NOTES

ADMITTING EVIDENCE OF CLIMATE CHANGE UNDER *DAUBERT*: CLIMATE EXPERTS AS RELIABLE, HYPER-QUALIFIED TECHNICIANS

EDMUND H.S. BROSE*

Climate change is here. Anthropogenic warming is currently increasing temperatures, the devastation of storms, and the incidence of droughts. If humanity continues on its current path, the next fifty years will see millions die due to extreme weather events, along with a drastic increase in the number of climate refugees seeking haven. In the face of this crisis, government inaction at all levels has fueled the flames. Private actors and state and municipal governments have stepped into the breach, bringing suits against polluters for the harms to their localities and citizens. The challenge that this Note seeks to address is how to take these dire predictions of the future, and damages of the present, and translate them into workable, reliable legal evidence that can be used in a court of law. While most courts have declined to allow suits to proceed on threshold questions, they will soon have to deal with scientific evidence of climate change as these suits grow more numerous and the plaintiffs more resourceful.

This Note serves as a plea to judges to approach climate modeling methods in the same way they approach comparable types of evidence. Under Federal Rule of Evidence 702, climate science should be admissible as sufficiently reliable, by a preponderance of the evidence. Climate science exists somewhere between pure science and specialized expert knowledge, due to the subjective nature of the discipline. While climate science may not be seen as sufficiently "scientific," if climate scientists are considered a group of experts, the discipline should easily pass muster under lower court interpretations of the Supreme Court's Kumho Tire *decision. By comparing climate science to criminal forensic methods, the case for admissibility becomes obvious. Thus, if judges take their roles seriously as neutral, consistent referees of justice, the admissibility of climate science should not be a serious hurdle for plaintiffs.*

 J.D., 2023, New York University School of Law; B.A., 2017, Columbia University. I am grateful to Judge Jed S. Rakoff for his guidance and good humor through the drafts of this Note, and to Professor Daniel J. Capra for inspiring an interest in evidence. Thank you to the phenomenal editors of the *New York University Law Review*, especially Soumya Kandukuri, whose cheery support was invaluable in getting this Note to publication, and Jonathan Goldberg, for all his thoughtful comments and insights into strengthening the piece. I am always grateful for my loving family, especially my parents, and for my niece Cece.

INTRODUCTION: LAWSUITS TO MITIGATE THE CLIMATE EMERGENCY

In 2015, a group of American youth plaintiffs made international headlines by suing the United States government. They claimed that

the government had failed to preserve the environment for their generation and future generations, and protect them from the dangers of climate change.[1] The district court found that the plaintiffs had standing to sue, writing that "plaintiffs have introduced sufficient evidence and experts' opinions to demonstrate a question of material fact,"[2] including declarations from eighteen expert witnesses on the issue of climate change.[3] The Ninth Circuit reversed and ruled that while the plaintiffs had met the standing requirements of injury in fact and causation, they failed with respect to whether Article III courts could provide redress for the injury alleged.[4]

The ruling is noteworthy, not because it dismissed the case as beyond the federal judiciary's competence (a common handwashing for disposing of such lawsuits),[5] but for its acceptance of the causal link between climate change and the injuries to the plaintiffs. However, such acceptance for standing purposes is rarely uniform—many federal courts have found that plaintiffs have failed to show causation.[6] Moreover, as climate litigation continues, cases are proceeding to trial, and courts will evaluate such evidence under increasing levels of scrutiny as litigation progresses. Assessing climate evidence will thus transition from only a pleading to an evidentiary admissibility question: Can climate science be used in a courtroom to prove that specific carbon emissions harmed a specific plaintiff?

Environmental protection groups, municipalities, and states have initiated litigation against the worst polluters,[7] seeking remuneration for alleged environmental degradation and destruction caused by the defendants' emissions.[8] Others, like the plaintiffs in *Juliana*, have

---

1 Complaint for Declaratory and Injunctive Relief, Juliana v. United States, No. 6:15-cv-01517-TC (D. Or. Aug. 12, 2015).

2 Juliana v. United States, 339 F. Supp. 3d 1062, 1101 (D. Or. 2018).

3 *Id.* at 1086.

