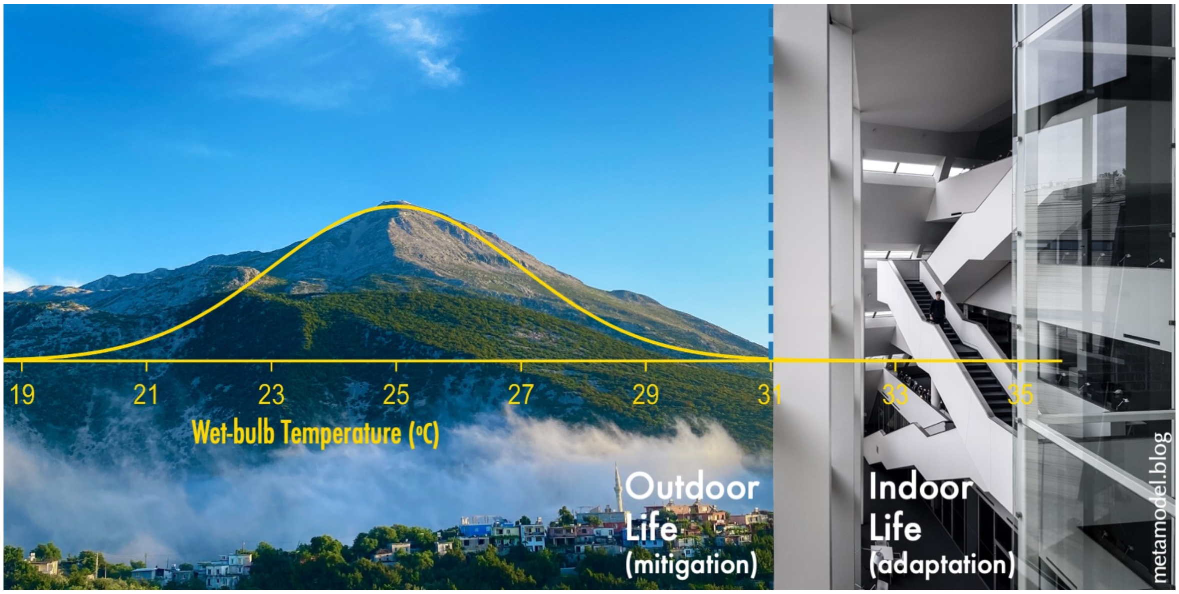
Metamodel.blog Recent posts 2023-03-01

Contents

1 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.[1] 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.[2]

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.[3] 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.[4] 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 flaw in the methodology of the McKinsey report.[5] You can read that tweet thread for

[1] Explaining the Fed's climate test (EEnews.net)

[2] Pilot Climate Scenario Analysis Exercise (FederalReserve.gov)

[3] What to expect when you're expecting a better climate model, Fig. 3 (Metamodel.blog)

[4] Communication of the role of natural variability in future North American climate (C. Deser et al., 2012; Nature Climate Change)

[5] IPCC AR6 WG1 report, Chapter 1: Framing, Context and Methods, p.198

more details and background information. These tweets motivated me to re-examine the report and also learn more about wet-bulb temperature.

1.1 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.[6] 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.[7]

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.

1.2 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. As described in a footnote,[8] the McKinsey report uses daily maximum temperature and daily average relative humidity to compute the wet-bulb temperature. The implicit as-

[6] IPCC AR6 WG1 report, Chapter 11: Weather and Climate Extreme Events in a Changing Climate, p.1588

[7] Tropical cyclones and climate change assessment part II: projected response to anthropogenic warming. (T. Knutson et al., 2019; Bulletin of the American Meteorological Society)

[8] Sandy and Its Impacts (NYC.gov)

sumption 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.

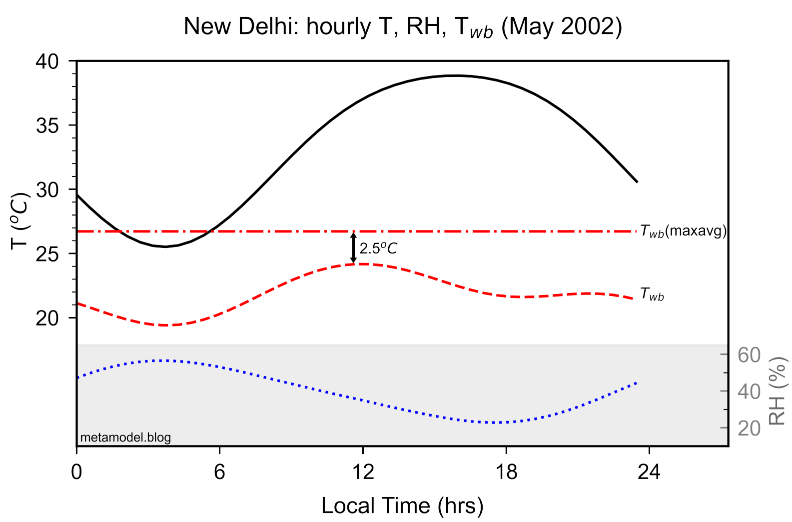


Figure 1. *Hourly surface temperature (black solid) and Relative Humidity (blue dotted) for New Delhi (India) during May 2002. Wet-bulb temperature calculated from hourly data (red dashed).*[9] *Wet-bulb temperature calculated using daily maximum temperature and daily average relative humidity (red dashdot).* [Fourier-smoothed data obtained from Patel et al. 2002 [10]]

When we combine the maximum temperature with average relative humidity, we over-estimate 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 temperature in New Delhi is about 2.5°C.[11] Let us assume that the standard deviation for daily

[9] Strange weather in the multiverse of climate (Metamodel.blog)
[10] The perils of predicting perils: (mis)calculating wet-bulb temperature (Metamodel.blog)
[11] Storylines: an alternative approach to representing uncertainty in physical aspects of climate change (T.G. Shepherd et al., 2018; Climatic Change)

maximum wet-bulb temperature is somewhat less, say 2.0°C.[12] 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.

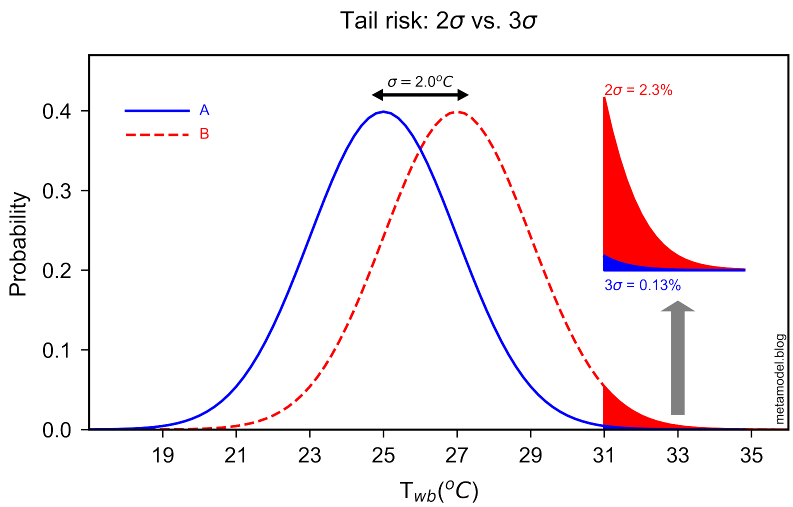


Figure 2. *Assumed normal probability distribution of wet-bulb temperature with standard deviation of 2°C. A. Centered at 25 °C (blue). B. Centered at 27°C (red). Probabilities of wet-bulb temperature exceeding 31°C are shaded, and the shaded areas under the curve are shown as percentages.*

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.

