Metamodel.blog All posts 2022-08-02

Contents

About this blog

> *Metamodel: a model that consists of statements about models.*

This is a blog about the language, science, and philosophy of predictive modeling. The aim is to be more whimsical than polemical, and mostly non-technical. The discussions will generally be about climate models, which are arguably the most complex models ever built. But models from related fields will make a guest appearance now and then. The discussions will frequently venture beyond science because the prediction of long-term climate change is now so intertwined with the cultural milieu that it is impossible to discuss it within a purely scientific context.

As a climate scientist, I have worked with a variety of models for over three decades, ranging from very simple to highly complex. Every time you, your company, or your government plan for the future, you are relying on the output of models, whether or not you are aware of it. But models are imperfect abstractions of reality. You need to understand models to use them properly. You don't need fancy mathematics or a massive supercomputer to understand models and their predictions. This blog will try to show that we can learn a lot from simple logical reasoning using basic physical and mathematical concepts.

Climate models are the essential tools used by IPCC for assessing our climate futures to help guide mitigation and adaptation. As statistician George Box observed, “all models are wrong; but some are useful.” It is only by analyzing how models are wrong can we figure out how best to use them. This blog will critically analyze climate and other models. The purpose of the critiques is not to diminish the serious threat of climate change, but to increase the efficacy of the urgent actions needed to mitigate it.

R. Saravanan

Twitter: @RSarava   Website: r.saravanan.us   Book: The Climate Demon

PS. This blog is mirrored on Substack, if you prefer to subscribe to it as a “newsletter”.

Why a blog?

Long-form blogs seem passé in this age of short-form Twitter and Tiktok. But long-form articles are still important, because many complex issues cannot be discussed efficiently using a short format. At the other extreme, one can use the really long-format of a book to discuss the science and philosophy of modeling. But books are not free, they take time to read, and do not address current developments. Blog posts are free, relatively quick to read, and can address emerging issues.

There are a few [climate blogs]({{< relref “links.md#other-climate-modeling-blogs” >}}) that are still around, but they are less active. Perhaps because climate denial has shifted from attacking the science to attacking the solutions. But models, climate and otherwise, continue to play an important role in climate solutions. There is perhaps still a role for a blog that discusses modeling and prediction.

The website metamodel.blog is the primary home of this blog. Keeping with current trends, posts will be announced on Twitter and you may also comment on posts by replying to the “official” announcement tweet. Due to algorithmic ranking, you may not see my tweets announcing new posts even if you follow me on Twitter. Also, not everyone is on Twitter. Therefore, all posts will be mirrored on metamodel.substack.com to provide a free subscription option for those who prefer to receive posts via email.

Motivation and format

Human nature abhors a prediction vacuum. People always want to know, and often need to know, what may happen in the future. If a particular source (say, astronomy) can’t provide that information, people will tend to go to a different source that is willing to provide that information (say, astrology). Until the weather service started issuing seasonal forecasts using computer models, people relied on folksy predictions from the Old Farmer’s Almanac or groundhogs named Phil. Today, predictions from scientific models are used to make decisions that affect millions of people, often saving many billions of dollars.

Models, scientific or otherwise, will always be used to make decisions. But models are frequently misunderstood by the general public. All scientists can do is to help ensure that the most appropriate models are used and that their predictions are interpreted with the appropriate caveats.

The landscape of models is somewhat like the wild west – there's the good, the bad, and sometimes even the ugly. It's not easy for an outsider to figure out which is which because there are no clear rules. The word *model* itself can mean very different things in diverse fields such as economics, epidemiology, physics, and climate science. This often results in outsider misconceptions about how models in a particular field work.

In climate science, models are not used only for predicting the future, but also to improve our understanding of phenomena. For example, simple nonlinear models are used for qualitative understanding of amplifying climate feedbacks, such as the release of methane from melting permafrost or the increased reflection of sunlight due to melting icesheets. But such simple models aren't necessarily good at making actionable quantitative predictions. Misunderstanding the limitations of models leads to people panicking about tipping points at specific temperature thresholds or believing in prophecies of imminent doom.

Complex models, which include numerous processes, are used by IPCC and others to make quantitative predictions of the future. But these complex models do not necessarily include all the climate feedbacks that can be studied using simpler models, because we often do not have sufficient data, or powerful enough computers, to accurately represent these feedbacks. To paraphrase a famous quote, we can only predict with the most comprehensive models we have, not with even more comprehensive models that we wish to have in the future.

As society has become increasingly reliant on models for climate risk assessment, there are many important questions need to be addressed, such as: - What phenomena can be usefully predicted by models? - How well can these phenomena be predicted? - How to choose the best model(s) to use? - Do we really need to use the most complex and expensive models for a particular problem of interest?

The purpose of this blog is to provide the background information to help answer these questions. What you can expect: - new posts at roughly 2—3 week intervals - some posts on fundamental, but unresolved, climate and modeling issues - some posts on recent developments and publications - "non-technical" discussions, featuring a mix of science and pop philosophy - (guest posts on modeling-related issues are welcome)

Technical stuff

As a programmer, I like to roll my own solutions and retain creative control (at the expense of inconvenience). This blog is implemented using open-source software on a small dedicated virtual linux server. It uses a static web site generator called Hugo, with the Blist theme and Nginx as the web server. The site is designed to be mobile and social-media friendly, in keeping with the times.

Comments on blog posts are handled using Remark42, a privacy-focused open-source commenting engine. Commenting on the site requires a "social login" to avoid spam. Alternatively, you can simply reply to the "official" tweet announcing the blog post to comment on it.

Modifications were made to the Blist theme and Remark42 integration to tweak the appearance and functionality of the blog. All the custom code modifications are available on Github, but not fully documented yet.

The simple markup language Markdown is used to format all the content on the local computer. After previewing locally, the content is pushed to the linux server. Markdown is also supported for comment entry.

Blog posts are mirrored on Substack as newsletters. Simply copying and pasting the Hugo-generated web output to the Substack editor appears to work fine for posting (except for additional formatting like footnotes). This extra bit of effort allows flexibility and avoids vendor lock-in, while still having access to the popular Substack platform.

Thanks to the miracle of the universal markup converter, Pandoc, all the posts on this blog are automatically converted to an eBook using the ePUB/PDF formats. You can download and view the eBook offline. A new book is created each time a new post is added. Individual articles are also downloadable as ePUB or PDF files.

Links

- Substack mirror of this blog

Personal

- Website
- Book page
- Twitter profile

A few other climate modeling blogs

- Real Climate
- …and Then There's Physics
- Serendipity
- Bryan Lawrence's Blog
- Isaac Held's Blog

Podcasts related to climate prediction

- Deep Convection

Useful climate resources

- CarbonBrief.org
- Reporting extreme weather and climate change: A guide for journalists (WorldWeatherAttribution.org — PDF)

---

1 Can we predict global climate tipping points?

Nonlinearity generates tipping points, but it also make them hard to predict

Metamodel.blog 2022-05-03

If the globe warms slightly beyond 2°C, will we cross a climate tipping point that leads to runaway warming or catastrophe? Are we doomed if we don't stop the warming by 2030 (or 2050)? Predictions of imminent climate tipping points often capture the imagination of the media and the public. Aren't the harmful consequences of a steadily warming climate and its effect on extreme weather bad enough to spur action? Do we need even more things to worry about?

One reason to worry about new dangers is that we may need to take additional preventive action. But the solution to avoid crossing potential tipping points is exactly the same as the solution to mitigate steady climate change: reduce carbon emissions as quickly as possible to stay close to our current climate equilibrium.

Another reason to talk about potential tipping points is that it can help underscore the urgency for mitigating action. But it would be better to discuss tipping points in general terms, without implying that there are precise global warming thresholds or mitigation time intervals. Numbers associated with tipping points typically come with many caveats about the uncertainties. If the caveats are lost in translation to the public, the numbers can end up feeding into doomist narratives predicated on faux certainty.

Dystopian headlines about doomsday glaciers and methane bombs attract attention and may perhaps spur more climate activism in some people. Casual talk of climate tipping points as if they were imminent can push other people past real emotional tipping points. This can result in debilitating climate anxiety and passive sharing of "doomer memes", rather than activism.

Climate tipping points are associated with *amplifying* (or positive) feedbacks that make for a dramatic story. An example is the ice-albedo feedback, which goes like this: unusually warm conditions; more ice melts; less sunlight reflected; more heating and warmth; rinse, repeat. In a geologic instant (i.e., centuries to millennia), we end up with a hot, ice-free planet. Sounds rather scary. But surely we have had some unusually warm summers over the past several thousands of years which could have triggered this feedback. Why aren't we already ice-free?

That's because there's more happening behind the dramatic scenes of an amplifying feedback. There are *stabilizing* or negative feedbacks that act to counter it. The simplest one goes like this: unusually warm conditions; planet emits more heat; planet cools down; end of story. The stabilizing feedbacks don't garner much media attention because they are banal, but they collectively overwhelm the amplifying feedbacks and keep the climate stable. If amplifying feedbacks are swashbuckling pirates, stabilizing feedbacks are the boring navy that keeps them in check.

While our climate has been stable for the last ten thousand years, paleoclimatic data tell us that it has undergone abrupt changes before that (by geologic standards).[1] The worry then is that future global warming may disrupt the balance between amplifying and stabilizing feedbacks, resulting in an amplifying feedback that "runs away" unfettered, at least for a while until the stabilizing feedbacks catch up. Will this happen at 2°C of global warming, 3°, 5°, or beyond? The complex IPCC models suggest that the answer is "beyond", but these models aren't perfect and may not capture the slow amplifying feedbacks well. We can build simplified models to understand the amplifying feedbacks that generate tipping points, but these simpler models may not capture all the stabilizing feedbacks accurately. This precludes attaching specific numeric global thresholds or dates to climate tipping points that may lie in our future.

We know there are absolute local temperature thresholds that are relevant to current and future climates. An important one has to with the human body. A metric called wet-bulb temperature, that combines temperature and humidity, is used as a measure of heat stress on humans. Extended periods with wet-bulb temperatures exceeding about 35°C would be intolerable for humans. (The wet-bulb temperature threshold is lower than the normal human body temperature of 37°C because the body cools itself by sweating and transferring heat to cooler surroundings.)

Human society has adapted to a certain range of temperatures and departing from these temperatures causes harmful impacts. Some regions of the global are closer to the absolute wet-bulb temperature threshold than others, and the anthropogenic warming itself varies regionally. Therefore, the relative warming thresholds for harmful impacts will vary with the region. There are also other region-specific temperature thresholds that affect agricultural and ecological systems. For example, corals are very sensitive to the ambient temperatures. Ice sheets and permafrost also respond to regional temperatures.

Will continued regional warming cause the climate to soon cross a global tipping point? Nonlinearity in the climate system is often touted as a reason to be concerned about tipping points, because a nonlinear system can potentially switch between multiple equilibrium states. But nonlinearity is a double-edged sword: it adds interesting threshold behavior to a system, but it also takes away predictability. As Edward Lorenz showed using a simple model of deterministic chaos, nonlinear error growth can lead to rapid

[1] A Physicist Explains Why Parallel Universes May Exist (NPR.org)

loss of predictive skill. Chaos associated with fast processes like weather reaches saturation for climate prediction, and can be quantified as stochastic noise or "certain uncertainty".[2] But this does not apply to slow climate processes like melting ice sheets and thawing permafrost, which are in the realm of "uncertain uncertainty". The initial conditions and the governing equations associated with these slow processes are poorly known. This means that nonlinear error growth will make it hard to accurately predict if and when any tipping points associated with these processes will be crossed.

Nonlinearity also prevents us from aggregating different local warming thresholds to come up with a single global warming threshold. Local thresholds associated with amplifying feedbacks can be studied using relatively simple models, but to answer the global question, we must use comprehensive global climate models. These models compute the combined impact of many different regional processes. When we add together many different nonlinearities in a complex system, the different nonlinear transitions can get smeared out, making the global system respond in a "near-linear" fashion with increasing emissions. This can help explain why the IPCC models do not predict that we will cross any tipping point soon, even as they predict that global warming and its impacts will get much worse without mitigation.[3]

Consider global average surface temperature, which often figures in discussions of tipping points. It is the most commonly used measure to characterize climate change, although it may not be the scientifically most discerning metric. Models can estimate the relative trend in the global temperature with fairly good accuracy to simulate the observed warming (Figure 1, line). But the errors in the absolute global average temperature in model simulations are rather high (Figure 1, sidebar). Among different global climate models, the absolute global average temperature can range between 13°C and 15°C. For one model, 2°C warming means warming globally from 13°C to 15°C, whereas for another model, it means warming from 15°C to 17°C. Since a global 2°C warming translates into different local warming for different models and different regions, it is not possible to identify a hard global warming threshold for catastrophic impacts using current models. All we can say is that if the globe continues to warm, the risk of catastrophic local damage will increase rapidly.

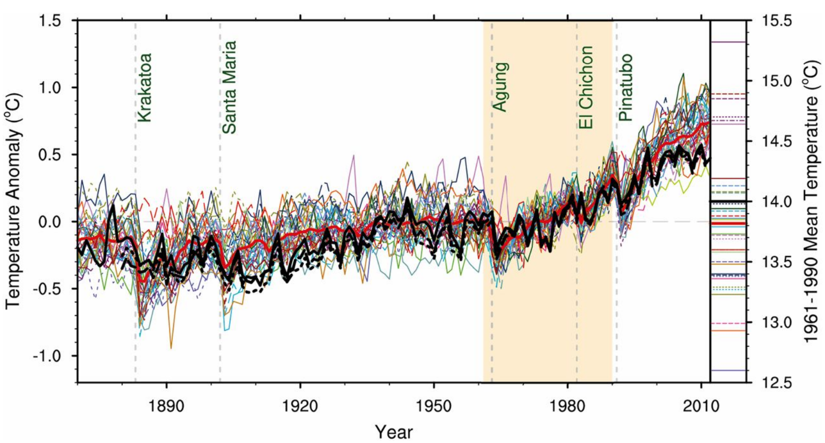


Figure 1. *Estimates of global-average surface temperature anomaly from model integrations performed in support of the fifth phase of the coupled model intercomparison project (CMIP5). Here, the*

[2]In normal climate terminology, we refer to the multiverse as the *ensemble*. We refer to individual universes as *members* of the ensemble. When the climate isn't changing, the time average is equivalent to the ensemble average, according to the ergodic hypothesis. In a changing climate, that is no longer the case.

[3]A Physicist Explains Why Parallel Universes May Exist (NPR.org)

observed anomaly (black) is estimated relative to the observed absolute time-average, while the model anomalies are estimated relative to each model's absolute time-average. Colored lines represent different models, with thick red denoting the model average, and the vertical dashed line denote volcanic eruptions. The small bar to the right of the figure shows the range of absolute global and time averaged model temperatures for the period 1961 to 1990. From Palmer and Stevens (2019)

Global climate models do not predict a climate cliff's edge located at specific numbers like 1.5 or 2°C of warming, or by specific dates. But the higher levels of global warming predicted for unmitigated emissions can lead to unbearably harsh weather and climate in many regions, even without crossing any tipping points. Climate harm is more likely to occur by a thousand cuts rather than in one fell swoop. Any planetary warming threshold for tipping points that we can identify will be fuzzy. Does that mean we should worry less about exceeding 2°C global warming, because the local thresholds may be further away than we think? Not quite. A fuzzier global threshold also means that local thresholds for harmful impacts may be closer than we think. So, we need to act as quickly as we can to eliminate carbon emissions.

(*Top image adapted from the poster for Pirates of the Caribbean: At World's End, using the Pieces of Eight font for the overlay text.*)

1.1 Related articles

- Why are the (climate) numbers so round? (Metamodel blog)
- Debate about communicating tipping points (And Then There's Physics)
- Runaway tipping points of no return (RealClimate.org)
- Superrotation, idealized models, and GCMs (Isaac Held)

1.2 Comments

Note: For updated comments, see the original blog post and the anouncement tweet.

- *Paul Pukite*:
  There are scales of non-linearity.
  - *R Saravanan*:
    True. The more nonlinear a system is, the more likely it will be to exhibit tipping point behavior, and the more difficult it will be to predict those tipping points (due to strong error growth).

---

2 How to judge a model beauty contest?

Model meritocracy is a good idea, but the devil is in the details

Metamodel.blog 2022-05-10

Every 6–7 years, major climate modeling centers around the world submit their climate simulations to an organization called the Coupled Model Intercomparison Project (CMIP). CMIP distributes the simulation data so that scientists around the world can analyze and compare the models. But what criteria should we use to judge or rank models? How do you decide whether one model is better than another? Do we care about superficial beauty or inner beauty? These questions raise fundamental issues relating to climate modeling.

A useful analogy that distances these issues from climate jargon is the college application process. How do you judge a college applicant? In many Asian countries, the numeric score on a single exam decides which colleges you can get into, effectively determining your whole career. In the U.S., college applicants submit grade point averages and (increasingly optional) test scores, along with essays and a resume of extracurricular activities. A "holistic" process weighing all this information is used to make admissions decisions in selective colleges. Often selective colleges receive many more qualified applicants than they can admit. Teachers and coaching companies "teach to the test" to help the students get ahead. One proposed solution to reduce the intense competitive stress is to first identify all applicants who pass an acceptability threshold and then use a lottery to select those who are admitted. So putting a lot of effort into obtaining scores above some threshold (or even perfect scores) would not really help.

