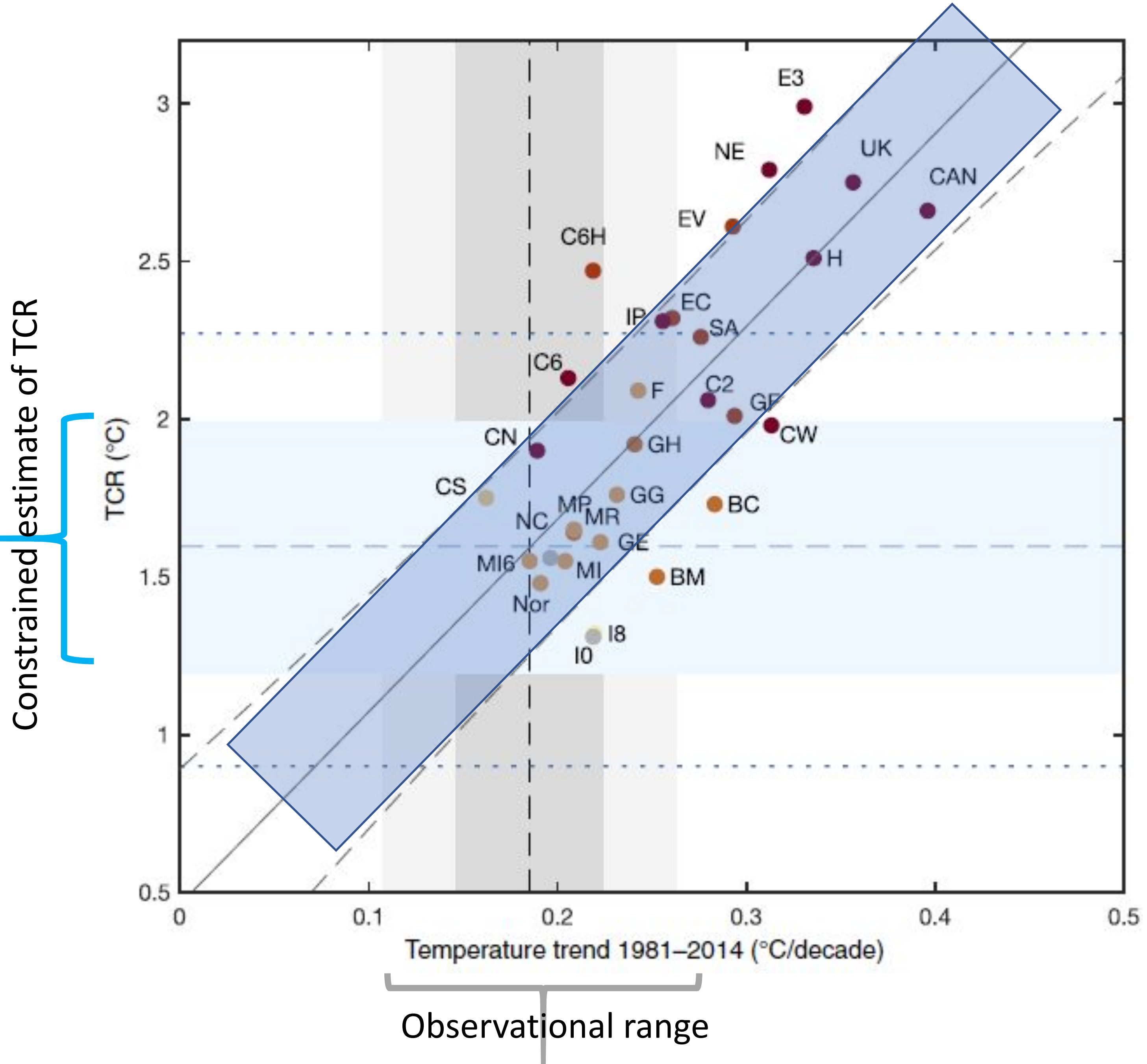
ryanoloughlin.org



Historical Constraints

Source: Tokarksa et al. 2020.

- Each dot represents a different model. Dark and light grey vertical bands show range of observational data plus error bars.
- Light blue horizontal bar represents range of emergent constraints TCR estimate.