1.3 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

[12] Emissions – the 'business as usual' story is misleading (Z. Hausfather and G.P. Peters, 2020; Nature)

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.[13]
- 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.[14]

1.4 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)

---

2 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

[13] 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)

[14] 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.[15]

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[16] 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?[17]

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.[18] That makes me rather wary of the hype surrounding this new breed of climate startups.[19]

[15] Explaining the Fed's climate test (EEnews.net)
[16] Pilot Climate Scenario Analysis Exercise (FederalReserve.gov)
[17] What to expect when you're expecting a better climate model, Fig. 3 (Metamodel.blog)
[18] Communication of the role of natural variability in future North American climate (C. Deser et al., 2012; Nature Climate Change)
[19] IPCC AR6 WG1 report, Chapter 1: Framing, Context and Methods, p.198

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.

2.1 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.[20] 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.[21]

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.[22] 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 has not gone away in carbon markets.[23]

When used for investment rather than for transactions, a bitcoin is rather like a digital tulip[24]—it has no intrinsic value beyond that determined by the market. It is not the

[20] IPCC AR6 WG1 report, Chapter 11: Weather and Climate Extreme Events in a Changing Climate, p.1588

[21] Tropical cyclones and climate change assessment part II: projected response to anthropogenic warming. (T. Knutson et al., 2019; Bulletin of the American Meteorological Society)

[22] Sandy and Its Impacts (NYC.gov)

[23] Strange weather in the multiverse of climate (Metamodel.blog)

[24] The perils of predicting perils: (mis)calculating wet-bulb temperature (Metamodel.blog)

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.[25]

2.2 Carbon offsets

Many carbon offsets, especially the cheap options that businesses like airlines offer,[26] remind me of the famous P. T. Barnum quote: "There's a sucker born every minute".[27] 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,[28] 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 [29] or even negative.[30] 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 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.

[25] Storylines: an alternative approach to representing uncertainty in physical aspects of climate change (T.G. Shepherd et al., 2018; Climatic Change)

[26] Emissions – the 'business as usual' story is misleading (Z. Hausfather and G.P. Peters, 2020; Nature)

[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)

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

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.

2.3 Comments

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

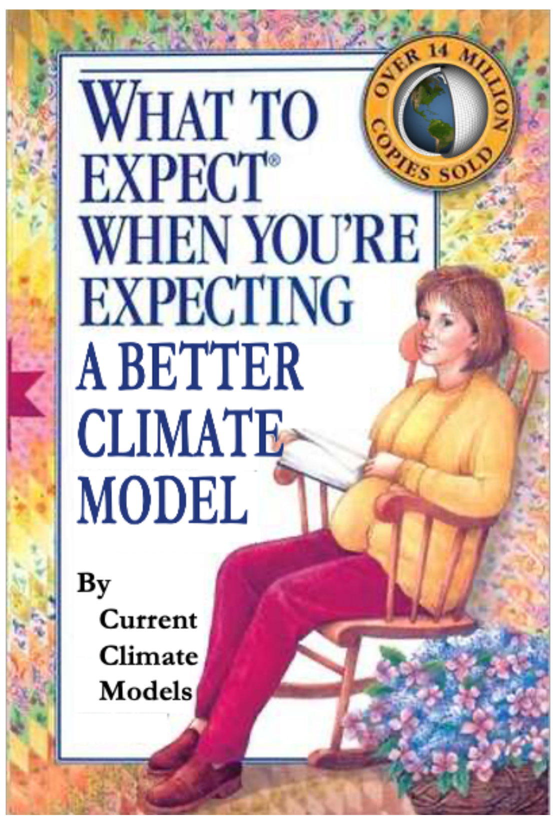
- *R Saravanan*:
  John Oliver's hard-hitting take on carbon offsets:
  Carbon Offsets: Last Week Tonight with John Oliver (YouTube video)

---

3 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*,[31] 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.[32] 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.[33]

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.

3.1 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.)

[31] Explaining the Fed's climate test (EEnews.net)
[32] Pilot Climate Scenario Analysis Exercise (FederalReserve.gov)
[33] What to expect when you're expecting a better climate model, Fig. 3 (Metamodel.blog)

There are two further uncertainties in climate prediction, and they do affect global average temperature.[34] 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.

3.2 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

[34] Communication of the role of natural variability in future North American climate (C. Deser et al., 2012; Nature Climate Change)

sense.

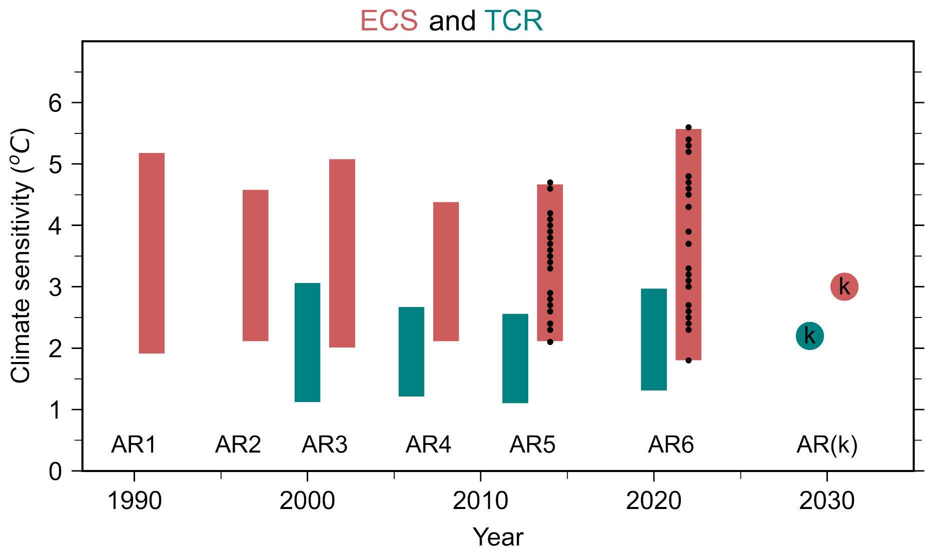


Figure 1. *Model-simulated values of equilibrium climate sensitivity (ECS; red) and transient climate response (TCR; teal) from successive IPCC Assessment Reports from AR1 to AR6. The bars show the spread of values estimated by different models, with black dots showing individual model values for AR5 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)][35]

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 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*.[36]