Consider, on the other hand, the stress-free process of attending a non-selective local college – say the Iowa Public Community College (IPCC) – that has a tradition of admitting all applicants. But one year, IPCC finds that some of the applicants have unusually low test scores (or be it grade point averages), even though they have good extracurricular activities on their resume. To deal with this, IPCC decides to suddenly become selective and notifies the applicants that only those with numeric scores above a threshold value will be admitted.

Although some current applicants may be miffed about the goalposts being moved after the ball has been kicked, the IPCC's decision may be acceptable as a short-term solution to maintain academic standards (to the extent measured by the numeric scores). But what are the long-term implications? Future applicants to IPCC may start to focus on improving the numbers that the college cares about, to the exclusion of factors like extracurricular activities that make them well-rounded. The newly selective community college should think long and hard before finalizing its new admissions policy

This is sort of the situation with the real IPCC, the Intergovernmental Panel on Climate Change, which uses the CMIP models for its assessments. In previous rounds of CMIP, explicit ranking of submitted models was not needed: Climate predictions were averaged among all the submitted models for assessment purposes, treating them equally. In the latest round of CMIP, the IPCC found that some of the submitted models were "running too hot", i.e., simulating too much warming in recent years, even if they were better in some other respects, like simulating regional climate features better. (The rationale for deciding what is "too hot" deserves its own discussion, but we'll just accept it for now.) If some models are running too hot, it will skew the average to be overly hot as well, resulting in "overprediction" of future warming when analyzing impacts.

To address this problem, the IPCC took a simple, if somewhat ad hoc, approach. Different models were weighted differently for averaging, based on how well they simulated the recent observed warming.[4] The models that overpredicted the recent warming were weighted less compared to the rest of the models.

Even after the IPCC report was released, many studies have continued to average across all the CMIP models equally, out of habit and due to convenience. A recent Comment[5] in Nature draws attention to this lack of awareness. The Comment reiterates that the models "running hot" should be downweighted when averaging. The issue is framed as "meritocracy" versus "democracy":[6] Treating all models as equal would mean a democracy, but assigning higher weights to the better models would be a meritocracy.

Computing and using model weights, as done in the IPCC assessment, can be a complicated process. What end users usually want is a simple recipe. Neither the IPCC nor the Nature Comment provide such a recipe, but a follow-up article[7] by the authors of the Comment suggests an alternative: *screening out models whose transient climate response (TCR) lies outside the likely (66% likelihood range) of 1.4C to 2.2C*. (TCR is the expected warming of global average temperature when the slowly increasing carbon dioxide concentration reaches double its value.)

The simple screening criterion is acceptable as a stopgap measure, as a practical "band aid" to fix an unexpected problem. But philosophically, it is a worrying development and should not be the long-term solution. It does not really address the hard question of why the physics-based models are "running too hot". The TCR-based screening crite-

[4] A Physicist Explains Why Parallel Universes May Exist (NPR.org)

[5] In normal climate terminology, we refer to the multiverse as the *ensemble*. We refer to individual universes as *members* of the ensemble. When the climate isn't changing, the time average is equivalent to the ensemble average, according to the ergodic hypothesis. In a changing climate, that is no longer the case.

[6] 14 movies and shows about the multiverse, from 'Spider-Man: No Way Home' to 'Everything Everywhere All at Once' (BusinessInsider.com)

[7] Three reasons why climate change models are our best hope for understanding the future (TheConversation.com)

rion goes further than the IPCC weighting approach by imposing a statistical constraint on predictions from physics-based models. (The IPCC approach uses model simulations of recent warming to compute the model weights.) Essentially the physics-based global climate models are no longer predicting global-average temperature, but merely serve to add regional climate detail to the statistically constrained global-average temperature prediction (a procedure referred to as dynamical downscaling).

There is the danger that a simple recipe like the TCR-screening could become the de facto metric for distinguishing "good models" from "bad models" in the world of model meritocracy. Like college applicants, model developers are Pavlovian. They will respond to behavioral incentives to develop "good models" and the climate science community should be careful to provide the right incentives. Established metrics are hard to dislodge even if they become counterproductive.[8] Hence this longish blog post.

Evaluating climate models

Much of our intuition about evaluating predictive models comes from simulations of precedented events occurring in relatively simple models. Global warming is an unprecedented event occurring in a highly complex system with many interacting components. By definition, our past prediction experience will be of limited use in characterizing unverifiable long-term predictions of an unprecedented event. We will need to reason from basic scientific principles to understand how best to do that. Here are some issues to consider:

Global-average-temperature-centric thinking: Global average temperature is an important and useful measure to study climate change, but it is not the only metric that's important. Climate impacts are determined by regional temperatures and rainfall, not the global average temperature. For example, a model could overestimate warming in the Northern Hemisphere and underestimate it in the Southern Hemisphere, but still end up with a small error in the global average. Such a model would be less useful than one that had the same global error uniformly, but would be weighted the same by a global-average metric. Similarly, a model that simulates the trends temperature well but not the trends in rainfall would also be less useful.

Model tuning and linear thinking: A climate model operates on a fairly coarse spatial grid, typically about 100x100 km (60x60 miles) in the horizontal, which cannot represent important processes like cloud formation. Approximate formulas, known as parameterizations, are used to represent clouds in models. The parameterizations have coefficients that are adjusted to make the simulations better fit observations – a process known as model tuning.[9] Often tuning is done explicitly, with varying degrees of effort and success, but sometimes it is implicit in the history of the modeling effort.[10]

It is commonly assumed that a model that simulates the recent observed global warming trend better should also be trusted make a more reliable prediction of the future trend. Strictly speaking, that is only true for linear models, where nonlinear interactions among different components are not important. Prediction models used in many fields, such as regression models, fall into this category of linear models. For nonlinear models that have been tuned to simulate spatially averaged quantities, there is an ambiguity when using the same averaged quantities for validation. We cannot be sure

[8] How we make our 2050 'forecasts', and why we do them (Uk Met Office)

[9] In 2020, the @metoffice produced a hypothetical weather forecast for 23 July 2050 based on UK climate projections. Today, the forecast for Tuesday is shockingly almost identical for large parts of the country. [Tweet by @SimonLeeWx](https://twitter.com/SimonLeeWx/status/1547957062000267267) (Twitter)

[10] Ch.2, The Climate Demon: Past, Present, and Future of Climate Prediction (ClimateDemon.com)

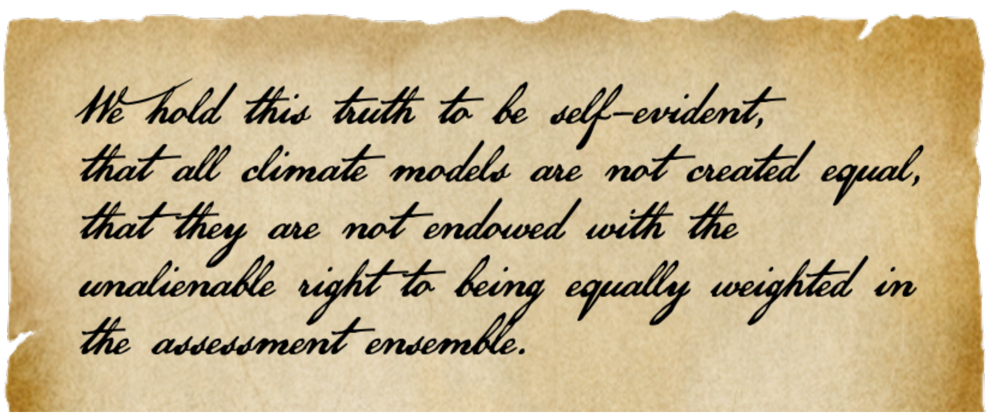
whether we are validating the fundamental accuracy of the representation of processes like clouds, or simply validating the efficacy of the tuning process.

Tuning is often described as model calibration. In simple models with few adjustable coefficients, the tuning process can estimate the “best” values of the unknown coefficients for each process, thus calibrating the model. In a complex nonlinear system with many adjustable coefficients, coefficients for one process may end up getting adjusted to cancel errors associated with a different process. Instead of calibration, we get compensating errors. The more averaged a tuning target is, the worse this problem.[11]

Consider a climate model with a poor cloud parameterization. This parameterization can have many well-adjusted coefficients tuned to compensate for errors in other components, enabling the climate model to simulate recent short-term warming well.[12] This simulation may even appear better than one using a more scientifically-sound, but less adjustable, cloud parameterization. But the long-term climate prediction using the poor parameterization can become less reliable, because the error compensation provided by the tuning is not guaranteed to be valid in a different climate.

Model tuning can definitely be beneficial in improving the fidelity of short-term (multidecadal) warming predictions of a model. But being able to tune parameterizations to *adequately simulate* recent warming should be considered a *necessary condition* for a good model rather than a *sufficient condition*.[13] There needs to be enough wiggle room in the definition of “adequately simulate” to allow a model with better parameterizations, but less successful tuning, to be considered acceptable.

Declaration of Meritocracy



With the increase in the number and complexity of climate models, the spread in their predictions has increased. Therefore, it makes sense to validate them carefully before using them for climate assessments. By assigning weights to model in AR6, the IPCC has thrown down the gauntlet on the notion of model democracy or treating all models as equal. [14] How do we transition to a model meritocracy?

It is easy to find fault with the scalar weighting metric used by the IPCC, but it will require a lot of constructive discussion to come up with more general merit criteria for

[11] How large does a large ensemble need to be? (S. Milinski et al., 2020; Earth System Dynamics)

[12] What to expect when you’re expecting a better climate model (Metamodel.blog)

[13] Insights from Earth system model initial-condition large ensembles and future prospects. (C. Deser et al., 2020; Nature Climate Change)

[14] Without human-caused climate change temperatures of 40°C in the UK would have been extremely unlikely (WorldWeatherAttribution.org)

models. It will be a challenge to keep the merit criteria simple enough for wide adoption but at the same time comprehensive enough to cover important aspects of the model. One option is a multifaceted threshold approach, where minimum benchmarks must be met in multiple metrics for a model to be considered acceptable. This may be better than a weighting approach because it won't incentivize overtuning (or overfitting) a model.

To return to the college admissions analogy, "teaching to the test" would be more acceptable if the test were broad enough and evaluated a range of skills. Using a single metric for assessing merit – like the ability to simulate the recent warming trend in global average temperature – is rather like buying a used car after a short test drive without looking under the hood. A well-tuned car will drive more smoothly, but will it also be reliable in the long haul? A thorough validation would require a mechanic to check engine parts under the hood of the car. A car that rattles a bit more during the test drive could still turn out to have more reliable parts under the hood and make for a better long-term purchase.

2.1 Comments

Note: For updated comments, see the original blog post and the anouncement tweet.

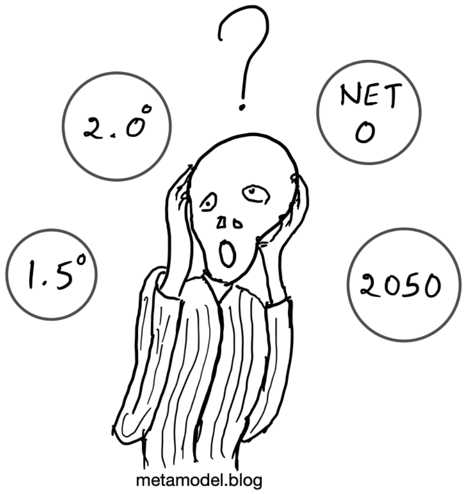
- *R Saravanan*:
  Not surprisingly, this post has attracted the attention of some who do not consider climate change to be a serious threat. Here's what the About page of this blog says:
  Climate models are the essential tools used by IPCC to assess our climate futures and guide mitigation and adaptation. As a climate scientist who has worked with many different models over decades, I am keenly aware of their strengths as well as their limitations. This blog will critically examine climate and other models. The purpose is not to diminish the seriousness of the threat of climate change, but to increase the efficacy of the urgent actions needed to mitigate it.
- *Jim White*:
  As The Yogi said: «Predicting is very hard, especially about the future.»

---

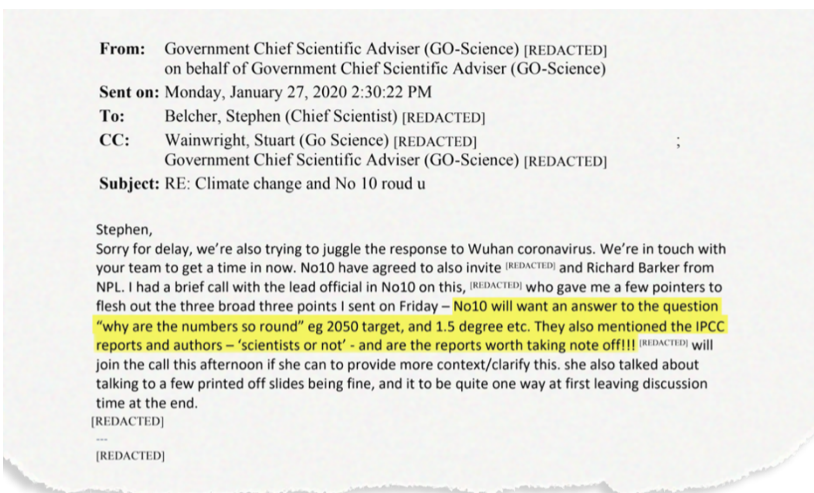
3 Why are the (climate) numbers so round?

Climate target numbers are approximate. Their roundness reflects that.

Metamodel.blog 2022-05-24



Note to non-UK readers: *No. 10, Downing Street is the official residence of the UK Prime Minister (like 1600 Pennsylvania Avenue for the US President). No. 11, Downing Street is the official residence of the Chancellor of the Exchequer, often considered the second-most powerful position (rather like the Vice President in the US, but with specific responsibilities similar to the Treasury Secretary).*



From: Government Chief Scientific Adviser (GO-Science) [REDACTED] on behalf of Government Chief Scientific Adviser (GO-Science)
Sent on: Monday, January 27, 2020 2:30:22 PM
To: Belcher, Stephen (Chief Scientist) [REDACTED]
CC: Wainwright, Stuart (Go Science) [REDACTED] ; Government Chief Scientific Adviser (GO-Science) [REDACTED]
Subject: RE: Climate change and No 10 roud u

Stephen,
Sorry for delay, we're also trying to juggle the response to Wuhan coronavirus. We're in touch with your team to get a time in now. No10 have agreed to also invite [REDACTED] and Richard Barker from NPL. I had a brief call with the lead official in No10 on this, [REDACTED] who gave me a few pointers to flesh out the three broad three points I sent on Friday – No10 will want an answer to the question "why are the numbers so round" eg 2050 target, and 1.5 degree etc. They also mentioned the IPCC reports and authors – 'scientists or not' - and are the reports worth taking note off!!! [REDACTED] will join the call this afternoon if she can to provide more context/clarify this. she also talked about talking to a few printed off slides being fine, and it to be quite one way at first leaving discussion time at the end.
[REDACTED]

[REDACTED]

Figure 1. From *The 11 slides that finally convinced Boris Johnson about global warming* CarbonBrief.org

From: Metamodel.blog
To: No. 10, Downing Street, London
Date: Tuesday, May 24, 2022
Subject: Why are the numbers so round?

Dear No. 10,

I read with interest the article about the scientific briefing on climate that changed your mind about global warming.[15] The briefing underscores the importance of making science accessible to decision makers. One email leading up to that pivotal briefing

[15] A Physicist Explains Why Parallel Universes May Exist (NPR.org)

includes an interesting nugget of a question (Figure 1). You asked, “why are the numbers so round”, referring to the 2050 target year, 1.5 degree warming, etc.? That is indeed an excellent question, appropriately coming from someone who lives in a house numbered 10.

Like a politician who wants to move into 10 Downing Street but ends up at 11 instead, shooting past round numerical targets has been the subject of much discussion after the recently released IPCC climate report. Are we likely to overshoot the global warming target of 1.5°C or the net-zero target date of 2050? If the warming ends up being 1.6 or 1.7°C or net-zero is reached 5-10 years later, would climate cross a tipping point?[16] As there appears to be another UK climate change briefing in the offing,[17] this letter attempts to explain the “roundness” of climate numbers.

The sentence in that email snippet above that follows your “roundness” question suggests you are not sure if actual scientists wrote the IPCC reports. Rest assured that scientists were involved in writing the IPCC reports and coming up with the climate numbers. The roundness of the numbers is itself perhaps proof of that. In business, it is considered good practice to add false precision as a negotiating tactic: “if one party gives a round number, it gives the signal that the party doesn’t really know what it’s doing.”[18] For example, faux precision is one explanation why Elon Musk offered to buy Twitter at $54.20 a share, instead of $54 or $55 a share.[19] But scientists aren’t businessmen. When scientific thresholds are approximate, it is normal to round the numbers up or down, to avoid giving a sense of false precision and to make them memorable.

Scientists have determined that the trapping of heat by increasing concentration of carbon dioxide and other atmospheric gases (known as greenhouse gases) is responsible for the global warming that is happening.[20] Burning of fossil fuels used in transportation, power generation, and other human activities is increasing the concentration of these gases. As the globe warms, the climate changes from what we are used to, leading to harmful impacts like more heatwaves, intense rainfall events, and rising sea levels. We need to reduce the emission of the greenhouse gases to zero as soon as possible to stop further warming. The current goal is to eliminate carbon emissions by 2050, and to keep the warming below 1.5°C.

You are right to wonder if 1.5 degrees sounds too round to be a scientific constant. Fundamental science constants typically have more digits in them. For example, the current hot controversy in fundamental physics is over whether the mass of the W boson is 80,357 MeV/c2 or 80,433 MeV/c2.[21] Whatever the correct value of that physics number, it will not affect government policy. Round climate numbers like 1.5 or 2.0, on the other hand, are quite important for policy even if they lack the exactitude of fundamental physical constants. They are inexact because climate is a highly complex system with many interacting physical, chemical, biological, and human components.