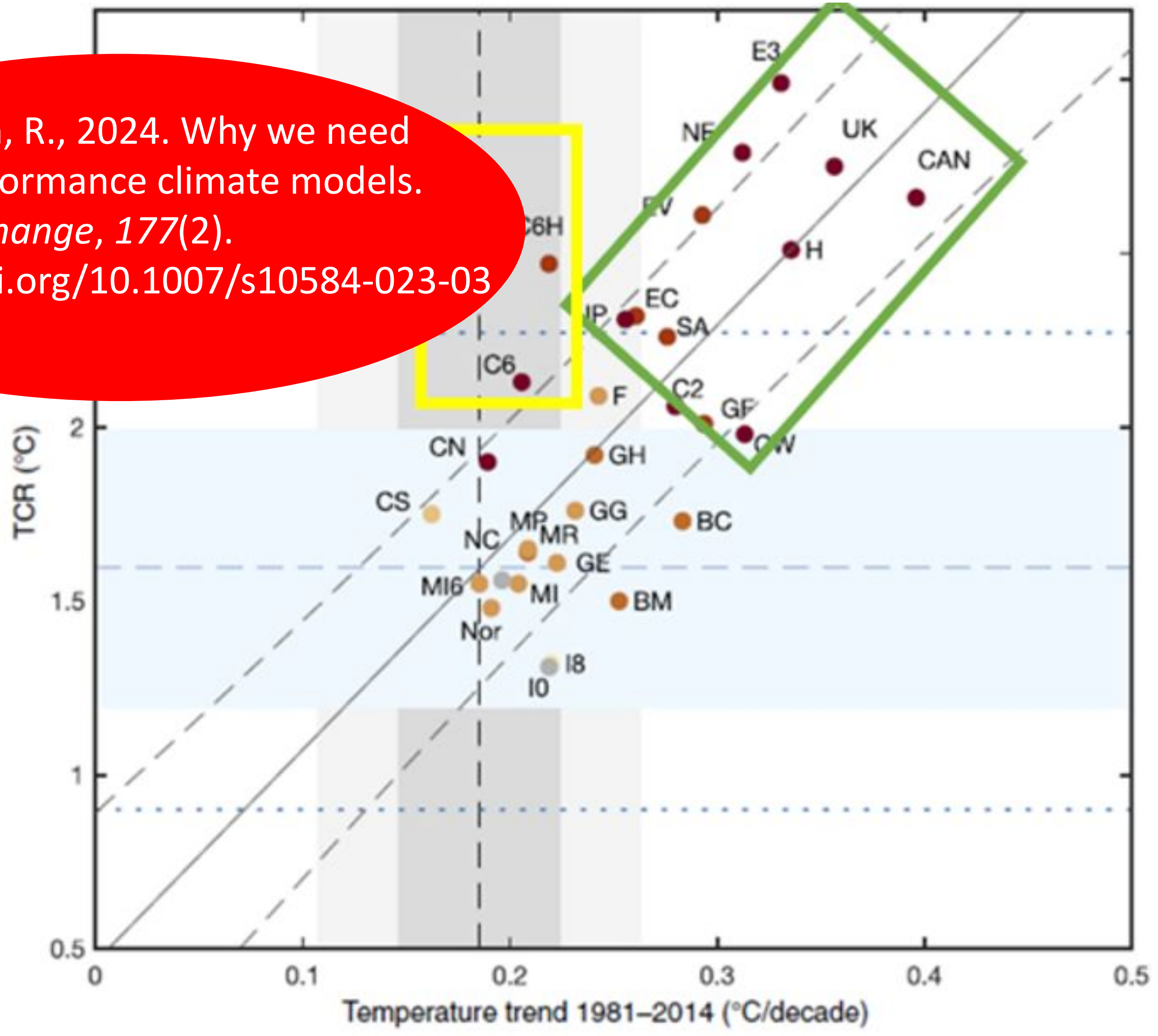


Historical Constraints

O’Loughlin, R., 2024. Why we need lower-performance climate models. *Climatic Change*, *177*(2). https://doi.org/10.1007/s10584-023-03661-7

Source: Tokarksa et al. 2020.

- Each dot represents a different model. Dark and light grey vertical bands show range of observational data plus error bars.
- Light blue horizontal bar represents range of emergent constraints TCR estimate.



Content

Key take-away points

1. *Convergence between climate models can be informally confirmatory* (model robustness)

2. *Failures in modeling can provide opportunities to yield new knowledge* (model antifragility)

Lingering questions

1. Can we give a formal account of how climate model convergence is confirmatory?

2. What are the different types of antifragility in scientific modeling and how can antifragility be enhanced in research practices?

3. Does the use of AI/ML in climate science render research **less** antifragile?

Thank you!

I look forward to any feedback and comments!

Acknowledgements: Thanks to Lisa Lloyd, Stu Gluck, Dan Li, Jonathan Baruch, Gavin Schmidt, and Wendy Parker,

Contact info:

Ryan O’Loughlin (roloughlin@qc.cuny.edu)