[35] IPCC AR6 WG1 report, Chapter 1: Framing, Context and Methods, p.198
[36] IPCC AR6 WG1 report, Chapter 11: Weather and Climate Extreme Events in a Changing Climate, p.1588

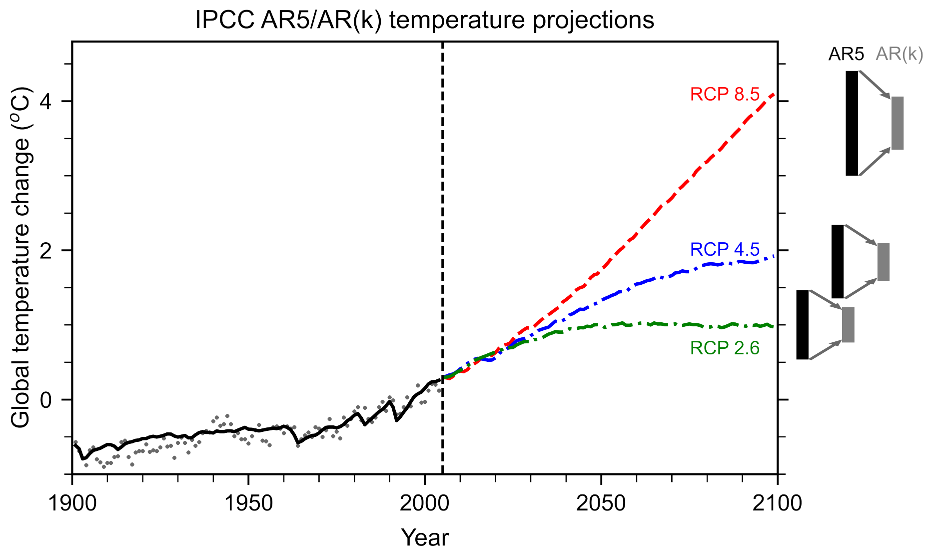


Figure 2. *IPCC AR5 multi-model average prediction of global-average surface temperature for three emission scenarios, high-end (RCP8.5; red), medium (RCP 4.5; blue) and low-end (RCP 2.6; green). The black bars show the AR5 model uncertainty, or the spread amongst models; the gray error bars show what 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][37]

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.

[37] Tropical cyclones and climate change assessment part II: projected response to anthropogenic warming. (T. Knutson et al., 2019; Bulletin of the American Meteorological Society)

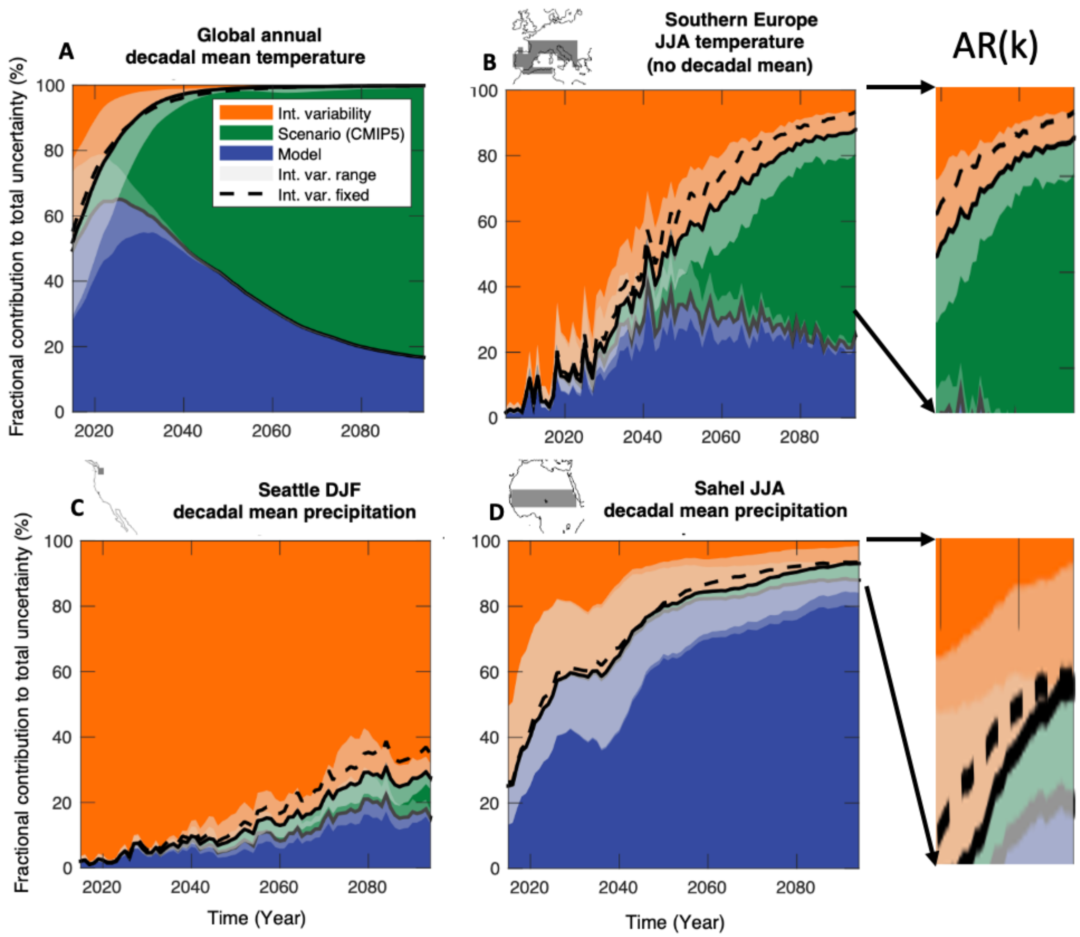


Figure 3. *Partitioning of the uncertainty (stochastic/internal-orange; scenario-green; model-blue) for decadal-average model predictions of: A. Global-average surface temperature; B. summer (Jun-Jul-Aug) temperature over southern Europe (no decadal average); C. winter (Dec-Jan-Feb) precipitation in 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][38]

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.)

[38] Sandy and Its Impacts (NYC.gov)

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.

3.3 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)[39]
2. *The small scales of the ocean may hold the key to surprises* (OCN)[40]

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 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

[39]Strange weather in the multiverse of climate (Metamodel.blog)
[40]The perils of predicting perils: (mis)calculating wet-bulb temperature (Metamodel.blog)

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.[41] 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![42]

3.4 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

[41] Storylines: an alternative approach to representing uncertainty in physical aspects of climate change (T.G. Shepherd et al., 2018; Climatic Change)

[42] Emissions – the ‘business as usual’ story is misleading (Z. Hausfather and G.P. Peters, 2020; Nature)

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.[43] 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!