[16] In normal climate terminology, we refer to the multiverse as the *ensemble*. We refer to individual universes as *members* of the ensemble. When the climate isn’t changing, the time average is equivalent to the ensemble average, according to the ergodic hypothesis. In a changing climate, that is no longer the case.

[17] 14 movies and shows about the multiverse, from ‘Spider-Man: No Way Home’ to ‘Everything Everywhere All at Once’ (BusinessInsider.com)

[18] Three reasons why climate change models are our best hope for understanding the future (TheConversation.com)

[19] How we make our 2050 ‘forecasts’, and why we do them (Uk Met Office)

[20] In 2020, the @metoffice produced a hypothetical weather forecast for 23 July 2050 based on UK climate projections. Today, the forecast for Tuesday is shockingly almost identical for large parts of the country. [Tweet by @SimonLeeWx](https://twitter.com/SimonLeeWx/status/1547957062000267267) (Twitter)

[21] Ch.2, The Climate Demon: Past, Present, and Future of Climate Prediction (ClimateDemon.com)

An example of a useful round number is the recommended social distancing threshold for Covid avoidance.[22] The World Health Organization recommends 1 meter distancing, and several countries follow that stringent recommendation. But some other countries recommend 1.5 meter distancing, the US recommends 1.8 meters (6 feet) and the UK recommends 2 meters.

Which is the correct number for social distancing? The answer would be “the largest practical one.” The greater the social distancing, the lesser the health risk. Different numbers reflect the different risk tolerances, and different length units in the different countries.[23] Some people take the Covid distancing thresholds literally, believing that their risk of catching Covid increases dramatically if they cross this threshold even slightly. But many other factors, such as the ventilation and mask efficacy, can have a larger impact on the risk of catching Covid than social distancing.

Although social distancing illustrates the unit-dependence of thresholds, it is not the best analogy for climate thresholds. Because of the nonlinearity of climate impacts, the difference in harm between 1.5 and 2 degrees of global warming is far greater than that between 1.5 and 2 meters of social distancing. Every half degree of warming matters,[24] and many more regions will face serious harm as local warming thresholds are crossed.[25]

A better analogy for a climate threshold is your doctor telling you to keep your bad cholesterol level below 4.0 mmol/L (about 160 mg/dL in the US),[26] rather than keep it below a threshold with more digits of precision, say, 4.123. There is no health “tipping point” that is triggered if you “overshoot”. You are unlikely to suffer a heart attack immediately if your bad cholesterol rises slightly above 4.0, but your risk will increase. If the average cholesterol level of the whole population increases, the number of cardiac disease-related deaths will increase rapidly.

---

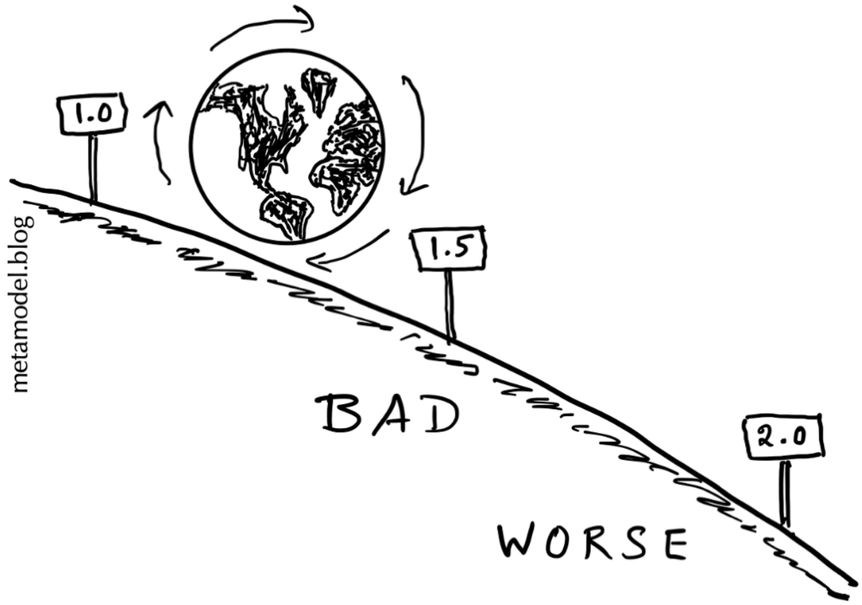
[22] How large does a large ensemble need to be? (S. Milinski et al., 2020; Earth System Dynamics)

[23] What to expect when you’re expecting a better climate model (Metamodel.blog)

[24] Insights from Earth system model initial-condition large ensembles and future prospects. (C. Deser et al., 2020; Nature Climate Change)

[25] Without human-caused climate change temperatures of 40°C in the UK would have been extremely unlikely (WorldWeatherAttribution.org)

[26] Factors other than climate change are the main drivers of recent food insecurity in Southern Madagascar (WorldWeatherAttribution.org)



The globe has already warmed 1.2°C above pre-industrial levels.[27] Our predictions are that the Earth's climate is currently barreling down the slope of ever worsening climate impacts rather than headed towards a cliff at 1.5°C (or 2.0°C).[28] The secret to success in climate mitigation, as it is in life, is to set challenging but achievable goals. A few years ago, 1.5°C appeared to be far enough in the future to serve as an achievable target for stabilizing global temperature, if aggressive steps to mitigate emissions were started immediately. It appears less achievable now, although exceeding the target in a single year is less worrying than the exceeding it in long term.[29]

The roundness of the warming targets depends upon the temperature scale. Those who live outside the United States surely know that water boils at 100°C, a rather round number. That's because the Celsius scale is defined that way. Water freezes at an even rounder 0°C. Again, that is because of the definition of the Celsius scale. These two scientific numbers completely define the temperature scale. *This means that no other scientific temperature numbers can be truly round — except by coincidence or because of rounding.*

On the Fahrenheit scale, the two commonly discussed warming thresholds, 1.5°C and 2°C, would correspond to 2.7 and 3.6°F respectively. In an alternate universe where everyone used the Fahrenheit scale, we might have chosen rounded warming targets of 3.0°F (1.67oC) or 3.5°F (and this letter might be addressed to 1600 Pennsylvania Avenue instead).

We should work hard – starting now – to keep global warming below our chosen target, be it 1.5°C or 2.0°C, or something in-between. But if we overshoot by a small fraction of a degree, the world will not end. Global warming thresholds should be taken **seriously,**

[27] Stop blaming the climate for disasters (E. Raju et al., 2022; Communications Earth & Environment), Politics of attributing extreme events and disasters to climate change (M. Lahsen and J. Ribot, 2021; WIREs Climate Change), and It's Not Just Climate: Are We Ignoring Other Causes of Disasters? (Yale Environment 360)
[28] The climate crisis can't be solved by carbon accounting tricks (The Guardian)
[29] Let's Not Pretend Planting Trees Is a Permanent Climate Solution (New York Times)

but not literally.

3.1 An even rounder target: net-zero

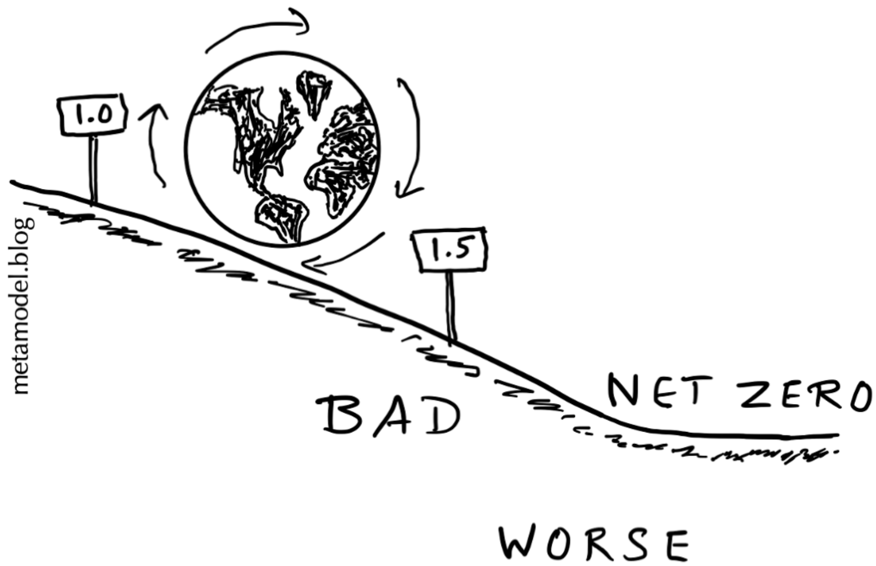
What about another very round number, net-zero, or reducing carbon dioxide emissions to zero? Why is the zero-emission target more appropriate than a target of, say, 3 Gigatons per year or -3 Gigatons per year? (For reference, current fossil-sourced carbon dioxide emissions are about 36 Gigatons per year.) The roundness of net-zero turns out to be coincidental.

Early climate mitigation research focused on keeping carbon dioxide concentrations constant, but that would have led to continued warming over centuries, until the ocean absorbed enough heat to reach equilibrium. This was referred to as "committed warming." Subsequent research showed that if we reduce carbon emissions to zero, the land and ocean will continue to absorb carbon dioxide and steadily lower its concentration.[30] Coincidentally, the cooling effect of this CO2 absorption roughly cancels the effect of continued ocean warming. This means that we can expect global temperatures to stabilize shortly after emissions go to zero. (Ideally, all carbon dioxide emissions should cease, but in practice some unavoidable positive emissions may need to be offset by yet-to-be-perfected negative emissions technology.)

To explain it better: If the Earth is like the human body, carbon emissions keep putting additional blankets on the body.[31] Land and ocean are continually removing about half of these extra blankets. The remaining extra blankets add to the warming of the body. If carbon emissions stop and the number of blankets stays constant, the warming will continue for several more minutes until the human body reaches a warmer equilibrium. In the case of climate, it would take several more centuries to reach the warmer equilibrium. But land and ocean will continue to remove the carbon dioxide blankets even after emissions stop, reducing the number of blankets. This permits temperature to reach equilibrium sooner — within a few seconds in the case of the human body and within a few decades for the climate system.

[30] The Climate Solution Actually Adding Millions of Tons of CO2 Into the Atmosphere (Propublica)

[31] The actual greenhouse effect is more complicated than this, due to the shortwave feedback. See The Greenhouse Effect (And Then There's Physics)



If it were not for the coincidental cancellation between atmospheric carbon dioxide reduction and ocean heat uptake, we would have a less round (and non-zero) emissions target to stabilize global temperature. Alternatively, we may have had to choose a non-zero rate of warming as a practical mitigation target. Note that because of the uncertainty in climate and carbon cycle models, we cannot be absolutely sure that net-zero emissions will stabilize temperatures exactly:[32] Global temperatures may still trend upward slightly when we reach net-zero, or possibly trend down a bit, if the uncertainty works in our favor.

The year 2050 is the notional target for reaching net-zero. It was chosen for practical (and political) reasons based on assessments of how quickly emission reductions could be achieved. If we can reach net-zero by 2040 or 2045, all the better. What if we only reach net-zero by 2055 or 2060? We'll then have to bear the increasingly harmful impacts of the continued warming, but we aren't likely to cross a global climate tipping point.

Sincerely,

metamodel.blog

3.2 Comments

Note: For updated comments, see the original blog post and the anouncement tweet.

- *Chris Wells*:
  Interesting, thought-provoking piece! Think this issue highlights the unavoidable connection between climate science and wider society, as well as the inherent trade-offs in different types of mitigation – and hence the need for a holistic overview contained in 1 (or a handful of) round number.
  Worth noting not everyone agrees Elon Musk is an «Intelligent Mind», and that his precise bid was likely influenced by other factors… https://www.nytimes.com/2022/04/14/business/e.

[32] Is there warming in the pipeline? A multi-model analysis of the Zero Emissions Commitment from CO2 (Biogeosciences)

---

4 Can we predict global warming using only statistics?

Not if it is unprecedented and nonlinear. You can't do statistics with a sample size of less than one. Science is our only hope.

Metamodel.blog 2022-06-07

> If you want to make a prediction, you better make sure that you're either in sample, or that you know the differential equation. [33]

In discussions of global warming, you sometimes hear arguments that you can't trust the complex climate models, that we should rely solely on data to predict the future, and that we should observationally constrain model predictions. This leads us to ask: Can we predict global warming using only statistics? We also ask a related question: Can we identify the causes of global warming using pure statistics?

Let us start by defining *temperature* as that which is measured by a thermometer, and *global warming* as a rise in the global average temperature of the order of 2°C over 200 years. It is the typical magnitude of warming that is expected to occur, say, between 1900 to 2100.

[33] A Physicist Explains Why Parallel Universes May Exist (NPR.org)

Is global warming unprecedented, i.e., has the planet spontaneously warmed in such a manner in the past? Reliable thermometers were only invented about 300 years ago, and accurate global measurements are available only since the late 19th century. So strictly speaking, we only have records for one event of global warming, the current one that is still ongoing. Therefore, the statistical sample count for the global warming events in the available data is a fraction that is less than one! This is true even though we have over a hundred years of temperature measurements. The measured temperature is the sum of the slow global warming event (*the signal*) superimposed with many fast events, such as El Niño (*the noise*) (Figure 1). We need multiple independent samples of global warming to separate the signal properties from the noise.

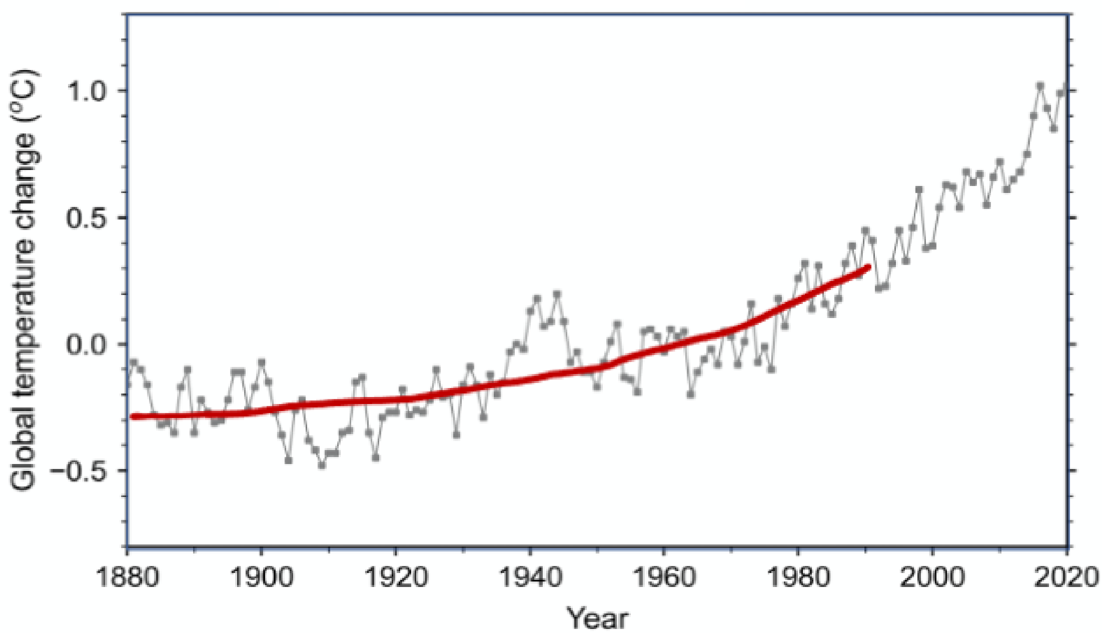


Figure 1 *Annual global average surface temperature change with reference to the1951—1980 base period (grey line with boxes). Thick red line shows the monotonic global warming signal.* [NASA GISTEMP v4 dataset]

What if we relax the definition of temperature to allow inferred measurements of temperature? For example, pollen, sediments, ice cores etc. serve as *proxy records* of indirect information about local temperatures. We then need to use several different models, involving assumptions about isotope fractionation, rainfall patterns, etc., to infer temperatures from these proxies. At this point, we would no longer be analyzing pure data, but model-filtered data.[34] This means that we have to take into account model-related errors into the data analysis.

Analyses of proxy records over the past several thousand years show that the current global warming event is unprecedented compared to climate variations during that period.[35] If we look further back in time, many millions of years, there were larger warm-

[34] In normal climate terminology, we refer to the multiverse as the *ensemble*. We refer to individual universes as *members* of the ensemble. When the climate isn't changing, the time average is equivalent to the ensemble average, according to the ergodic hypothesis. In a changing climate, that is no longer the case.

[35] 14 movies and shows about the multiverse, from 'Spider-Man: No Way Home' to 'Everything Everywhere All at Once' (BusinessInsider.com)

ing (and cooling) events, but they occurred over many millennia, not a few centuries (as far as we can tell using imperfect data). So the rapid 2°C warming over 2 centuries remains unprecedented in pace, if not in amplitude.

Once we conclude that the current global warming is occurring in an unprecedented manner, we can no longer rely solely on data to predict its future time evolution. Trying to predict global warming using a purely statistical model is like trying to predict the very first observed El Niño event using a statistical model. Simple linear extrapolation of the observed warming in the tropical Pacific associated with El Niño would have missed the quasi-cyclical aspect of the phenomenon. Now that we have data from many El Niño events, we can build skillful statistical models for El Niño prediction that account for the quasi-cyclical aspect.