4 Juliana v. United States, 947 F.3d 1159, 1169–74 (9th Cir. 2020).

5 In a similar vein to redressability, many cases have also been thrown out on political question doctrine grounds. *See* California v. Gen. Motors Corp., No. C06-05755 MJJ, 2007 WL 2726871 (N.D. Cal. Sept. 17, 2007) (dismissing a state suit against six car companies for alleged emissions contributing to climate change as a nonjusticiable political question).

6 *See, e.g.*, Wash. Env't Council v. Bellon, 732 F.3d 1131, 1143–44 (9th Cir. 2013).

7 *See generally* Geetanjali Ganguly, Joana Setzer & Veerle Heyvaert, *If at First You Don't Succeed: Suing Corporations for Climate Change*, 38 Oxford J. Legal Stud. 841 (2018) (outlining the different generations of climate litigation, especially emphasizing changing legal strategies in light of new scientific evidence).

8 Note that this type of litigation aims not only at private corporations, but also suing governments for inaction. *See* Joana Setzer & Catherine Higham, Global Trends in Climate Change Litigation, 2021 Snapshot, Grantham Rsch. Inst. 23 (July 2021), https://www.lse.ac.uk/granthaminstitute/publication/global-trends-in-climate-litigation-2021-snapshot [https://perma.cc/L7WJ-EBLM] ("More than half of this group of cases (37) build on the approach taken in the landmark case of *Urgenda* . . . the first . . . litigation to

launched suits against governments for inaction in the face of clear climate warnings. Such litigation could play a crucial role in helping to shift governmental priorities and increase public awareness of the dangers of climate change, spurring action on a wider scale. Internationally, climate suits have resulted in concrete wins, including forcing the largest polluters to cut their carbon emissions.[9]

Like the *Juliana* lawsuit, these fights often play out on legal landscapes highly inhospitable to plaintiffs.[10] American courts have thrown out these suits on standing[11] or political question issues.[12] At this point few cases have even made it to trial. However, as more cases do proceed as society and courts realize the damage such emissions cause, courts will have to contend with issues of causation, a necessary component in any suit.

To admit this evidence, courts must accept climate science, a discipline which seeks to understand the changing global environment through complex modeling of the earth's physical and atmospheric conditions. This science can furnish that critical link between local emissions and global harms. Courts must decide whether to admit this crucial type of evidence under Rule 702 of the Federal Rules of Evidence (FRE) and the famous *Daubert* test. State courts where *Daubert* is the relevant rule of expert admissibility must also answer this question.[13]

A recent case highlights the importance of admitting and carefully considering climate evidence in climate suits. In August 2023, youth plaintiffs of Montana won a suit against the state, forcing it to take into

---

successfully challenge the adequacy of a national government's overall approach to reducing emissions. The whole-of-system approach . . . [can be] described as 'systemic mitigation' cases.").

[9] *See, e.g.*, HR 20 december 2019, NJ 2020, 425 m.nt. Van Mierlo (State of the Netherlands/Stichting Urgenda) (Neth.) (ruling for plaintiff's emissions targets in a private tort suit against Dutch government inaction on climate change); *see also* Rb. Den Haag 26 mei 2021, JOR 2021, 208 m.nt. SJMB (Milieudefensie/Royal Dutch Shell PLC) (Neth.) (ruling for environmental groups against Dutch oil company Shell Group, mandating that it must cut its CO_2 emissions by forty-five percent compared to 2019 levels).

[10] *See* Maria L. Banda, Env't L. Inst., Climate Science in the Courts: A Review of U.S. and International Judicial Pronouncements vi (2020) ("Courts, especially in the United States, have generally exercised restraint and deferred to the representative branches of government to devise solutions to the climate challenge.").

[11] *See, e.g.*, Native Village of Kivalina v. ExxonMobil Corp., 663 F. Supp. 2d 863, 880–81 (N.D. Cal. 2009), *aff'd*, 696 F.3d 849 (9th Cir. 2012).

[12] *See, e.g.*, California v. Gen. Motors Corp., No. C06-05755 MJJ, 2007 WL 2726871 (N.D. Cal. Sept. 17, 2007).