3.5 Comments

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

- *R Saravanan*:
  Important paper about uncertainty in predicting climate extremes:
  For precip extremes, "past observations … can provide almost as accurate a picture of future extreme occurrences as even the best … climate models.."
  "large ensembles should become the standard" for models
  Tweet thread

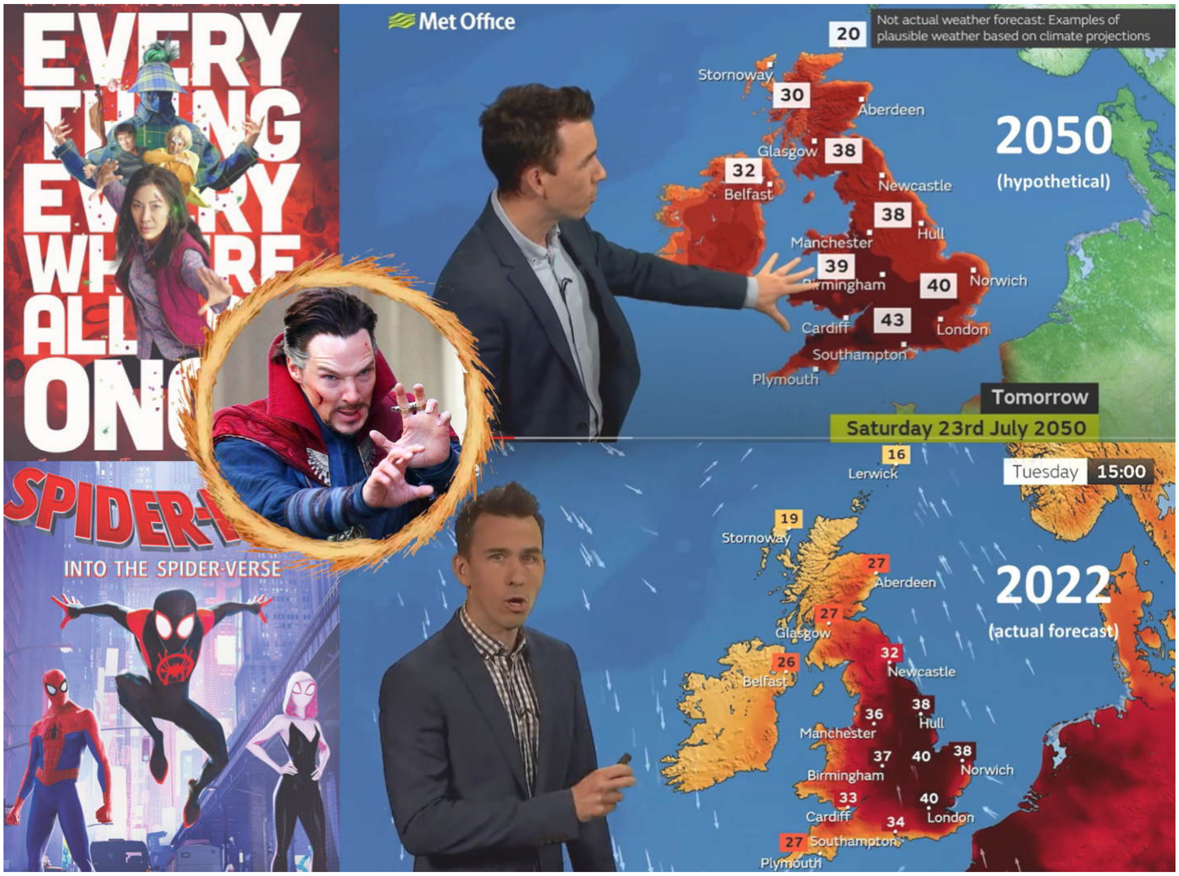
---

4 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

[43] 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.[44] 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.[45] 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.

[44] Explaining the Fed's climate test (EEnews.net)

[45] Pilot Climate Scenario Analysis Exercise (FederalReserve.gov)

Although other sci-fi movies have relied on the multiverse before,[46] *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.[47])

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).[48] 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.

[46] What to expect when you're expecting a better climate model, Fig. 3 (Metamodel.blog)

[47] Communication of the role of natural variability in future North American climate (C. Deser et al., 2012; Nature Climate Change)

[48] IPCC AR6 WG1 report, Chapter 1: Framing, Context and Methods, p.198

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][49]

4.1 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 Demon.[50] Laplace's Demon could predict the future of our single universe, and there

[49] IPCC AR6 WG1 report, Chapter 11: Weather and Climate Extreme Events in a Changing Climate, p.1588

[50] Tropical cyclones and climate change assessment part II: projected response to anthropogenic warming. (T. Knutson et al., 2019; Bulletin of the American Meteorological Society)

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.

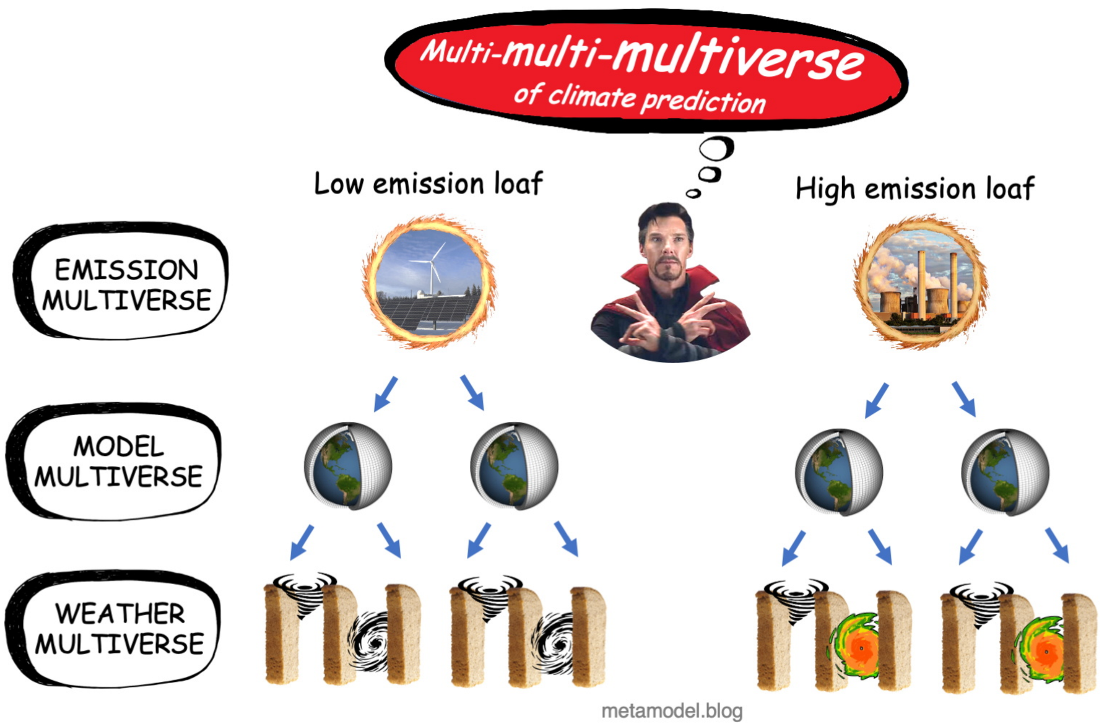


Figure 2 *Three types of multiverses in climate prediction. The bread slices at the bottom represent different predicted universes with random sequences of weather events. Assuming our models are good, the real universe will be one of those slices, but we can never tell exactly which one. By controlling 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 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.[51]

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!