Over any short enough recent period, say a couple of decades, we can approximate global warming as a linear trend. Extrapolating that linear trend for a couple of more decades may not be a bad approximation. But beyond that, will the trend stay linear? Or will it start to bend upwards parabolically (or amplify exponentially) at some future time? Or will it start to slow down like the square-root function or flatten like a sinusoid curve? (Figure 2) Data alone cannot tell us how the ongoing global warming event will evolve because we have no examples of past global warming events to fit our statistical models.

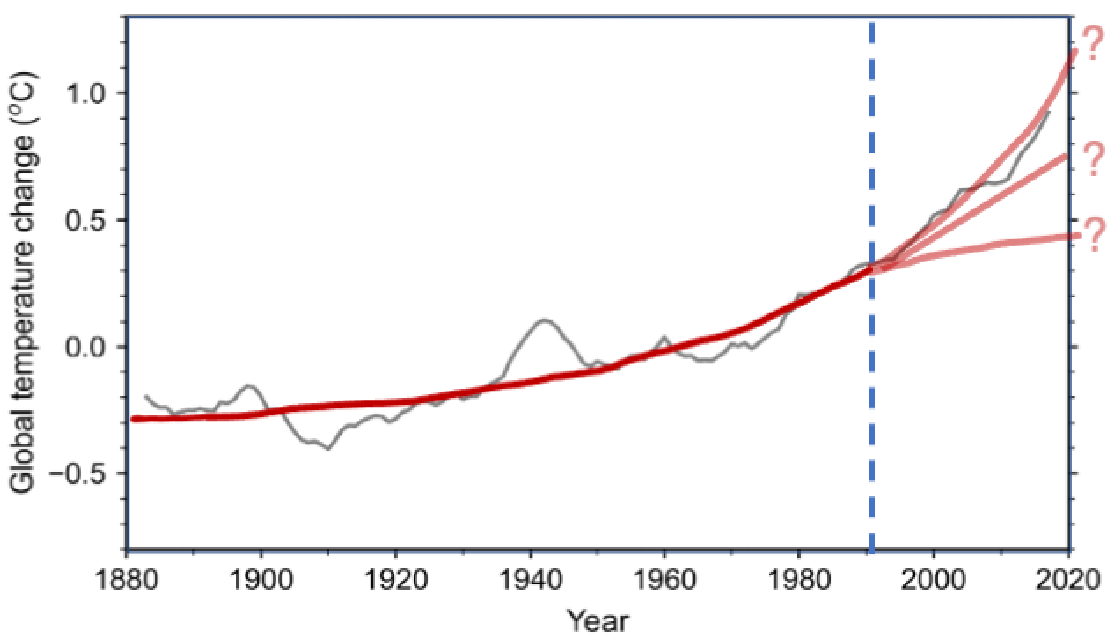


Figure 2 *Seven-year moving average of the annual global average surface temperature change in Figure 1 (grey line). Thick red line shows a schematic of three alternative predictions of global warming after 1990. Without precedents, statistics cannot choose between the divergent predictions*

Science-based models should be able to predict an unprecedented nonlinear event, but these models suffer from many imperfections. The coarse spatial grid of a global climate model cannot capture important fine-scale processes like clouds. Therefore, approximate formulas called *parameterizations* are used to represent clouds. Other processes, including the carbon cycle, are also represented using approximate formulas that are calibrated using data. The uncertainties associated with these approximations result in a considerable spread in the predictions of science-based models. The spread turned out to be particularly large in the latest generation of science-based climate models, known as CMIP6.[36]

Frustration with the growing complexity of global climate models and the uncertainties in their parameterizations can motivate scientists to look for alternatives — perhaps there is a statistical approach that relies on pure data to analyze and predict global warming, uncontaminated by imperfect models. One can learn a lot from statistical analysis of data. But when it comes to actual prediction of global warming, the grass may not really be greener on the statistical side.

To highlight some important issues in applying statistics to global warming, we consider two recent studies that resort to statistical approaches, but in rather different ways. One is the comprehensive IPCC AR6 report[37] that applies statistical "observational constraints" to model predictions. The other is a more narrow study, a recent paper by Koutsoyiannis and co-authors (K22) that uses a purely statistical approach to address the causality of climate change.[38]

A common refrain of those favoring statistical approaches is to let "the geophysical records speak for themselves" (a direct quote from the K22 paper). Like a chatty old person at a party, data can speak and tell you many stories, but they may not be interesting to you. You may want to know how fast the current centennial scale warming will progress, but data may tell you instead about past warming that occurred at a more "glacial" multi-millennial timescale. Often, these data stories need interpretation and translation. You need scientific context to interpret stories about correlations. You need models to translate stories about pollen into statements about temperature.

4.1 Adjusting model predictions using data

First, we consider the IPCC assessments. Previously, the IPCC relied solely on science-based global climate models for predicting global average temperature. But for the most recent IPCC assessment (AR6), some of the newer models were predicting much more warming than the previous generation of models, which was at variance with new observational constraints. To deal with this, the IPCC has introduced a complicated hybrid approach that includes statistically constrained "emulator" models.

The emulators are highly simplified climate models with adjustable coefficients. The coefficients in the emulator equations are "calibrated" using results from complex science-based global climate models. For example, if the global models were simulating too much warming during the observed period, the emulators would also do that. But we can adjust some coefficients in the emulators to constrain the rate of simulated warming. The adjusted emulators can then be used to extrapolate into the future. (Ideally, the global models should be adjusted at the process level to improve their simulated warming.[39] Presumably the IPCC resorted to a short-cut because adjusting the innards of global climate models is much harder than adjusting the globally averaged parameters of emulators.)

[36] Three reasons why climate change models are our best hope for understanding the future (TheConversation.com)

[37] How we make our 2050 'forecasts', and why we do them (Uk Met Office)

[38] In 2020, the @metoffice produced a hypothetical weather forecast for 23 July 2050 based on UK climate projections. Today, the forecast for Tuesday is shockingly almost identical for large parts of the country. [Tweet by @SimonLeeWx](https://twitter.com/SimonLeeWx/status/1547957062000267267) (Twitter)

[39] Ch.2, The Climate Demon: Past, Present, and Future of Climate Prediction (ClimateDemon.com)

Are these emulators superior to the global climate models? Yes, and no. Yes, because they provide an ad hoc fix for the immediate problem of some models predicting too much warming. No, because they are still calibrated using the global models. The observational constraints cannot fix all the deficiencies inherited by the emulators from the global models. Also, the emulators predict just the globally averaged temperature. Only comprehensive global models can provide the regional detail needed for risk assessment. (The IPCC introduces the notion of a global warming level (GWL) to combine predictions from emulators with global climate models, but the GWL approach does not provide the temporal information needed for risk assessment.[40])

Can't we directly calibrate the emulator using data, thus avoiding the reliance on global climate models? No, because global warming is an ongoing, unprecedented event. The emulator can emulate the known past but not the unknown future of this event. For a given emission scenario, whether the warming trend of the recent decades will continue as a linear trend, or accelerate/decelerate nonlinearly at some future time (Figure 2), will be determined by the behavior of global models used to train the emulator.[41] Even using statistics to constrain long-term equilibrium properties like climate sensitivity cannot fully determine the short-term evolution of global warming. To have more trust in the nonlinear aspects of the emulators, we need to improve the parameterizations in global climate models. However, just making the parameterizations more complex may not reduce the uncertainties, as the recent increase in the spread of model predictions indicates.

4.2 Determining the causes of global warming using just statistics

The K22 paper aims to let the data speak for itself by using pure statistics to study global warming. The paper analyzes two different types of precedented events, ice age cycles and year-to-year climate variability. Correlations between temperature and carbon dioxide are computed to determine which is leading which.

For the ice core data, K22 finds that temperature leads carbon dioxide during the multi-millennial ice age cycles. For the year-to-year variability, they also find that temperature leads carbon dioxide. The paper then claims that this result "contradicts common opinion" because the correlations seem to imply that the observed warming is driving the observed increase in carbon dioxide. But this is a misinterpretation arising from the conflation of processes occurring at different timescales. Both correlations identified in the paper are well-known relationships at their respective timescales, and are not inconsistent with the notion that carbon dioxide leads temperature on centennial timescales. (The year-to-year correlations between temperature and carbon dioxide have been discussed in many earlier studies.[42])

With no scientific context, it is easy to misinterpret statistical correlations because:

- *Correlation does not imply causation*: If correlation analysis reveals that changes in variable A lead changes in variable B, it does not mean A causes B. There could be a third variable C that is controlling both. For the ice core correlations analyzed by K22, the third variable is the Earth's orbital perturbations that affect the amount

[40] How large does a large ensemble need to be? (S. Milinski et al., 2020; Earth System Dynamics)
[41] What to expect when you're expecting a better climate model (Metamodel.blog)
[42] Insights from Earth system model initial-condition large ensembles and future prospects. (C. Deser et al., 2020; Nature Climate Change)

of sunlight received.[43] Warming at the end of an ice age releases stored carbon dioxide into the atmosphere. For the interannual correlations analyzed by K22, the third variable is likely the El Niño index, which affects both temperature and carbon dioxide.

- *Correlation between variables depends on timescales*: The climate system has different processes, with different causal mechanisms and different lead-lag relationships between temperature and carbon dioxide, that co-exist and operate at different time scales. This includes ice age cycles, the seasonal cycle, and the year-to-year El Niño variations, and global warming. To the extent data of multiple realizations of a process are available, statistics can capture the lead-lag relationships of a specific process associated with a particular timescale. We have such data for the ice ages, the seasons, and El Niño events, but not for the rapid global warming that is happening now.
- *A trend is correlated with every other trend*: You can correlate an increase in global temperatures with the increase in the stock market index or just about any other quantity that has increased over the last century. Monotonic trends are a nuisance in correlation analysis and are usually subtracted out before computing correlations. Therefore, we cannot infer anything causal about the current global warming event from a purely statistical analysis of trends in recent data.

Only by introducing a comprehensive scientific framework can we extract useful information from observed correlations and trends. A partial scientific framework can still lead to misinterpretation of results from data analysis if important processes are neglected.

4.3 Conclusions

We cannot predict or explain an unprecedented nonlinear phenomenon using statistics alone.[44] A comprehensive science-based model can predict an unprecedented event because it solves the differential equations governing the phenomenon. But how much we should trust that prediction depends upon how accurately all the relevant equations are known. Statistical analysis can help improve the equations in the science-based models, and constrain some of their properties, but it cannot replace them for predicting the temporal and spatial detail of future global warming.

Winston Churchill said that “democracy is the worst form of government except all those other forms that have been tried”. We could say the same of complex global climate models — they are imperfect but alternatives such as purely statistical models or highly simplified models are even less scientifically defensible for predicting climate. And we need models to plan for the future.

Note:

Some outside mainstream climate science would argue that the current global warming is not unprecedented and that warming events of similar pace and amplitude have occurred naturally before. Multiple lines of evidence do not support this argument,[45] but we can consider the hypothetical scenario where global warming has precedents in

[43] Without human-caused climate change temperatures of 40°C in the UK would have been extremely unlikely (WorldWeatherAttribution.org)

[44] Factors other than climate change are the main drivers of recent food insecurity in Southern Madagascar (WorldWeatherAttribution.org)

[45] 14 movies and shows about the multiverse, from ‘Spider-Man: No Way Home’ to ‘Everything Everywhere All at Once’ (BusinessInsider.com)

the data record. In that case, some of the statistical limitations described in this article would no longer apply. However, there is no escaping the need for models.

As noted before, pure data cannot be used to determine whether global warming has precedents, since we do not have global temperature measurements prior to the 19th century. We have to use models to calibrate proxies for temperature in making any such determination. Even if we identify enough past global warming events to provide a sizeable statistical sample, we would need to have accurate temperature data with at least decadal time resolution to train a statistical prediction model. If we manage to do that, we would still require good science-based global climate models to add the necessary spatial detail to any statistical prediction of global average temperature.

4.4 Comments

Note: For updated comments, see the original blog post and the anouncement tweet.

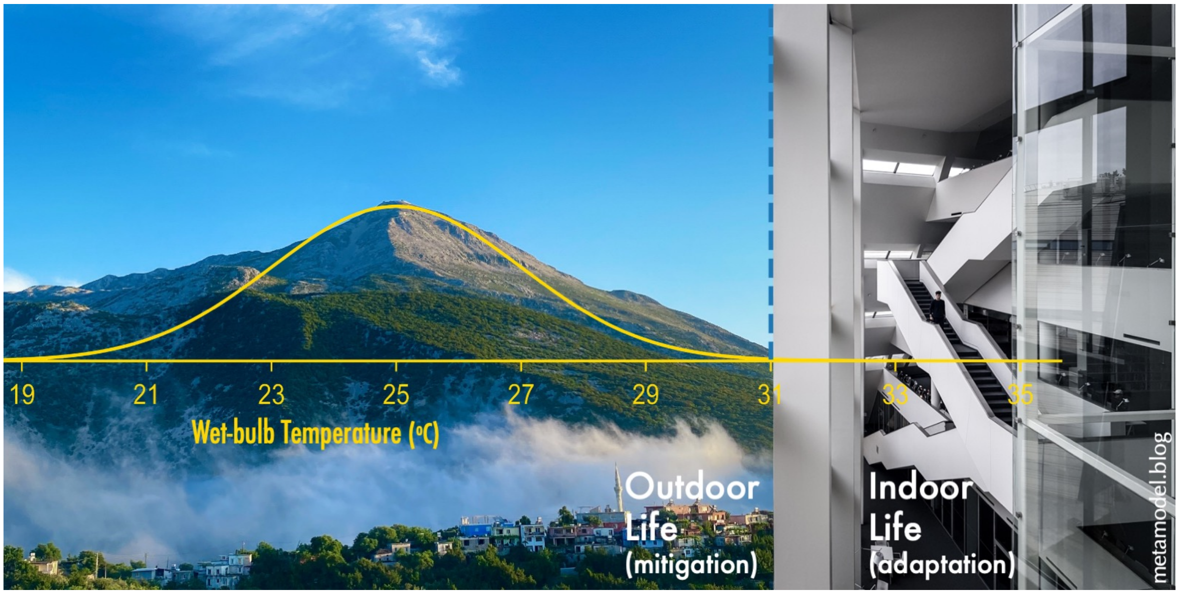
- *Dikran Marsupial*:
  I believe the error with the modern temp/CO2 data is primarily due to differencing the time series before analysis. This means the long term trend becomes a constant offset in the time series, and their causal inference algorithm can't infer a causal relationship between constant offsets in temperature and lnCO2. So all they are looking at is the already known ENSO-CO2 correlation (first noticed back in the 70s by Bacastow) and they can infer precisely nothing about the long term increase as their algorithm wouldn't be able to detect it.
  They are far from being the first to make this mistake:
  https://skepticalscience.com/salby_correlatio...
  and
  https://skepticalscience.com/roys_risky_regre...

---

5 The perils of predicting perils: (mis)calculating wet-bulb temperature

Many assumptions lurk behind climate risk assessments. Small errors in the average can mean large errors in the tail risk.

Metamodel.blog 2022-06-22



Is life worth living? It depends on the liver. So goes the old joke. Is life worth living in a warming climate? The answer might be: It depends on the wet-bulb temperature. Like the state of the liver determines whether you can drink alcohol, the value of the wet-bulb temperature determines whether you can survive without air conditioning — if the value exceeds 35°C, you just cannot.

Wet bulb temperature is an obscure meteorological metric that has gained prominence lately because it measures human survivability. Many recent media articles, such as this one in *The Economist*, use wet-bulb temperatures to characterize the severity of heat waves.[46] Kim Stanley Robinson's cli-fi novel, *The Ministry for the Future*, begins with a graphic description of a lethal heat wave in India with a high wet-bulb temperature.[47]

Wet-bulb temperature can provide many useful insights, but is complicated to calculate. A seemingly small (and innocuous) mistake in the calculation can have a big impact on risk estimation. This may have been the case with a high-profile report on climate change issued by the international consulting company McKinsey in 2020.[48] The report predicted that by 2030 that hundreds of millions people could be living in regions that will experience heat waves that threaten human survivability (under a particular emission scenario). The fine print in the report details a crucial assumption about how wet-bulb temperature was calculated. If that assumption is incorrect, it could affect the conclusions of the report.

At the time McKinsey report was released, I was writing a book about climate prediction.[49] I decided to use a quote from the report in my book as an example of climate risk assessment. I did not evaluate the claims in the report or read the fine print. I was merely using the quote to make the point that such reports are very influential. But about a week ago, climate scientist Patrick Brown tweeted about a potentially serious

[46] A Physicist Explains Why Parallel Universes May Exist (NPR.org)

[47] In normal climate terminology, we refer to the multiverse as the *ensemble*. We refer to individual universes as *members* of the ensemble. When the climate isn't changing, the time average is equivalent to the ensemble average, according to the ergodic hypothesis. In a changing climate, that is no longer the case.

[48] 14 movies and shows about the multiverse, from 'Spider-Man: No Way Home' to 'Everything Everywhere All at Once' (BusinessInsider.com)

[49] Three reasons why climate change models are our best hope for understanding the future (TheConversation.com)

flaw in the methodology of the McKinsey report.[50] You can read that tweet thread for more details and background information. These tweets motivated me to re-examine the report and also learn more about wet-bulb temperature.