4.2 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 to the different model or emission multiverses.

[51] Sandy and Its Impacts (NYC.gov)

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.[52] 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.

4.3 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.

[52] Strange weather in the multiverse of climate (Metamodel.blog)

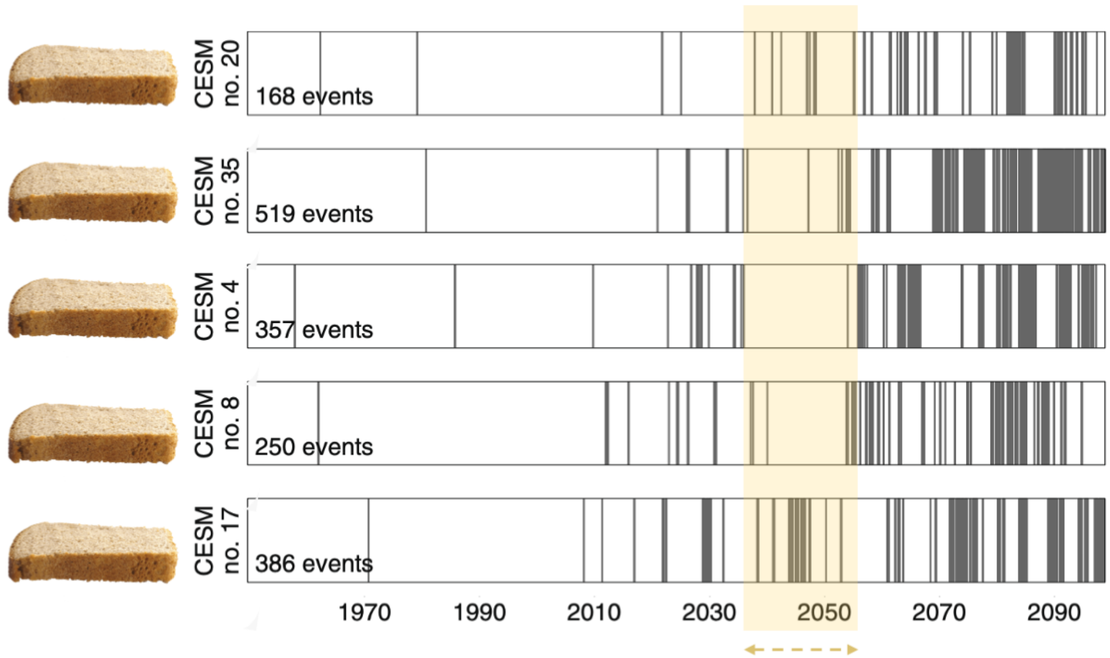


Figure 3 *Occurrence of extreme summer heat in the weather multiverse, with each bread slice denoting a single universe. Vertical bars mark the occurrence of July days that exceed the historical (1950–1999) 99.9th temperature percentile for the model grid box containing Dallas, Texas, in five simulated 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][53]

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.[54] 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.

Climate change did not significantly affect the 2021 drought in Southern Madagascar,

[53] The perils of predicting perils: (mis)calculating wet-bulb temperature (Metamodel.blog)
[54] Storylines: an alternative approach to representing uncertainty in physical aspects of climate change (T.G. Shepherd et al., 2018; Climatic Change)

according to the WWA, even though some media headlines claimed otherwise.[55] 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.[56] 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.)

4.4 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.

We don't have the superpower to choose which weather universe we will live in, because

[55] Emissions – the 'business as usual' story is misleading (Z. Hausfather and G.P. Peters, 2020; Nature)

[56] 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)

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.[57] 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.

4.5 Comments

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

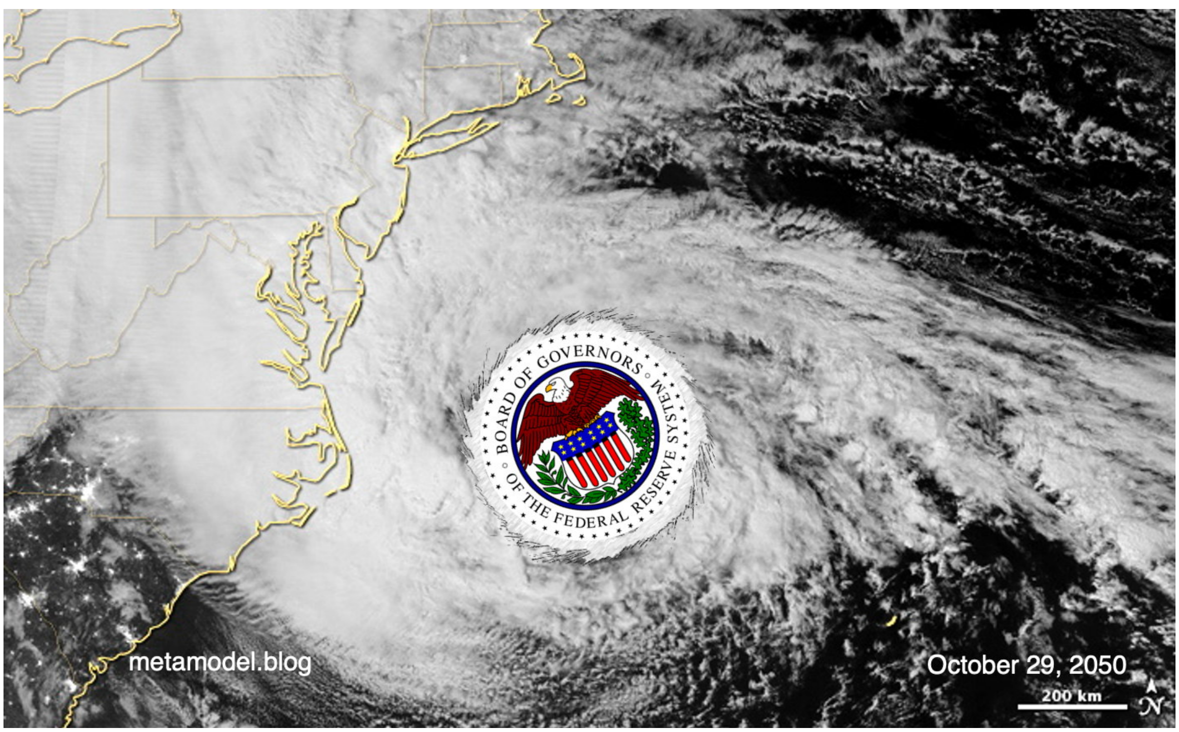
- *R Saravanan*:
  Interesting blog post on The heatwave that never was
  Also discussed in this tweet thread

---

5 Hurricane Fed: The New Climate Stress Test for Banks

The Fed's new hurricane-based risk assessment is well-intentioned but poorly formulated. Since future hurricane probabilities are hard to predict, a simple storyline approach would have been better.