5.1 Wet-bulb temperature

What is wet-bulb temperature? It is literally the temperature measured by a special thermometer that has a “wet bulb” — the bulb is the portion of the thermometer that senses temperature. A normal thermometer has a dry bulb. A wet-bulb thermometer has a bulb covered with a wet cloth, which is analogous to a human body that is sweating. As the sweat evaporates, it cools the body and therefore it feels cooler than the temperature of the surrounding air. The wet-bulb thermometer essentially measures how cool it feels, sort of like a heat index.

The wet-bulb temperature is always cooler than the actual air temperature. How much cooler depends upon how humid the surrounding air is. If the air is dry, water can evaporate easily, cooling the thermometer. When the humidity approaches 100%, the wet-bulb temperature will approach the air temperature, because water can no longer evaporate as the air becomes saturated with moisture.

Why all the recent interest in this somewhat arcane meteorological measure? It’s because wet-bulb temperature has implications for outdoor activity and survivability of humans.[51] As warm blooded creatures, we continuously generate heat that must be expelled to maintain our body temperature at 37°C (98.6°F). When the air is cool or dry, sweating allows us to do just that. The opposite of “cool or dry” is “hot and humid”. When it is hot and humid, sweating becomes inefficient as a heat loss mechanism.

The wet-bulb temperature is a combined measure of heat and humidity that tells us whether we can continue to cool our bodies naturally. If the wet-bulb temperature exceeds a hard theoretical threshold of 35°C, it becomes impossible to do that. We cannot survive without air conditioning. There is also a softer practical threshold of about 31°C beyond which outdoor activities will need to be severely curtailed.[52]

We calculate the wet-bulb temperature using a formula that takes surface temperature and relative humidity as inputs. Temperature and relative humidity vary throughout the day. Dew forms in the pre-dawn hours of the morning, when temperature is the lowest but relative humidity is the highest. As temperature increases during the day, the relative humidity typically falls. This anti-correlation between temperature and relative humidity affects wet-bulb temperature.

5.2 Fine print in the McKinsey report

The McKinsey report relies on predictions of how wet-bulb temperatures will change in the future for many of its headline findings. To calculate wet-bulb temperature accurately, we should ideally use hourly information of surface temperature and relative humidity. But such detailed hourly information is not always saved for model predictions of the future. Often, only the daily average or maximum/minimum values are available.

[50] How we make our 2050 ‘forecasts’, and why we do them (Uk Met Office)

[51] In 2020, the @metoffice produced a hypothetical weather forecast for 23 July 2050 based on UK climate projections. Today, the forecast for Tuesday is shockingly almost identical for large parts of the country. [Tweet by @SimonLeeWx](https://twitter.com/SimonLeeWx/status/1547957062000267267) (Twitter)

[52] Ch.2, The Climate Demon: Past, Present, and Future of Climate Prediction (ClimateDemon.com)

As described in a footnote,[53] the McKinsey report uses daily maximum temperature and daily average relative humidity to compute the wet-bulb temperature. The implicit assumption is that this procedure provides a good estimate of the daily maximum wet-bulb temperature.

We cannot check the procedure used by the McKinsey report simply by looking at the formula for wet-bulb temperature. The formula is so complicated and nonlinear that it is hard to figure out how variations in temperature and relative humidity affect the wet-bulb temperature. The strong nonlinearity means that inadvertent time averaging could potentially introduce large errors.

To check the procedure used in the report, we need to analyze data. We consider hourly variations in temperature and relative humidity during May 2002 at one location, New Delhi, India, which lies in our region of interest.

Figure 1 shows that temperature and relative humidity at New Delhi vary in an anti-correlated manner throughout the day. The wet-bulb temperature computed using hourly data varies between 19.4°C and 24.2°C. The straight line shows the wet-bulb temperature computed using the daily maximum temperature and daily average relative humidity. This has a value of 26.7°C, which is about 2.5°C higher than the maximum wet-bulb temperature computed from the hourly data.

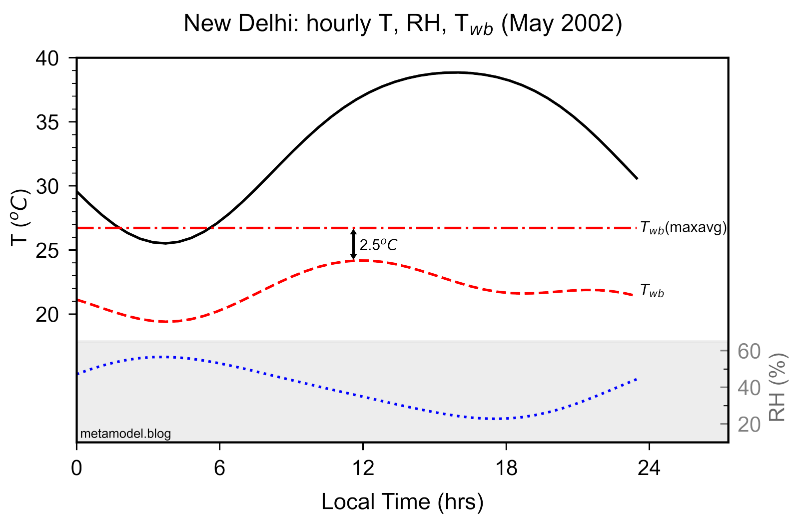


Figure 1. *Hourly surface temperature (black solid) and Relative Humidity (blue dotted) for New Delhi (India) during May 2002. Wet-bulb temperature calculated from hourly data (red dashed).*[54] *Wet-bulb temperature calculated using daily maximum temperature and daily average relative humidity (red dashdot).* [Fourier-smoothed data obtained from Patel et al. 2002 [55]]

When we combine the maximum temperature with average relative humidity, we overestimate the wet-bulb temperature. To understand the implications of this, let us consider a hypothetical normal probability distribution of wet-bulb temperature for May in New Delhi, centered at 25°C. The standard deviation (σ) for daily maximum tempera-

[53] How large does a large ensemble need to be? (S. Milinski et al., 2020; Earth System Dynamics)
[54] What to expect when you're expecting a better climate model (Metamodel.blog)
[55] Insights from Earth system model initial-condition large ensembles and future prospects. (C. Deser et al., 2020; Nature Climate Change)

ture in New Delhi is about 2.5°C.[56] Let us assume that the standard deviation for daily maximum wet-bulb temperature is somewhat less, say 2.0°C.[57] Then the probability distribution for wet-bulb temperature corresponds to distribution A in Figure 2. If we overestimate the daily wet-bulb temperature by 2.0°C, the distribution will shift to the right by one standard deviation or one-σ, forming distribution B in Figure 2.

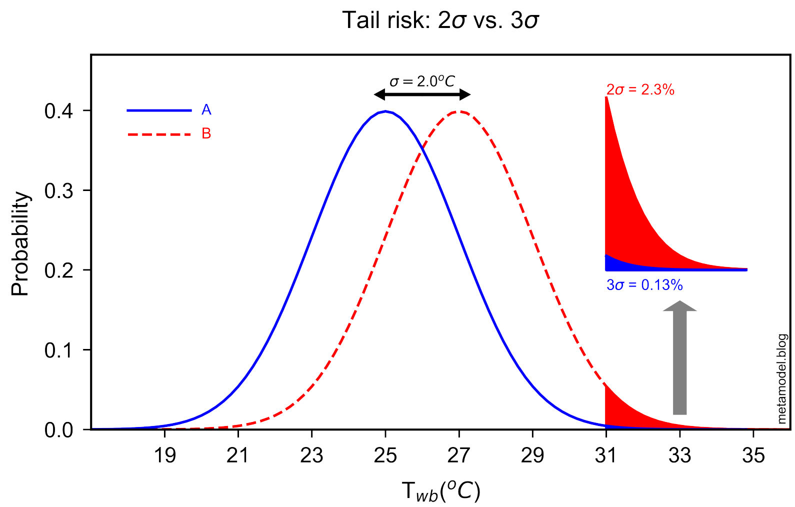


Figure 2. *Assumed normal probability distribution of wet-bulb temperature with standard deviation of 2°C. A. Centered at 25 °C (blue). B. Centered at 27°C (red). Probabilities of wet-bulb temperature exceeding 31°C are shaded, and the shaded areas under the curve are shown as percentages.*

Suppose we are interested in extreme values of the wet-bulb temperature, say exceeding a soft threshold of 31°C (the McKinsey report actually uses a more severe threshold of 34°C to define "lethal" heat waves.) For distribution A, the 31°C threshold is three-σ away from the mean of 25°C. This means that there is a 0.13% probability that the soft threshold will be exceeded. For distribution B, the 31°C threshold is only two-σ away from the mean of 27°C. This means that there is a 2.3% probability that the soft threshold will be exceeded.

As illustrated in Figure 2, a 2.0°C error in estimating the wet-bulb temperature can lead to a factor of 17(=2.3/0.13) error in estimating the probability of exceeding the 31°C threshold. Recall that using the average relative humidity resulted in an overestimation of the maximum wet-bulb temperature by 2.5°C. This 1.25σ error would overestimate the probability of exceeding the threshold by a factor of 39. If errors of this magnitude are present in the McKinsey report, then its conclusions will need to be revised. Whether or not that's the case, this analysis serves as a cautionary tale in estimating the tail risk of climate change.

[56] Without human-caused climate change temperatures of 40°C in the UK would have been extremely unlikely (WorldWeatherAttribution.org)

[57] Factors other than climate change are the main drivers of recent food insecurity in Southern Madagascar (WorldWeatherAttribution.org)

5.3 What do we learn from this analysis?

- Climate risk reports from consultants often contain precise sounding impact numbers and probability estimates. But these numbers may depend on various assumptions, not all of which may even be noted in the report. Reading the fine print and requesting any additional supporting information is a good idea. It is also worth carrying out simple back-of-the-envelope calculations to check the numbers using data at selected locations.
- Consultant reports that have not been thoroughly peer-reviewed by experts should be considered less authoritative than scientific reports like IPCC assessments that have undergone extensive peer review.
- In risk assessment, a modest error in the estimate of the average can alter tail risk probabilities by an order of magnitude. This is the flip-side of the argument made about the frequency of extreme events in a changing climate, that even a small increase in the average temperature can lead to a big change in the frequency of extreme heat waves.[58]
- The difficulty in quantifying the tail risk means that we should perhaps not take many numeric estimates of tail risk literally, even as we take the overall risk assessment itself seriously.

Climate change is a serious threat, and we don't always need fancy computer models or voluminous consultant reports to help us appreciate that. Reinterpreting Figure 2, we can see that a 2.0°C rise in the average wet-bulb temperature will increase the probability of exceeding the 31°C threshold by a factor of 17. A 1°C rise, which is well within the realm of possibility based on current warming trends, will increase the probability of exceeding the threshold by a factor of 4.6, which still amounts to a large increase in risk.[59]

5.4 Comments

Note: For updated comments, see the original blog post and the anouncement tweet.

- *R Saravanan*:
  One reader sent me this funny (and insightful) cartoon about wet-bulb temperature: Why do we care about wet bulb temperature and could they have given it a better name?
  First Dog on the Moon (The Guardian)

---

6 Climate Startups, Carbon Offsets, and Crypto

Theranos, WeWork ... Carbonos? With startups selling products that affect the planet's health, physical checks are more important than fancy crypto. The wild west of carbon offsets needs a sheriff.

Metamodel.blog 2022-06-28

[58] Stop blaming the climate for disasters (E. Raju et al., 2022; Communications Earth & Environment), Politics of attributing extreme events and disasters to climate change (M. Lahsen and J. Ribot, 2021; WIREs Climate Change), and It's Not Just Climate: Are We Ignoring Other Causes of Disasters? (Yale Environment 360)

[59] The climate crisis can't be solved by carbon accounting tricks (The Guardian)

What happens when the startup mantra "fake it till you make it" is taken to the extreme? It doesn't end well. It happened with WeWork, a real estate company that was pretending to be a tech company. It also happened with the Silicon Valley unicorn Theranos that promised to revolutionize blood testing using a pinprick, but faked test results and failed to make it. The high-profile startup was brought down by an investigative reporter and an inspector from a "boring" government agency that regulated medical testing.[60]

The puncturing of the hype surrounding WeWork hurt only its investors and employees. The hype perpetuated by Theranos endangered the health of ordinary human beings. What about the hype from startups that are selling products that affect the health of the entire planet?

As a climate scientist, I have mixed feelings about the recent surge in private funding directed at climate-related projects. Billons of dollars are flowing to startups are seeking to offset carbon emissions[61] or assist with carbon disclosures mandated by new financial regulations. It is inspiring that they are developing products to help improve the planet's health. However, unlike earlier startups that focused on energy technology like batteries, the newer startups are selling less tangible products like accounting and offsets. How much should we trust these commercial products?[62]

Some of the new climate startups also tout their use of crypto and the blockchain. Like many, I believed in the Theranos hype about helping humanity because I didn't know any better. But I know how the climate system works and I also have some experience with cryptography, having worked on open-source email encryption.[63] That makes me

[60] A Physicist Explains Why Parallel Universes May Exist (NPR.org)

[61] In normal climate terminology, we refer to the multiverse as the *ensemble*. We refer to individual universes as *members* of the ensemble. When the climate isn't changing, the time average is equivalent to the ensemble average, according to the ergodic hypothesis. In a changing climate, that is no longer the case.

[62] 14 movies and shows about the multiverse, from 'Spider-Man: No Way Home' to 'Everything Everywhere All at Once' (BusinessInsider.com)

[63] Three reasons why climate change models are our best hope for understanding the future (TheConversa-

rather wary of the hype surrounding this new breed of climate startups.[64]

Private innovation plays an important role in climate solutions. Pledges for curbing emissions rely on advances in electric vehicles and the development of efficient batteries, solar panels, and wind turbines. The price of renewable energy has dropped dramatically over the last decade because of such innovative products, which are subject to the *caveat emptor* or "buyer beware" principle. You can verify the range of the electric car or the efficiency of a solar panel after you buy it. But products like carbon offsets and carbon accounting are quite different. Like diagnostic blood tests, only an independent authority can certify if the product works. Just as a false negative test can hide the progression of disease, a flawed carbon offset can prolong global warming and the associated harm.

6.1 Climate and crypto

Startup culture tends to lean libertarian because regulations can stifle innovation in many areas. But in the climate space, promising "solutions" may be worthless without strong regulations. Combining climate solutions with hot trends like crypto, as some startups are doing, is not a good idea.

Using computers to mine proof-of-work crypto currencies consumes prodigious amounts of energy, which often results in increased carbon emissions directly or indirectly.[65] Other crypto features like the blockchain facilitate anonymous transactions without middlemen, but these obfuscatory features may actually work against the accountability essential for climate solutions.[66]

Consider the following transaction: MegaCorp purchases 10,000 tons of carbon offsets for $1 million from the green startup Carbonos using bitcoin. The beauty of the blockchain is that it allows the money transfer to be validated without the need for a trusted intermediary. But what about the validity of the carbon offsets purchased? This cannot be verified by doing math on a blockchain. You need boots on the ground to verify that the captured carbon stays underground. When it comes to carbon capture, the "blockchain" one should trust is the physical chain of custody of blocks of carbon from extraction to permanent sequestration.

If MegaCorp is buying offsets merely for public relations to claim carbon neutrality, or to hawk cheap offsets to its retail customers, it may be a case of "buyer doesn't care" rather than "buyer beware". Both parties may be quite satisfied with a shoddy but cheap product.[67] Only an independent regulatory authority, much like the "boring" government agency that inspected Theranos, can be trusted to verify offsets and protect the health of the planet.

Once we have a trusted authority to verify carbon accounting and offsets, there is really no need for the trappings of crypto, except perhaps to impress venture capitalists. An important reason for investor fascination with crypto is the lure of quick profits. Although that fascination may have faded a bit with the recent downturn in crypto valuations, it

---

tion.com)

[64] How we make our 2050 'forecasts', and why we do them (Uk Met Office)

[65] In 2020, the @metoffice produced a hypothetical weather forecast for 23 July 2050 based on UK climate projections. Today, the forecast for Tuesday is shockingly almost identical for large parts of the country. [Tweet by @SimonLeeWx](https://twitter.com/SimonLeeWx/status/1547957062000267267) (Twitter)

[66] Ch.2, The Climate Demon: Past, Present, and Future of Climate Prediction (ClimateDemon.com)

[67] How large does a large ensemble need to be? (S. Milinski et al., 2020; Earth System Dynamics)

has not gone away in carbon markets.[68]

When used for investment rather than for transactions, a bitcoin is rather like a digital tulip[69]—it has no intrinsic value beyond that determined by the market. It is not the cryptographic features of bitcoin that allow investors to make a quick buck, but its limited supply and appreciation potential. If you can predict the future value of any tradable commodity, you can profit from that. That is how you profit from trading bitcoin, which is no different from trading stocks, or emissions/offsets in a carbon market. There is no need for a complicated blockchain to buy and sell stock listed in an exchange.[70]

6.2 Carbon offsets

Many carbon offsets, especially the cheap options that businesses like airlines offer,[71] remind me of the famous P. T. Barnum quote: "There's a sucker born every minute".[72] As a human being worried about climate change, I too would like to believe that we can miraculously find a way to suck carbon out of the atmosphere at an affordable cost. But as a climate scientist aware of the complexity of the carbon cycle, I am not going to believe that we have sucked out carbon until an independent expert confirms it. I'd rather contribute to a charity that purchased solar panels for developing countries than waste money buying cheap carbon offsets of dubious provenance.

Direct air capture of carbon can potentially be an effective and verifiable solution to offset emissions, but it is not currently affordable and there is no guarantee that it will become so in the future. This and other carbon capture technologies are certainly worth researching, but with the understanding that failure is an option. Focusing too much on capture technologies can distract us from the most effective way to reduce carbon emissions,[73] which is to eliminate the demand for fossil fuels by providing affordable alternatives that use carbon-free energy sources.

Nature-based conservation efforts to offset carbon emissions sound sustainable, but their hard-to-quantify mitigation benefits may turn out to be transitory [74] or even negative.[75] Like describing plastic apparel as vegan leather, eco-friendly names for crypto tokens may just be creative marketing. If you truly care about nature, don't rely on nature-based solutions to offset carbon emissions, but continue to support them as you have done in the past, simply as worthy efforts of nature conservation. In other words, grow your natural beans but don't count them as part of any carbon offset budget.

We don't leave the certification of medical drugs and treatments to industry self-regulation, because the stakes are too high to allow mistakes. We should apply the same standard when the health of the planet is at risk. An independent regulatory authority is essential to *measure, report, and verify* all carbon accounting and offsetting

[68] What to expect when you're expecting a better climate model (Metamodel.blog)

[69] Insights from Earth system model initial-condition large ensembles and future prospects. (C. Deser et al., 2020; Nature Climate Change)

[70] Without human-caused climate change temperatures of 40°C in the UK would have been extremely unlikely (WorldWeatherAttribution.org)

[71] Factors other than climate change are the main drivers of recent food insecurity in Southern Madagascar (WorldWeatherAttribution.org)

[72] Stop blaming the climate for disasters (E. Raju et al., 2022; Communications Earth & Environment), Politics of attributing extreme events and disasters to climate change (M. Lahsen and J. Ribot, 2021; WIREs Climate Change), and It's Not Just Climate: Are We Ignoring Other Causes of Disasters? (Yale Environment 360)

[73] The climate crisis can't be solved by carbon accounting tricks (The Guardian)

[74] Let's Not Pretend Planting Trees Is a Permanent Climate Solution (New York Times)

[75] The Climate Solution Actually Adding Millions of Tons of CO2 Into the Atmosphere (Propublica)

practices. Since there is an incentive to game the offset system even at the national level, the regulator may need to be a trusted international authority.

The verification authority must be able to carry out unannounced physical audits of the carbon offsetting process following established scientific protocols. Lopsided funding for glamorous carbon startups without commensurate support for unglamorous verification infrastructure is a recipe for attracting Theranos-like business models.

Crypto isn't some magic pixie dust. A sprinkling of crypto dust on climate solutions is superfluous at best and can be harmful at worst—if it reduces transparency or increases energy consumption. Crypto enthusiasm may never go away in finance and investing, but keep it out of climate solutions for the planet's sake.

6.3 Comments

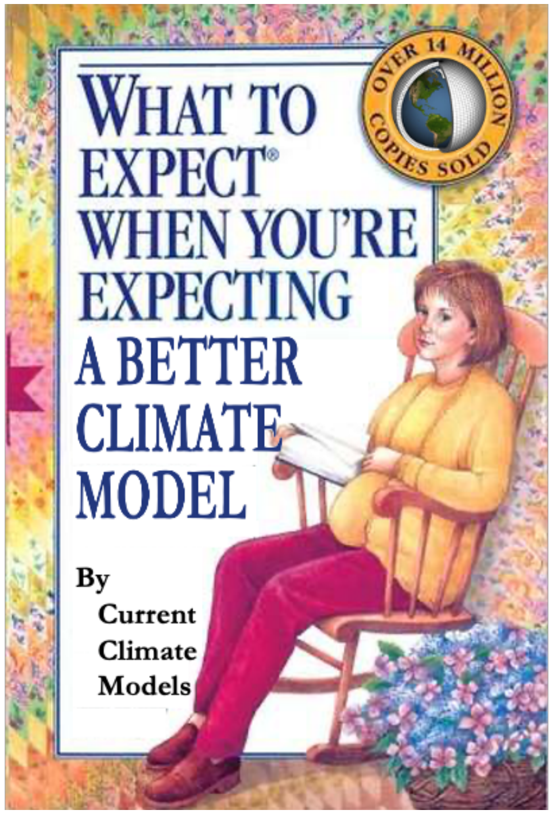
Note: For updated comments, see the original blog post and the anouncement tweet.

---

7 What to expect when you're expecting a better climate model

Irreducible uncertainties associated with internal variability and human actions limit our ability to predict long-term climate change. Higher model resolution can help, but it is not a silver bullet.

Metamodel.blog 2022-07-13



If we build a gigantic supercomputer, ask it the ultimate question, and receive a single number as an answer, what have we learned? Without context, not much. A single number, whether it is 42, as in *The Hitchhiker's Guide to the Galaxy*,[76] or 3°C for Earth's climate sensitivity, doesn't mean much unless we know how it was calculated and what its uncertainty is.

This provides a nice segue to the recent blog discussion about a concerted international effort to build a climate model with a 1-km (k-scale) horizontal grid.[77] That would be a big jump from the current generation of climate models, which typically use a 50-km grid. The common expectation is that a million-fold increase in computer power available for modeling will lead to a quantum leap in our predictive capabilities, thus better informing policy-makers. The headline of a recent Wall Street Journal article, "Climate Scientists Encounter Limits of Computer Models, Bedeviling Policy", reflects this sentiment.[78]

To what extent can better climate models inform policies, and exactly what policies can they help inform? The phrase "actionable predictions" is frequently used in this context, but often without elaboration. How much improvement in predictions can we expect from much better climate models of the future? Will they reduce the error bar by 10%, 50% or 90%? It turns out that our current models have something to say something about that.

[76] A Physicist Explains Why Parallel Universes May Exist (NPR.org)

[77] In normal climate terminology, we refer to the multiverse as the *ensemble*. We refer to individual universes as *members* of the ensemble. When the climate isn't changing, the time average is equivalent to the ensemble average, according to the ergodic hypothesis. In a changing climate, that is no longer the case.

[78] 14 movies and shows about the multiverse, from 'Spider-Man: No Way Home' to 'Everything Everywhere All at Once' (BusinessInsider.com)

7.1 Limits and uncertainties of climate prediction

From our familiarity with weather forecasts, we know there are limits to weather prediction. We don't expect forecasts to be accurate beyond about a week. That's because we have imperfect knowledge of the initial condition for a weather forecast. Small errors in the initial condition grow exponentially over time leading to large errors in the forecast after several days. This property of chaos, known as the Butterfly Effect, limits weather prediction to about two weeks. Even a perfect weather model cannot predict beyond this limit.

Is there a corresponding limit to climate prediction? The usual answer is that the Butterfly Effect does not apply to climate prediction because we are not predicting individual weather events but the statistics of future weather. That's technically true, but what happens to the Butterfly Effect beyond two weeks? The error associated with the Butterfly Effect eventually stops growing and saturates in amplitude, morphing into *stochastic uncertainty* or *internal variability* in climate prediction. Since we can never be rid of it, we could call it the Cockroach Effect. Even that may be misleading because we could reduce roach numbers with pesticides but the stochastic uncertainty is fundamentally irreducible—it will persist even in a perfect climate model. We can estimate the amplitude of stochastic uncertainty by carrying out climate predictions with different initial conditions.

You may not have heard much about stochastic uncertainty because it's not important when predicting global average temperature, which dominates popular discussions of global warming. Predicting societal impacts, or even tipping points, requires prediction of regional climate, which is where stochastic uncertainty becomes important. (If ice sheet instabilities and/or oceanic overturning circulation instabilities turn out to be more important on centennial timescales than currently believed, that will likely increase the amplitude of global chaotic/stochastic uncertainty.)

There are two further uncertainties in climate prediction, and they do affect global average temperature.[79] The next is *scenario uncertainty*. This arises from unpredictable human actions that determine the scenario of future carbon emissions and thus the magnitude of the resulting global warming. This uncertainty cannot be characterized probabilistically and is scientifically irreducible. Even a perfect climate model will exhibit this uncertainty—only human actions (including technological developments) can reduce it. We estimate this uncertainty by carrying out predictions with different emission scenarios.

The third uncertainty in climate prediction is *model uncertainty* which arises from structural differences in the representation of small-scale processes like clouds in climate models. Since these processes occur on scales too fine to be resolved by the coarse spatial grids of the climate, they are represented using approximate formulas known as parameterizations. The errors in these parameterizations lead to spread in predictions using different models. This is the only scientifically reducible error in climate prediction. Using a model with a finer grid, such as a k-scale model, can decrease this uncertainty because fewer processes will be poorly represented. We estimate this uncertainty by carrying out predictions with climate models using different parameterizations.

[79] Three reasons why climate change models are our best hope for understanding the future (TheConversation.com)

7.2 Meta-prediction: Predicting the future of prediction

Analyzing the partitioning between the three different types of uncertainty in our current models allows us to calibrate our expectations for better models. Two important measures of how quickly the globe might warm are transient climate response (TCR) and equilibrium climate sensitivity (ECS). These two measures are basically rough estimates of how much warming doubling of carbon dioxide will cause by the end of this century and over many centuries, respectively. As we see in Figure 1, the spread in these measures has not decreased as the models have gotten "better" over the years. If anything, the ECS spread has increased in recent decades. Figure 2 shows the multi-model average of the global warming projected for three different emission scenarios. The error bars show the model uncertainty for each scenario. Note that the scenario uncertainty is comparable to, or larger than, the model uncertainty.

New let us perform a thought experiment. Suppose we have a future IPCC Assessment Report AR(k) based on a single k-scale model. That means we have a model that predicts climate out to year 2100 using a 1-km spatial grid. As we see in Figure 1, we would have an additional estimate each for TCR and ECS, respectively. But without multiple independent k-scale models, we cannot assess the model uncertainty, i.e., the spread in TCR or ECS. We'd have no way of knowing if the AR(k) estimates are superior in any sense.

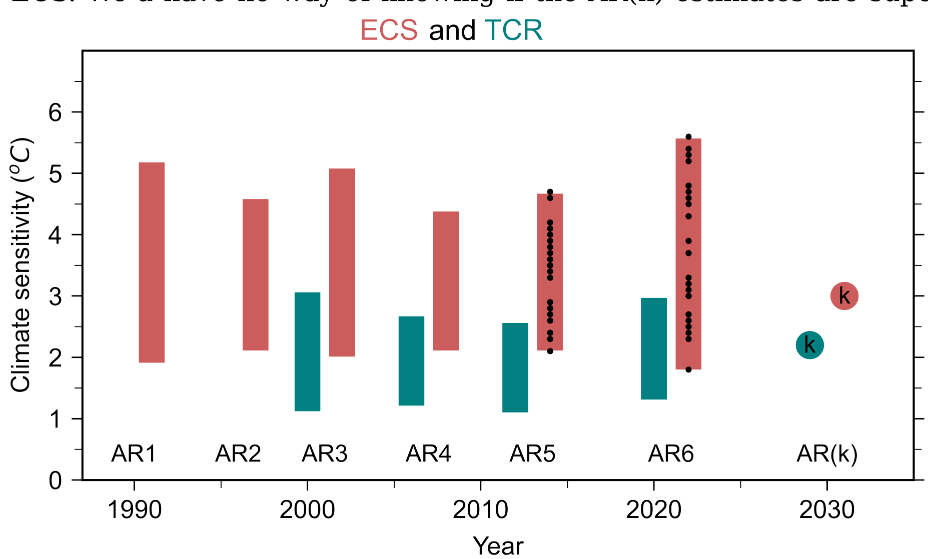


Figure 1. *Model-simulated values of equilibrium climate sensitivity (ECS; red) and transient climate response (TCR; teal) from successive IPCC Assessment Reports from AR1 to AR6. The bars show the spread of values estimated by different models, with black dots showing individual model values for AR5 and AR6. The solid circles show ECS and TCR value assessed for a hypothetical IPCC AR(k) in 2030 using a single k-scale model.* [Adapted from Meehl et al. (2020)][80]

Let us be optimistic and assume further that we are able to afford to run many independent k-scale models for the hypothetical IPCC AR(k) and the spread between these models has reduced by a factor of 2 (say). As we see in Figure 2, the spread in predicted warming by 2100 for different scenarios will become the dominant uncertainty, and will persist even if we had the perfect climate model. Mitigation policy decisions will not benefit very much from reduced model uncertainty or narrower estimates of climate

[80] How we make our 2050 'forecasts', and why we do them (Uk Met Office)

sensitivity, because scenario uncertainty dominates. When it comes to predicting how much the globe will warm by the end of the century, *the biggest uncertainty is us*.[81]

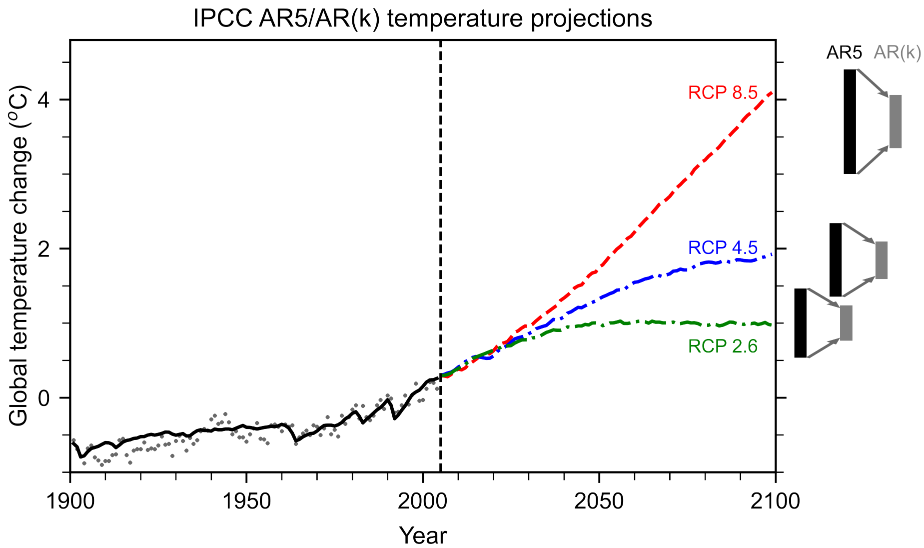


Figure 2. *IPCC AR5 multi-model average prediction of global-average surface temperature for three emission scenarios, high-end (RCP8.5; red), medium (RCP 4.5; blue) and low-end (RCP 2.6; green). The black bars show the AR5 model uncertainty, or the spread amongst models; the gray error bars show what it would look like if the spread was reduced by a factor of 2 by better models in the hypothetical AR(k). (AR5 projections are shown rather than AR6, because AR6 uses model weighting to shrink its larger error bars to resemble AR5 anyway.)* [Adapted from Knutti and Sedlaček, 2013][82]

The dominance of scenario uncertainty for centennial prediction of global temperature is illustrated more vividly by the evolution of the uncertainty partitioning over time (Figure 3a). Scenario uncertainty grows monotonically but is relatively small for the first decade-and-a-half of the prediction, while model uncertainty peaks around that time. Therefore, reducing model uncertainty would have the biggest (fractional) benefit for global predictions on decadal timescales.

[81] In 2020, the @metoffice produced a hypothetical weather forecast for 23 July 2050 based on UK climate projections. Today, the forecast for Tuesday is shockingly almost identical for large parts of the country. [Tweet by @SimonLeeWx](https://twitter.com/SimonLeeWx/status/1547957062000267267) (Twitter)

[82] Ch.2, The Climate Demon: Past, Present, and Future of Climate Prediction (ClimateDemon.com)

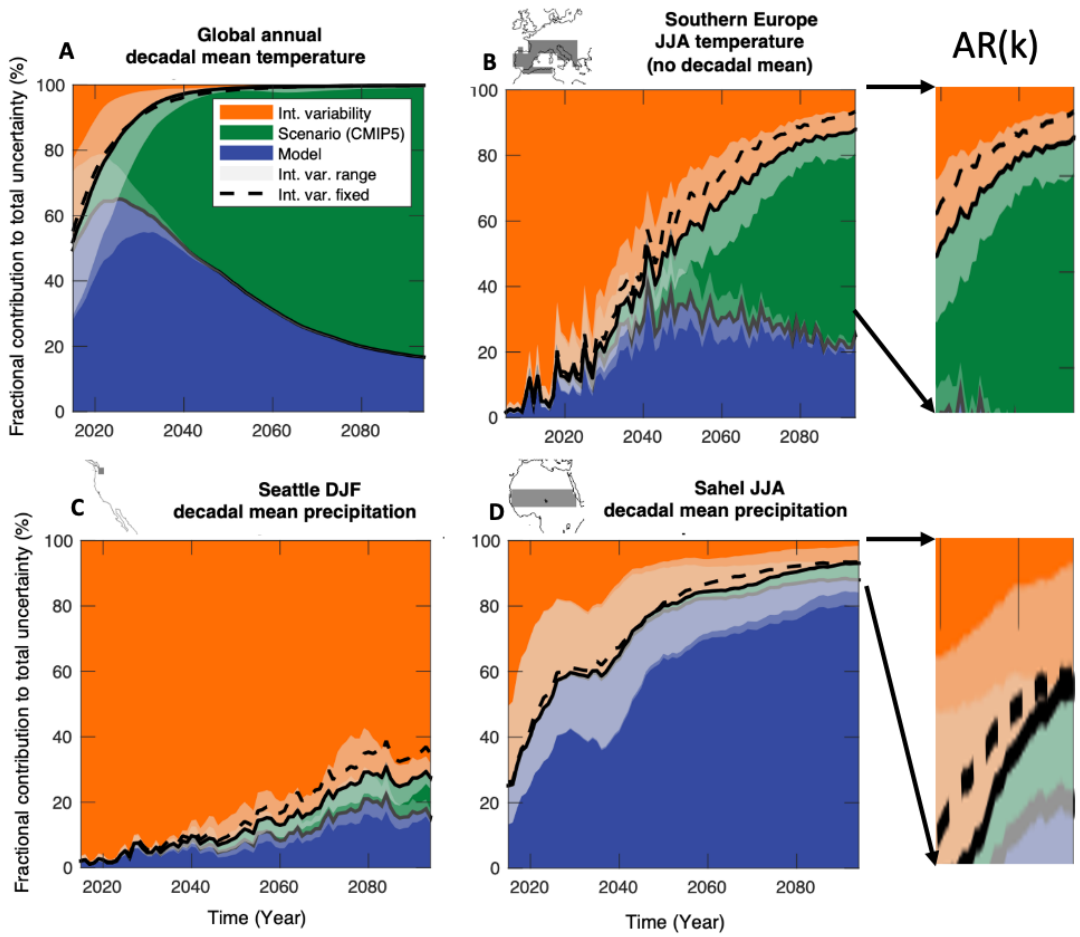


Figure 3. *Partitioning of the uncertainty (stochastic/internal-orange; scenario-green; model-blue) for decadal-average model predictions of: A. Global-average surface temperature; B. summer (Jun-Jul-Aug) temperature over southern Europe (no decadal average); C. winter (Dec-Jan-Feb) precipitation in Seattle, Washington (USA); D. summer (Jun-Jul-Aug) rainfall over the Sahel region of Africa. The lighter shading denotes the higher-order uncertainty in model estimates of stochastic internal variability. If we had a perfect model, the model uncertainty fraction (blue) would vanish, but other uncertainties would remain. The two "blow-ups" on the right illustrate this for a hypothetical AR(k) with greatly reduced model error.* [Adapted from Lehner et al., 2020][83]

Improved prediction of just the global averages is not very useful for assessing societal impacts, which depend on the details of regional climate change. Say we are interested in predicting summer temperatures in southern Europe. The dominant uncertainty is associated with the emission scenario (Figure 3b). Model error accounts for only 30% of the prediction uncertainty. That means even a perfect model would reduce the total uncertainty by no more than 30%. (The regions where we can expect model improvements to provide the most "bang for the buck" are those where model error is the dominant uncertainty and emission scenarios are the second-most important uncertainty, such as over the Southern Ocean.)