Metamodel.blog 2023-03-01



[57] Strange weather in the multiverse of climate (Metamodel.blog)

Recently, the US Federal Reserve ("Fed") issued guidance for a pilot exercise on how the six largest US banks should analyze their exposure to climate risk.[58] The Fed calls the study exploratory and says the results will have no capital implications. Nevertheless, this exercise is essentially a climate stress test—like the financial stress tests to check whether banks have enough capital to survive economic shocks. Other central banks around the world are taking similar steps. Climate change can lead to increased economic risk. Planning to deal with this risk, taking into account all the uncertainties, is prudent for the Fed and other central banks to advocate.

The Fed proposes a probabilistic assessment of the risk of an extreme hurricane event making landfall in the Northeast US in 2050 under two different carbon emission scenarios: a medium emission scenario, RCP4.5 and a high-emission scenario, known as RCP8.5. This sounds straightforward at the surface, but when you dig deeper the proposed assessment turns out to be quite complicated, involving many assumptions.

Predicting future climate and economic conditions requires using complicated models and scenarios. However, this doesn't mean that risk assessment needs to be framed in a convoluted fashion. When we ask bankers (and other lay people) to test for climate risk, it would be best provide guidance in the simplest terms possible. The Fed's risk estimation exercise fails the stress test of simplicity. As explained later in this article, a better way for banks to assess climate risk would be to use a simple "storyline approach" with a few well-defined assumptions about future climate change.

5.1 Problems with the Fed's guidance

In January 2023, the Fed issued instruction for the *Pilot Climate Scenario Analysis Exercise*[59], which asks the banks to assess physical climate risk in the following manner:

> [F]or the iterations of the common shock component, participants should estimate the impact of a hurricane event(s) within the Northeast region with the following characteristics:
>
> 1. Climate conditions broadly consistent with possible future climate conditions in 2050 as characterized by the SSP2-4.5 (or RCP 4.5) pathways with a 100-year return period loss...
> 2. Climate conditions broadly consistent with possible future climate conditions in 2050 as characterized by the SSP5-8.5 (or RCP 8.5) pathways with a 200-year return period loss...
> 3. Climate conditions broadly consistent with possible future climate conditions in 2050 as characterized by the SSP5-8.5 (or RCP 8.5) pathways with a 200-year return period loss...
>
> To estimate the impact of the hurricane event(s) in 2050 across the three iterations above, participants may need to make additional assumptions around the state of climate and the related chronic physical features in 2050, including, but not limited to, an increase in surface temperatures, sea level rise, and precipitation levels.

Taken at face value, this sounds like the right approach. The Fed's reasoning appears to rest on the following chain of assumptions: *(i) risk assessments should be based on probabilities; (ii) extreme weather is expected to worsen with climate change; (iii)*

[58] Explaining the Fed's climate test (EEnews.net)
[59] Pilot Climate Scenario Analysis Exercise (FederalReserve.gov)

considering climate in a specific region rather than global average climate makes the assessment relevant to the US; (iv) decadal timescales are more important for financial risk than centennial timescales; and (v) different emission scenarios allow us to span the range of policy responses.

Most people's intuition about climate change is based on the popular discourse on climate change, which often focuses on the global average temperature increasing over centuries. There is a clear separation of the global warming signal on centennial timescales as predicted by climate models for different emission scenarios—the globe will be much warmer under RCP8.5 than RCP4.5 by the year 2100. There is always random noise associated with unpredictable ("stochastic") variations associated with weather, but it is relatively small when temperature is averaged globally, because of cancellations amongst regional variations.[60]

However, nobody lives in "global-average-land"! While global warming thresholds of 1.5 (or 2.0)°C dominate newspaper headlines, what will affect your life is the warming in the region you live in. In the middle latitudes, the regional greenhouse gas warming signal remains roughly the same as the global warming signal, but the random noise will be a lot stronger, because the cancellation benefit of global averaging is lost. That makes it much harder to discern the regional warming signal. If we are interested in short-term warming (by year 2050, say), the warming signal becomes even weaker whereas the noise amplitude remains the same (Figure 1).

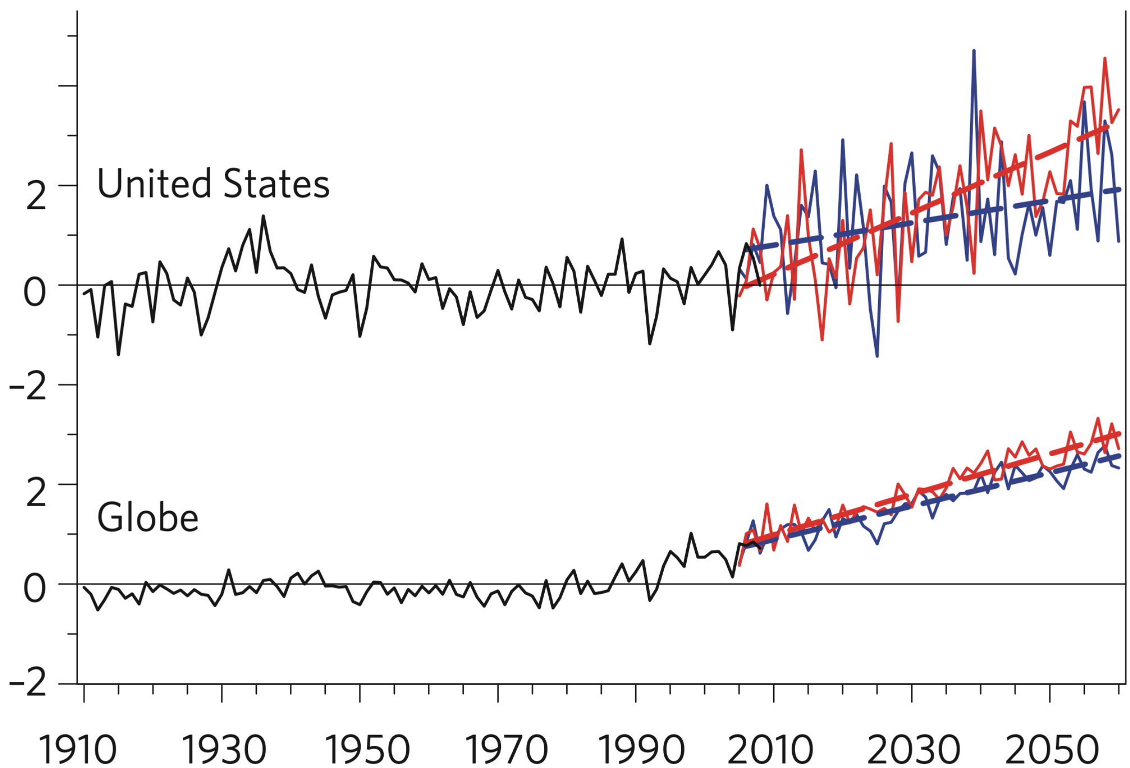


Figure 1 *Temperature change averaged over the globe and over the contiguous United States in observations until 2008 (black) and for a single emission scenario (A1B; similar to RCP6.0) until 2060, for the weather realization with the largest (red) and smallest (blue) future trends. The smaller the averaging region, the larger the random*

[60]What to expect when you're expecting a better climate model, Fig. 3 (Metamodel.blog)

noise due to weather. [Adapted from Figure 2b of Deser et al., 2020][61]

The popular discourse on climate change also leaves the impression that just because we have models that predict future global warming, we can compute precise probabilities of all future extreme events like heat waves, heavy rainfall or hurricanes. Climate models simulate different phenomena with differing degrees of realism. Large-scale warming patterns and heat waves are simulated better than small-scale rainfall events. Uncertainties associated with different climate models make it hard to estimate precise future probabilities of regional extreme rainfall, for example (Figure 2).

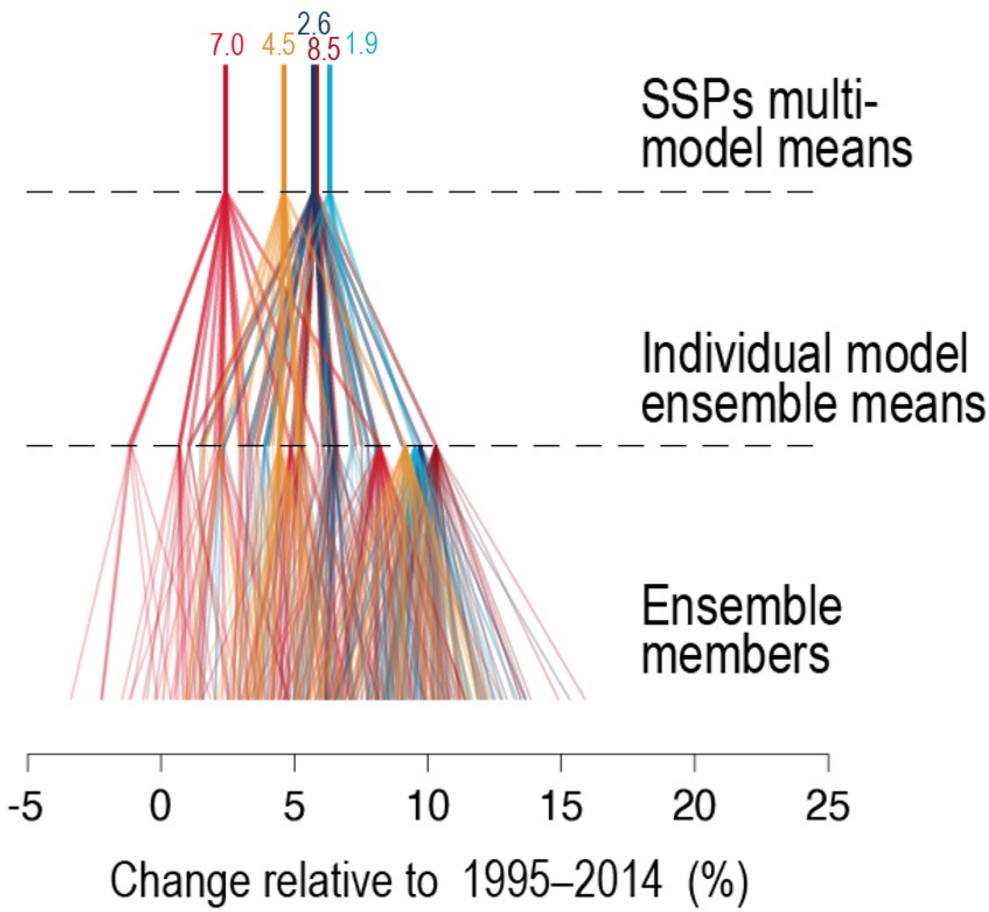


Figure 2 *Cascading uncertainties in IPCC model projections of East Asia summer (June-August) rainfall for multiple emission scenarios (SSP 1.9, 2.6, 4.5, 7.0, 8.5) for years 2041-2060. Top row shows average for each SSP; next row shows the average for individual models; bottom row shows the rainfall for each individual realization of weather in each model (representing the random variability). It is hard to distinguish between emission scenarios 4.5 and 8.5 when looking at the spread of simulated rainfall in the year 2050. Interestingly, the lowest emission scenario, 1.9, shows the largest rainfall changes in the near term due to the effect of aerosols in that scenario.*

[61] Communication of the role of natural variability in future North American climate (C. Deser et al., 2012; Nature Climate Change)

[Adapted from Figure 1.15 of IPCC AR6 WG1 report][62]

With the above caveats in mind, we focus on weak links in the chain of assumptions (i)—(v) that underlie Fed's guidance.

(i) *Risk assessments should be based on probabilities*. By specifying numeric probabilities (e.g., 100-year return period loss), the Fed aims to make the risk assessment precise. That would work if we knew the precise probabilities of hurricanes making landfall in the Northeast US in the year 2050. But we don't. Our current global climate models have too coarse spatial resolution to estimate the precise probability of future hurricane landfalling events. The current horizontal grid of climate models is at best about 50kmX50km—not enough to resolve the eyewall of a hurricane. Making additional assumptions and using simpler/regional models, we can come up with numbers for future hurricane probabilities, but the answers will be sensitive to the assumptions.[63]

(ii) *Extreme weather is expected to worsen with climate change*. This statement is generally true, but the devil is in the details. Large-scale weather extremes like heat waves will uniformly get worse but the picture is more complicated for hurricanes because they involve small-scale moist processes. According to our current scientific understanding, the strength of hurricanes and the associated rainfall are expected to increase but their total number may actually decrease (Figure 3). The strongest hurricanes are expected to get even stronger and may increase in number. If the total number of hurricanes does not increase, it implies that weaker hurricanes, such as Category 1 or 2, will decrease in number.

[62] IPCC AR6 WG1 report, Chapter 1: Framing, Context and Methods, p.198

[63] IPCC AR6 WG1 report, Chapter 11: Weather and Climate Extreme Events in a Changing Climate, p.1588