[83] How large does a large ensemble need to be? (S. Milinski et al., 2020; Earth System Dynamics)

Next, we consider two regions with contrasting behavior for regional precipitation prediction: the rainy city of Seattle in Washington state, USA and the dry Sahel region of Africa (Figures 3c,d). In both regions, the scenario uncertainty fraction is small, but the model uncertainty fraction is quite different.

If we are interested in predicting Seattle rainfall for the end of the century, current models tell us that better models may not make much of a difference—unpredictable and irreducible stochastic uncertainty accounts for over 70% of the total, meaning that rainfall changes will remain hard to predict (Figure 3c).

Predicting Sahel rainfall for the end of the century tells a different story (Figure 3d). Spread among different models plays a dominant role in the uncertainty. This is the manifestation of a common problem in climate modeling—the large biases in the simulated climate in certain regions. The focus on global average temperature often masks these large regional biases. Higher resolution models would definitely be helpful in reducing these biases.

What if k-scale models were able to substantially reduce the model spread in the Sahel region? Figure 3d suggests that this would cause internal variability to become the dominant uncertainty in the Sahel region. With a better model, Sahel rainfall may still be mostly unpredictable on centennial timescales, but we will be able to say that with more confidence and a much smaller error bar.

We have considered changes in time-averaged temperature and rainfall. But extremes in temperature and rainfall are also very important because they can have severe impacts. Currently, our coarse-resolution climate models cannot predict rainfall extremes very well, because rain is determined by small-scale air motions and microphysical processes. With finer resolution and parameter tuning, k-scale models should be able to do a better job of simulating these extremes in our current climate. The extent to which k-scale models can better predict how rainfall extremes will change in a future climate is an open question—it will depend upon how big a role uncertainties in the still unresolved microphysical processes will play.

7.3 Deconstructing the promise of k-scale

We have outlined what we might expect from better climate models with regard to reducing uncertainty. Now we consider the two recent *Nature Climate Change* articles about k-scale modeling that triggered the blog discussions, one about the atmosphere and the other about the ocean. Their titles are:

1. *Ambitious partnership needed for reliable climate prediction* (ATM)[84]
2. *The small scales of the ocean may hold the key to surprises* (OCN)[85]

As is often the case in climate discussions, ambiguities in language can lead to a mismatch between what the public thinks that science can deliver and what the science is actually capable of delivering. Therefore, it is worth deconstructing what these articles actually say about the benefits of k-scale modeling.

For example, consider the phrase "climate prediction", which appears in the title of the ATM article. Climate scientists use this phrase even for predictions of the average

[84] What to expect when you're expecting a better climate model (Metamodel.blog)

[85] Insights from Earth system model initial-condition large ensembles and future prospects. (C. Deser et al., 2020; Nature Climate Change)

weather for the next season, because climate is the average weather. But the public is more likely to associate "climate prediction" with IPCC and predictions of global warming extending to the end of the century (absent additional qualifiers like "seasonal climate prediction"). This conflates two very different types of prediction: one where initial conditions provide the signal and another where they become the noise.

The ATM article talks about reliable predictions from "daily weather to decadal variability, conditioned by global warming trends". Reliability cannot be assessed for centennial timescale predictions, due to lack of data. Therefore, the article seems to be implicitly focusing more on reducing model biases to improve predictions of El Niño and other phenomena up to the decadal timescale. This is the timescale where reducing model uncertainty will be most beneficial in improving global predictive skill (Figure 3a).

K-scale models should be able to better predict the future statistics of local extreme events on shorter timescales because they can resolve fine-scale fluid motions associated with cloud processes. On longer timescales, however, errors in other non-fluid components of the climate system—such as microphysical processes in clouds or the carbon cycle—will play an increasing role. The direct benefit of k-scale modeling in reducing the uncertainty of centennial climate predictions would therefore be more limited. There would still be the indirect benefit of increasing our confidence in such long-term climate predictions.

The OCN article, on the other hand, does not even mention predictions and instead talks about projections, implying longer timescales. (The title actually refers to "surprises", which is quite the opposite of prediction.) Higher resolution can improve ocean simulation in critical regions that affect possible tipping point behavior associated with Atlantic ocean circulation. Current comprehensive climate models do not exhibit tipping points, but it is possible that higher resolution models could exhibit more nonlinear or threshold behavior. The suggestion, therefore, is that current models could be underestimating oceanic internal variability.

Despite their contrasting views on prediction, the common thread in both articles is the utility of higher spatial resolution to reduce biases in models and improve our understanding of the climate system. This will improve our confidence in climate predictions but should we expect it to significantly reduce the spread in predictions?

Our everyday experience with prediction comes from weather forecasts. We expect that a better weather model using more powerful computers will make predictions with a smaller "error bar". This error bar, which we can calculate using past observations, has indeed decreased over time with better weather models.

Climate prediction is fundamentally different. Since centennial-scale global warming is an unprecedented event, we cannot use observational statistics to compute its error bar.[86] Therefore, the same climate models that make predictions are also used to estimate the spread or the "error bar" associated with their predictions. Better climate models can give us more knowledge because more processes are added or represented better, but the associated error bar could be larger. More knowledge may not always lead to more certainty![87]

---

[86] Without human-caused climate change temperatures of 40°C in the UK would have been extremely unlikely (WorldWeatherAttribution.org)

[87] Factors other than climate change are the main drivers of recent food insecurity in Southern Madagascar (WorldWeatherAttribution.org)

7.4 One model to rule them all?

Contrary to some media headlines, it's not the lack of better models that bedevils climate mitigation policy, but the lack of political will. More computing power for models can help improve the skill of short-term (seasonal-to-decadal) predictions, but that would not be relevant to climate policy.

The understanding gained from better short-term predictions can help improve models used for long-term prediction by reducing biases, especially in their precipitation simulations. Depending upon the relative strengths of internal variability and model error in each region, these improvements may or may not significantly reduce the quantified uncertainty of long-term prediction (Figure 3). Nevertheless, better models would increase our confidence in long-term predictions and provide a sounder basis for climate adaptation policies.

The ATM article recommends spending certain dollar amounts to support k-scale modeling, but doesn't spell out exactly how they should be spent. Should the money be used to build a giant supercomputer associated with a single, international modeling center, or should it be distributed among many centers? Let us consider the former option, i.e., creating the climate-equivalent of CERN, the international facility dedicated to experimental particle physics with an order of magnitude more resources than any national facility:

- *A single k-scale CERN for centennial climate prediction*: This would be a bad idea. Such a Climate-CERN will gain de facto authority because its model will be considered "better" and its climate prediction will be considered official. Since it will contribute only one data point in Figure 1, there'll be no way to estimate the error spread. Of course, the Climate-CERN could develop multiple model structures to estimate the spread. But to do that well, it may require at least 10-20 different model structures. It would be better for these model structures to be developed at separate modeling centers under independent management. (As anyone who has worked at a modeling center could tell you, human factors affect the choice of model structure as much as scientific factors.) Collaboration and standardization of coding structure between multiple modeling centers would certainly be beneficial. Sharing a single supercomputer to run independent models would also be fine.
- *A single k-scale CERN for seasonal-to-decadal prediction*: This could be a good idea, serving as a proof-of-concept for the touted benefits of k-scale modeling. The goals and performance benchmarks of such a SeaDec-CERN would need to be clearly defined, to avoid "mission creep". Limiting predictions to shorter timescales would also prevent the dilution of computing resources. The short-term predictions would provide public benefits, but may not help mitigation or adaptation policies. A SeaDec-CERN may also gain authority because it has a "better" model, but there's a self-correcting mechanism. We'll know soon if the k-scale El Niño forecasts are substantially better than competing models with fewer resources. If they are not, which is quite possible, then SeaDec-CERN will lose its authority. If the forecasts improve substantially, then the knowledge gained can help reduce biases in long-term climate prediction models.
- *Black swans, unknown unknowns, and fundamental research*: We have focused so far on what to expect from better models. But what is unexpected—the "surprises" alluded to in the OCN article—could be more interesting. We know that our climate models are imperfect representations of the complex climate system. In our

climate future, we may encounter a black swan event that was never anticipated or cross a tipping point that was unpredictable. Having the most comprehensive model, *but not necessarily the most complex model*, would help us be better prepared when we encounter unknown unknowns. A good example is the discovery of the Antarctic ozone hole.[88] Without good atmospheric chemistry models that were already available, it would have taken us much longer to understand the mechanism of the ozone hole. Even though these models never predicted the emergence of the ozone hole, they could be modified to predict its future evolution. Model development for the sake of better understanding is typically considered fundamental research, because it does not provide "actionable predictions". Rather than be obsessed with predictions, one can argue that it is important for society to support fundamental climate research as a form of planetary defense—on par with, or even exceeding, other big science projects like space-borne telescopes, planetary missions, and particle accelerators.

Note: As noted in a blog comment, an international center for long-term climate prediction that builds a model at current spatial resolution may be justified for a very different reason. Scientists from developing countries lack the resources to build and use climate models to answer questions that are most relevant to them. Having an international climate modeling center dedicated to their needs would be a great idea!

7.5 Comments

Note: For updated comments, see the original blog post and the anouncement tweet.

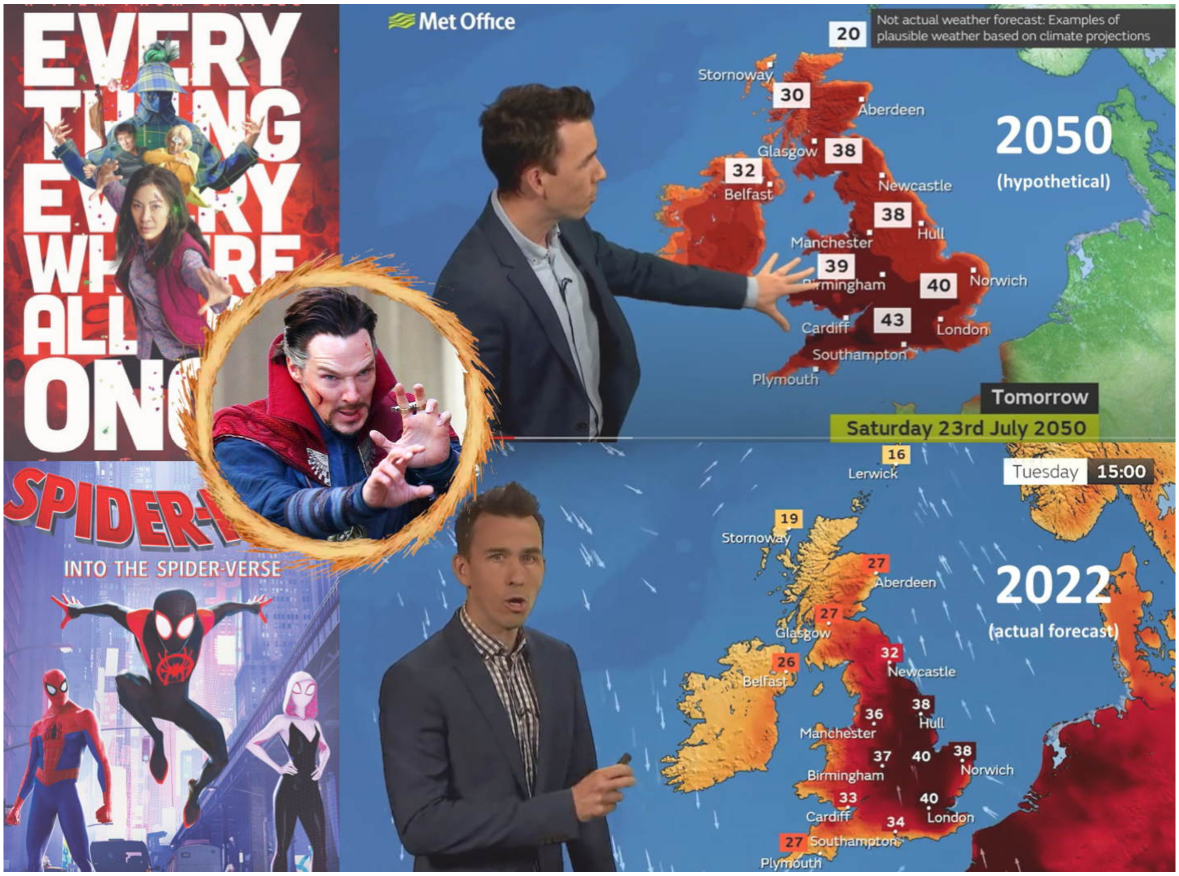
---

8 Strange weather in the multiverse of climate

We cannot predict our weather universe but we can choose our emission multiverse

Metamodel.blog 2022-08-02

[88] Stop blaming the climate for disasters (E. Raju et al., 2022; Communications Earth & Environment), Politics of attributing extreme events and disasters to climate change (M. Lahsen and J. Ribot, 2021; WIREs Climate Change), and It's Not Just Climate: Are We Ignoring Other Causes of Disasters? (Yale Environment 360)



Imagine that our universe is just one slice of bread in the grand cosmic loaf of the multiverse.[89] That's a popular description of the physics concept of the multiverse. But the multiverse is not considered essential for everyday applications of physics, even if it makes for good pop-sci narratives. If one were to use Occam's Razor to slice up the multiverse loaf, one could even argue that the concept of the multiverse adds unnecessary complexity.

Although it may be speculative in physics, the multiverse can be quite useful in understanding climate prediction. We usually define climate as the time average of weather, typically over thirty years or so. When climate itself is changing over that period, this definition becomes less useful. Enter the multiverse.

Imagine that our weather universe is just one slice of bread in the grand loaf of the climate multiverse. The same weather events—like heat waves or hurricanes—occur across the multiverse, but in a different order in each weather universe. We can then define climate as the average across the multiverse. As climate changes over time, the multiverse average also changes.[90] We cannot predict which weather universe we will live in, but we can try to predict the average properties of the multiverse we will live in. This is a complex scientific concept that is often hard to explain to a lay audience. Thankfully, the slew of recent movies about the multiverse, or multiple versions of the universe, may make it easier.

[89] A Physicist Explains Why Parallel Universes May Exist (NPR.org)

[90] In normal climate terminology, we refer to the multiverse as the *ensemble*. We refer to individual universes as *members* of the ensemble. When the climate isn't changing, the time average is equivalent to the ensemble average, according to the ergodic hypothesis. In a changing climate, that is no longer the case.

Although other sci-fi movies have relied on the multiverse before,[91] *Spiderman: Into the Spider-Verse* was the first to use it in its title. If you are into Marvel blockbusters, watching *Spiderman: No Way Home* or *Dr. Strange and the Multiverse of Madness* is good preparation for this blog post about the climate multiverse. If you prefer something more arty (or downright weird), then surviving a viewing of *Everything Everywhere All at Once* may be even better preparation. (After all, climate models have been described as trying to predict everything everywhere all at once.[92])

Not appreciating the multiverse aspect of climate prediction can lead to confusion about the impact of climate change on extreme weather. In July 2022, Britain experienced unprecedented heat waves, with temperatures exceeding 40°C in some locations. Ironically, in 2020, the UK Met(eorological) Office had predicted a similar heat wave as hypothetically occurring in July 2050, using computer models, as part its forecasts from the future program (Figure 1).[93] Does the fact that such a strong heat wave occurred 28 years earlier than "predicted" mean that our climate models are underpredicting the severity of climate change? That is indeed one possible explanation. But there is an alternative explanation—and that involves the multiverse.