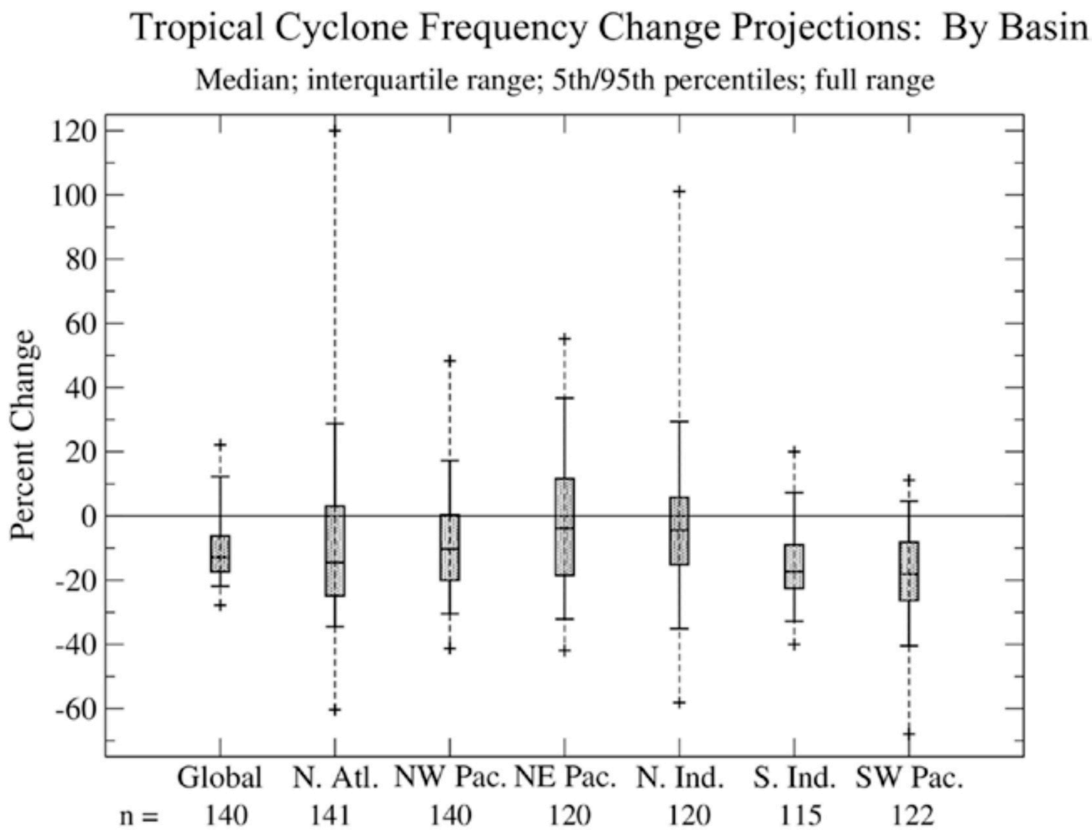


Figure 3 *Projected changes in the frequency of tropical cyclones (known as hurricanes in the North Atlantic) for each ocean basin for 2°C of additional global warming compared to current conditions. Hurricane frequency in the North Atlantic is projected to decline about 15% on the average, but the uncertainty range is huge!* [Adapted from Fig. 1b of Knutson et al., 2020][64]

The Fed's choice to focus on an extreme hurricane affecting the Northeast US is presumably motivated by an actual natural disaster, Hurricane Sandy, that wreaked havoc in that region in October 2012, killing hundreds of people and inflicting many tens of billions of dollars in damage.[65] But Sandy was not very strong, as hurricanes go; it barely reached a peak intensity of Category 3 and made landfall as a Category 1 storm. Bankers may be surprised to learn that climate change might actually make weak storms like Sandy rarer in the future.

Of course, a stronger hurricane than Sandy could impact the Northeast in the near future. Such a rare event could happen purely by chance, even in the absence of additional warming between now and 2050. Future global warming could be responsible for amplifying the storm, but the distinction between different emission scenarios may not be very discernible by the year 2050.

(iii) *Considering climate in a specific region rather than global average climate makes*

[64]Tropical cyclones and climate change assessment part II: projected response to anthropogenic warming. (T. Knutson et al., 2019; Bulletin of the American Meteorological Society)

[65]Sandy and Its Impacts (NYC.gov)

the assessment relevant to the US. Yes, but the regional focus also greatly increases the strength of random (stochastic) noise as discussed earlier. To probabilistically assess regional climate change for a small area like the Northeast US, we would need a large ensemble of climate model simulations to quantify the random noise[66] — something the Fed's guidance fails to note.

(iv) *Decadal timescales are more relevant for financial risk than centennial timescales*. True, but focusing on shorter timescales means the global warming will be weaker and less sensitive to the scenario being considered. Coupled with the much larger amplitude of random variability, as noted in the (iii) discussion, there may be no point in trying to distinguish between the signals of RCP8.5 vs. RCP4.5 by 2050, as the Fed recommends.

(v) *Different emission scenarios allow us to span the range of policy responses*. Yes, but it would become a moot issue with the lower signal-to-noise ratio, as noted in the discussion of points (iii) and (iv) above.

The Fed's guidance may sound well-defined and numerically precise on the surface, but it isn't quite so. The numerical precision of the specified return period loss becomes irrelevant if our estimates of the probability of the physical hazard, i.e., future landfalling hurricanes, are themselves imprecise. Since banks are free to make numerous additional assumptions needed to estimate the physical hazard, the guidance is poorly formulated. Even seemingly small errors in these assumptions can lead to big errors in the estimation of tail risk of the physical hazard.[67] The uncertainties associated with the assumptions can be lost in translation, leading to faux precision in the final risk assessment.

5.2 The super-Sandy storyline

Since climate risk is real and important, is there is a better way to assess its impact on the banking sector? One that does not involve hard-to-compute probabilities and a multitude of assumptions? Instead of hiring consultants to make many small assumptions involving a cascade of models that we can't keep track of, can't we just make a few big assumptions? Wouldn't it be more transparent if we are upfront about the uncertainties?

An alternative, and simpler, way to frame the risk assessment is to start with a known extreme event, say Hurricane Sandy that affected the Northeast US in 2012, and ask how a stronger version of this storm occurring in the near future would affect bank finances. Such an approach, often referred to as a "storyline", is better suited to describing future high-impact low-likelihood events whose probabilities are hard to quantify.[68]

Scientific research shows that global warming amplifies the water cycle, meaning that extreme storms like Sandy can become more intense as the atmosphere becomes moister. One can assess the financial risks of a stronger "super-Sandy" hurricane, say 10% stronger than Sandy, making landfall in the Northeast US. Such an assessment should also include estimates of higher sea-levels by the year 2050, which would amplify the coastal impacts.

The storyline framing possesses several advantages. It sidesteps the contentious issue

[66] Strange weather in the multiverse of climate (Metamodel.blog)

[67] The perils of predicting perils: (mis)calculating wet-bulb temperature (Metamodel.blog)

[68] Storylines: an alternative approach to representing uncertainty in physical aspects of climate change (T.G. Shepherd et al., 2018; Climatic Change)

of whether the high-emission RCP8.5 scenario—required by the Fed—is even a plausible future.[69] According to the latest international climate assessment from the IPCC, recent trends in the energy sector mean that the likelihood of the RCP8.5 scenario is low. The storyline framing focuses on the narrower question of whether a super-Sandy storm impacting the Northeast US by 2050 is a plausible event. The combination of climate change and random variability can make such an event plausible. (If need be, we can consider different plausible strengths of super-Sandy to span the range of physical risk.)

The current Fed proposal allows the use of *a la carte* assumptions to assess climate risk. Banks may end up making different modeling assumptions. A smorgasbord of assumptions will mean that comparing risk assessments from different banks will be like comparing apples to oranges. A simpler alternative is to assess financial impacts using a well-defined storyline with fewer assumptions. That can provide a clearer picture of the climate risk faced by banks.

Keep it simple, Fed!

(*Top image shows a NASA satellite view of Hurricane Sandy approaching the Northeast US on October 29, 2012, with the Federal Reserve logo in the eye of the storm.*)

5.3 Comments

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

---

[69] Emissions – the 'business as usual' story is misleading (Z. Hausfather and G.P. Peters, 2020; Nature)