---

[91] 14 movies and shows about the multiverse, from 'Spider-Man: No Way Home' to 'Everything Everywhere All at Once' (BusinessInsider.com)

[92] Three reasons why climate change models are our best hope for understanding the future (TheConversation.com)

[93] How we make our 2050 'forecasts', and why we do them (Uk Met Office)

Figure 1 *Top panel shows a hypothetical heat wave forecast for 23 July 2050 (as simulated on a model) that was published in 2020 by the UK Met Office. Bottom panel shows the actual heat wave forecast for 19 July 2022.* [From a tweet][94]

8.1 A multitude of multiverses

If we had a perfect model of the universe and perfect knowledge of its current state, could we predict the future perfectly? Philosophers once believed this was possible, and they named the super-intellect that could make such a prediction as Laplace's De-

[94]In 2020, the @metoffice produced a hypothetical weather forecast for 23 July 2050 based on UK climate projections. Today, the forecast for Tuesday is shockingly almost identical for large parts of the country. [Tweet by @SimonLeeWx](https://twitter.com/SimonLeeWx/status/1547957062000267267) (Twitter)

mon.[95] Laplace's Demon could predict the future of our single universe, and there would be no need to invoke the multiverse. However, quantum uncertainty and classical chaos dashed the prospects of there being a Laplace's Demon, opening the door to the multiverse of predictions.

We only have imperfect models of a subset of the universe, called climate models, and we can never measure the current state of the climate perfectly. Therefore, we can never predict the future perfectly. To account for our imperfect knowledge, we predict the future of a multiverse, rather than our single real universe. The hope is that the set of future predicted universes, the predicted multiverse, includes the future of our real universe.

In climate prediction, we deal with three types of multiverses (Figure 2). The first type is the *weather multiverse*. Since we do not know the initial climate state perfectly, we carry out predictions for several slightly different initial states. Due to the Butterfly Effect of chaos, even minor differences in the initial state will lead to completely different weather conditions after a few weeks, generating the weather multiverse.

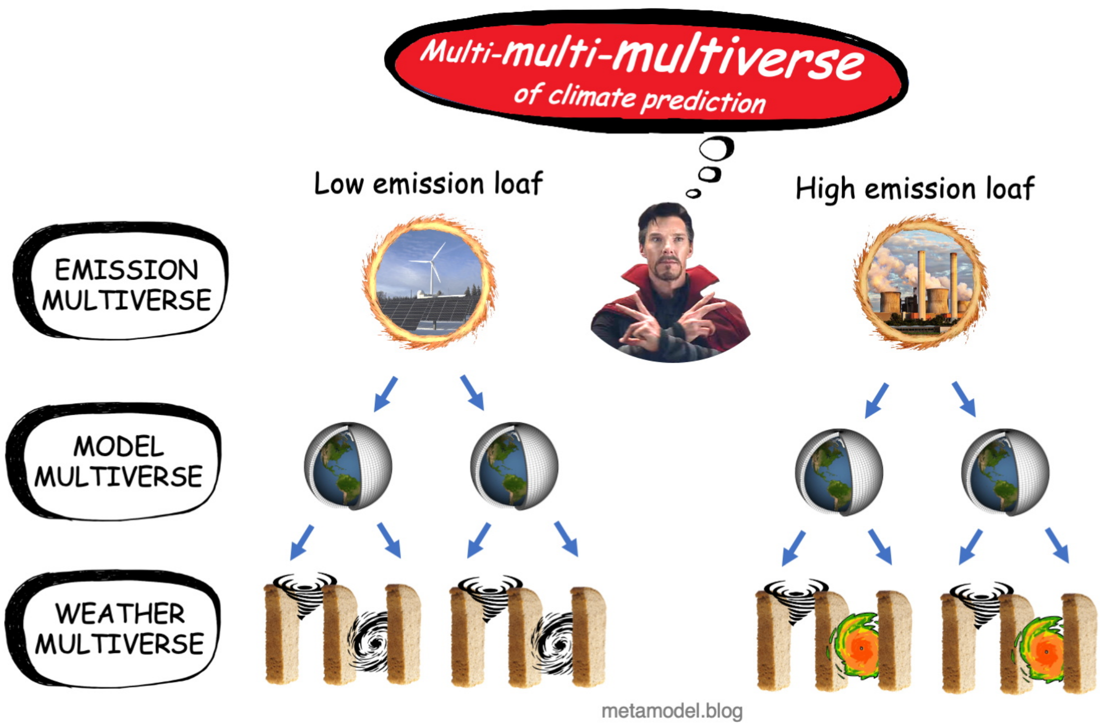


Figure 2 *Three types of multiverses in climate prediction. The bread slices at the bottom represent different predicted universes with random sequences of weather events. Assuming our models are good, the real universe will be one of those slices, but we can never tell exactly which one. By controlling emissions, we select the loaf that the slice will be chosen from. (The color of the hurricane graphic in the high emission loaf indicates that some weather events will be stronger in a warmer world.)*

Say we make a prediction starting from 2020 using a climate model. One predicted universe may have an extreme heat wave (with 40°C temperatures) occurring in July 2050, but another predicted universe may have it occurring in July 2022 (Figure 1). If we simulate only a few predicted universes, then we may miss out on the one where the

[95] Ch.2, The Climate Demon: Past, Present, and Future of Climate Prediction (ClimateDemon.com)

heat wave occurs earlier. This could explain why the UK Met Office made a hypothetical prediction of the extreme heat wave in July 2050, but a real event occurred much earlier. The larger the weather multiverse, the more likely that it includes the real universe. It has been estimated that we may need 50 or more universes in the weather multiverse to adequately span the range of weather variations.[96]

There can be another reason the extreme heat wave occurred earlier in the real universe than in the predicted multiverse. If the climate model is imperfect, and tends to systematically underpredict the warming, then even a larger multiverse may not capture the extreme heat waves. To handle model imperfections, we need another type of multiverse, and we can call it the *model multiverse*. We construct several climate models, each with somewhat different structures for scientific equations. The expectation is that while some models may underpredict the warming, others will overpredict it to compensate. For example, one model may predict that the Arctic will be ice free by 2050 whereas another may predict slower Arctic ice loss. We carry out predictions with different climate models to generate the model multiverse.

There is the need to invoke yet another multiverse type. Our climate models represent just a subset of the universe, because they predict only the physical, chemical and biological aspects of the climate system using scientific equations. But the rest of the universe also affects climate. This includes human activities resulting in carbon emissions. There are no scientific equations to predict human actions a century into the future. So we simply make different sets of plausible assumptions, called *scenarios*, about how humans may behave in the future and then calculate the resulting carbon emissions. Thus we generate the *emission multiverse*, where we predict the future for different carbon emission scenarios.

To top it all, the three types of multiverse are not additive; they are multiplicative (Figure 2). Say there are 50 universes in the weather multiverse, corresponding to different initial states. We may have 20 different equation structures in the model multiverse. We may choose 4 scenarios for the emission multiverse. This means that all the loaves in the grand multi-multi-multiverse of climate will have a total of $50 \times 20 \times 4 = 4000$ slices, each corresponding to a different predicted universe!

8.2 Risk assessment and the multiverse

To properly assess climate risk, we need to consider all three types of multiverses. This can be quite complicated, rather like a cross between the multilayered plot of the movie *Inception* and the multiverse plot of *Everything everywhere all at once*.

Quantitative risk assessment requires assigning probabilities to each universe in a multiverse. For the weather multiverse, we can assume an equal probability or likelihood for each universe, because the memory of the initial state is quickly lost and the distribution becomes random. That's why impact risk assessment using past weather data can be quite accurate up to a decade or so, when climate change effects are still small. We don't need to consider different emission scenarios because the scenarios would not yet have diverged sufficiently. We may still need to consider different models, but global model errors would still be small because they haven't yet had time to build up.

Beyond a few decades, risk assessment gets more complicated because the different emission scenarios diverge and global model errors build up. *Purely probabilistic assessment of risk is no longer possible*, because we cannot assign objective probabilities

[96] How large does a large ensemble need to be? (S. Milinski et al., 2020; Earth System Dynamics)

to the different model or emission multiverses.

For model differences, we can assess the spread of among the models but we cannot assign a specific likelihood to a model universe that is appropriate for all predicted variables. For emissions, we can consider the worst-case scenario, the best-case scenario, and a few scenarios in between. Risk assessments frequently consider just a single, typically the worst-case, emission scenario rather than the full emission multiverse. This can be misleading because it could lead to the worst-case scenario being treated as the most likely scenario, by default.

Often, risk assessments ignore the weather multiverse, even though it is usually the largest of the three multiverses, because it is not important for predicting global average temperature.[97] But accurate risk assessment requires consideration of regional climate change, not just the global averages. Models also continue to exhibit large errors in their simulation of regional climate, underscoring the need for a sufficiently large model multiverse to assess uncertainty. Trimming (or ignoring) the weather/model multiverse types can lead to underestimation of the spread in risk, especially for climate impacts that depend nonlinearly on temperature or rainfall.

8.3 Extreme weather in the multiverse

In recent years, it has become increasingly common to attribute individual extreme and unprecedented weather events, such as heat waves, cold spells, droughts, floods, or hurricanes, to climate change. How do we scientifically make this attribution? To answer that, we need to consider not just whether the event is extreme or unprecedented in our weather universe, but also whether it is so in the multiverse.

Consider five simulated weather universes for the period 1950—2100 using a single climate model for a high emission scenario. Figure 3 shows the predicted occurrence of extreme hot days in Dallas, Texas, during the month of July. We see that the likelihood of extreme hot days increases as global warming continues unabated, but their occurrence is quite irregular among the different universes. Inhabitants in the top universe may be less worried about climate change in 2022, because they experience fewer extreme hot days than inhabitants in the bottom weather universe, although both suffer the same amount of global warming.

[97] What to expect when you're expecting a better climate model (Metamodel.blog)

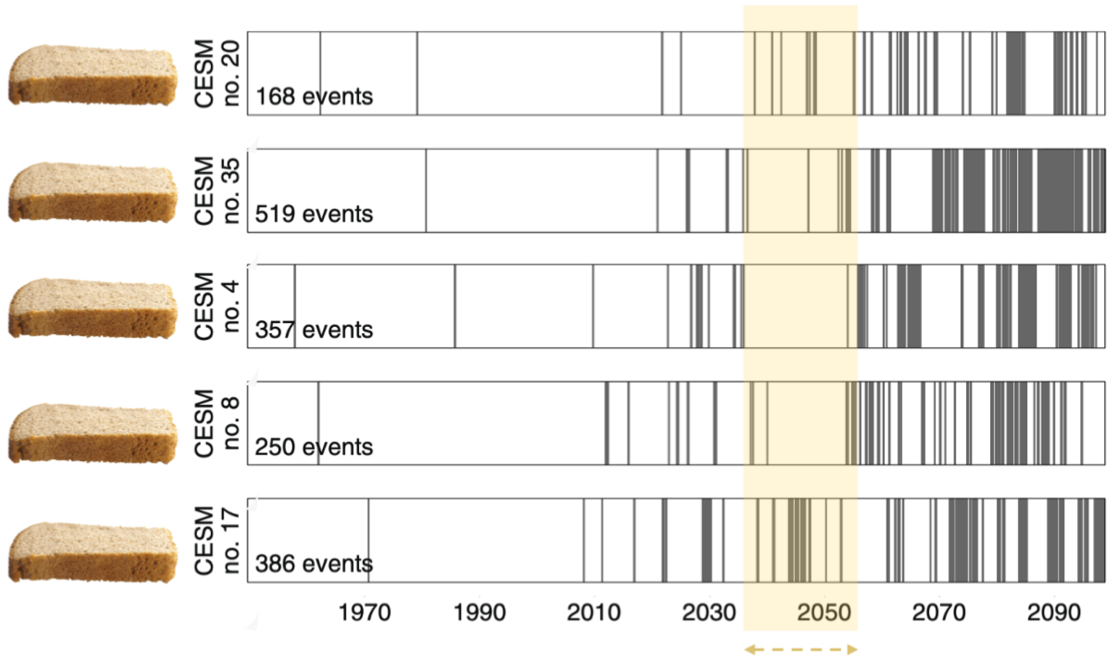


Figure 3 *Occurrence of extreme summer heat in the weather multiverse, with each bread slice denoting a single universe. Vertical bars mark the occurrence of July days that exceed the historical (1950–1999) 99.9th temperature percentile for the model grid box containing Dallas, Texas, in five simulated weather universes of the CESM climate model between 1950-2100, under a high emission scenario (RCP 8.5; now considered implausible). Lightly shaded region denotes the period 2035—2055. (Note that exceeding the monthly 99.9th percentile is roughly a one-in-30-year event before 2000 but happens more frequently later.)* [Adapted from Deser et al., 2020][98]

Note that even a decade from now, between 2035—2055, the middle universe experiences few extreme hot days (Figure 3), which could lead its inhabitants to conclude that global warming isn't affecting Dallas. But the inhabitants of the bottom universe, which experiences many extreme hot days, would draw a different conclusion. This underscores how the randomness of weather can dominate locally, even as the average temperature warms globally.

The rareness and irregularity of extreme events, as illustrated in Figure 3, means that we should carry out careful statistical and modeling analysis before reaching conclusions about the relationship between global warming and local weather. We should not just rely on our personal intuition or experience to draw such conclusions.

There is an international organization of scientists, the *World Weather Attribution* (WWA), that carefully analyzes extreme weather events. The WWA has concluded that global warming makes all heat waves more frequent, as was indeed the case with the 2022 UK heat wave.[99] Rainfall is also becoming more intense, although it is often harder to quantify exactly by how much. For some other types of extreme events, such as droughts, climate change may not always be a major factor.

[98] Insights from Earth system model initial-condition large ensembles and future prospects. (C. Deser et al., 2020; Nature Climate Change)

[99] Without human-caused climate change temperatures of 40°C in the UK would have been extremely unlikely (WorldWeatherAttribution.org)

Climate change did not significantly affect the 2021 drought in Southern Madagascar, according to the WWA, even though some media headlines claimed otherwise.[100] Extreme cold spells are also frequently blamed on climate change, even though the scientific argument for changes in the polar vortex amplifying cold spells is far from settled.

Global warming makes many extreme weather events more frequent and intense. Drawing public attention to climate change by linking it to extreme weather is therefore a good thing. But just as we shouldn't consume too much of a good thing like sugar, we should also be wary of "overattribution" of extreme weather. Reflexively and dramatically blaming every weather-related disaster on climate change can have negative consequences like amplifying climate anxiety and climate fatalism. Attributing disasters primarily to global warming can also divert attention from other, more easily fixable, local socioeconomic vulnerabilities that amplify those disasters.[101] For example, blaming climate change for flooding events can detract from a history of poor urban planning.

To make proper attribution, we need to determine scientifically if an extreme weather event, say event X occurring in 2022, was *significantly* affected by climate change. For unprecedented extreme events, we lack sufficient data to statistically analyze past events similar to X. Therefore, we have to use models. We use one or more climate models to generate two weather multiverses from 1850 to 2022: 1. A *factual weather multiverse* where greenhouse gases increased to their current concentrations from their 1850 pre-industrial values. This multiverse experiences global warming, as recorded in the historical data. 2. A *counter-factual weather multiverse* where we go back in time to 1850 and deliberately hold greenhouse gas concentrations fixed at their pre-industrial values. This multiverse experiences no global warming.

For each weather multiverse, for the year 2022, we count the number of times events similar to event X have occurred in the different universes. If the factual multiverse has many more events similar to X than the counter-factual one with the manipulated timeline, then we can blame global warming for its more frequent occurrence. The larger our multiverse populations and the better our climate models, the more accurately we can assign such blame. (Assigning blame for heat waves is easier than assigning blame for floods or droughts, because models are much better at predicting temperature changes than rainfall changes.)

8.4 Fate and free will in the multiverse

Climate prediction is extremely complex. It differs greatly from many simpler kinds of prediction that you may be familiar with from other disciplines. The pop culture notion of the multiverse allows us to illustrate this complexity, which is often glossed over by those predicting inevitable climate doom with certainty. Predictions with such fateful certainty can only happen in a simplified model universe that does not really belong in the multiverse of comprehensive models.

If you are a decision maker and someone presents you with predictions of future climate or assessments of climate risk, it is worth asking how they handled the three multiverse types. Hopefully, a better understanding of the climate multiverse can help you make more informed decisions in tackling the serious and urgent threat of climate change.

---

[100] Factors other than climate change are the main drivers of recent food insecurity in Southern Madagascar (WorldWeatherAttribution.org)

[101] Stop blaming the climate for disasters (E. Raju et al., 2022; Communications Earth & Environment), Politics of attributing extreme events and disasters to climate change (M. Lahsen and J. Ribot, 2021; WIREs Climate Change), and It's Not Just Climate: Are We Ignoring Other Causes of Disasters? (Yale Environment 360)

We don't have the superpower to choose which weather universe we will live in, because the dice roll of fate makes that choice. We have some power to trim the model multiverse with more research, but progress is not guaranteed.[102] We do have the superpower (i.e., free will) to control which emission multiverse we will live in. If we act to reduce emissions quickly, we will end up with a slightly warmer multiverse with fewer extreme heat waves and heavy rainfall events. If not, we will end up with a much warmer multiverse with many more (and stronger) such events.

8.5 Comments

Note: For updated comments, see the original blog post and the anouncement tweet.

---

[102] What to expect when you're expecting a better climate model (Metamodel.blog)