[13] *See* J.L. Hill, *The States of* Daubert *After Florida*, LexVisio (May 6, 2020), https://www.lexvisio.com/article/2019/07/09/the-states-of-daubert-after-florida [https://perma.cc/7GQ5-S63J] (characterizing Montana as a "Modified *Daubert*" jurisdiction).

account greenhouse gas emissions when permitting fossil fuel projects.[14] While the state did not file a motion attempting to bar the evidence of climate change or climate modeling (and even failed to present the testimony of their main expert witness during the trial),[15] the court not only credited the general postualtes of climate science and human-induced climate change,[16] but also climate models and projections that forecast the damage anthropogenic climate change will cause.[17] By accepting the causal link between Montana's actions or inactions,[18] Judge Seeley's ruling is the "strongest decision on climate change ever issued by any court."[19]

While the *Held* case stands as a strong affirmation of the legal significance and judicial acceptance of climate science, the ruling is expected to be appealed, and proving causation in tort suits still proves no straightforward task.[20] Not every judge will be as well versed or willing to engage with the climate evidence, and many defendants will put on more complete and vigorous defenses, attacking particularly the science of climate change. Therefore, this Note's call for judicial competence in climate science rests not only on a normative proposition of facilitating justice for those most affected by climate change, but also on the practical projection that the volume of climate cases will substantially increase.

It is surprising that such an important question has received relatively little scholarly attention. The most eloquent case for the strengths and weaknesses for admissibility comes from Engel & Overpeck.[21] They offer a balanced explanation of the science, compatibility with the law, and recommendations to aid judges in considering climate

---

14 David Gelles & Mike Baker, *Judge Rules in Favor of Montana Youths in a Landmark Climate Case*, N.Y. Times (Aug. 14, 2023), https://www.nytimes.com/2023/08/14/us/montana-youth-climate-ruling.html [https://perma.cc/8NKM-ADNV].

15 *Id.*

16 Findings of Fact, Conclusions of Law, and Order ¶¶ 65–92, Held v. State, No. CDV-2020-307 (Mont. Dist. Ct. Aug. 14, 2023), https://westernlaw.org/wp-content/uploads/2023/08/2023.08.14-Held-v.-Montana-victory-order.pdf [https://perma.cc/8P6L-MDJ8].

17 *Id.* ¶¶ 93–99.

18 *Id.* ¶¶ 99–139.

19 Micah Drew, Flathead Beacon & Amanda Eggert, *'This Changes Everything': Experts Respond to Held v. Montana Climate Ruling*, Mont. Free Press (Aug. 17, 2023), https://montanafreepress.org/2023/08/17/this-changes-everything-experts-respond-to-landmark-youth-climate-ruling [https://perma.cc/2P5T-24CE].

20 *See generally* Kimberly M.S. Cartier, *Climate Litigation Has a Big Evidence Gap*, Eos (July 23, 2021), https://eos.org/articles/climate-litigation-has-a-big-evidence-gap [https://perma.cc/KFU4-TV9M] (describing the lack of communication between climate scientists and climate lawyers in proving specific causation in climate suits).

21 Kirsten Engel & Jonathan Overpeck, *Adaptation and the Courtroom: Judging Climate Science*, 3 Mich. J. Env't & Admin. L. 1 (2013).

science under *Daubert*.[22] In contrast, academics and practitioners skeptical of admissibility take issue with different aspects of climate science. Best presented by Morrison et al.,[23] skeptics argue that the lack of statistical significance,[24] the number of assumptions made by climate researchers,[25] the inaccuracy of predictions,[26] and the inability to connect wider climate change to local conditions all should preclude admission by courts.[27] Critics particularly worry about this last problem, the cutting-edge practice of "downscaling," that seeks to link climate change to effects in smaller localities.[28] Others particularly take offense at the subjectivity of climate modeling methods.[29] Engel & Overpeck actually agree with several of these critiques, but do not hold that they are grounds for preclusion.[30]

Morrison et al. center the Intergovernmental Panel on Climate Change Assessment Reports (IPCC ARs), U.N.-sponsored reports on global warming, as the main battlefield for climate science admissibility.[31] The IPCC reports are important as they represent the "gold standard" of climate science findings, because of the "meticulous procedures followed by a broad range of climate scientists in order to generate" their findings.[32] Thus, while the admissibility of the report itself may not be implicated in every trial, to the extent certain claims are based on the report's findings or methods utilized, it may also come under scrutiny.[33]

---

22 Their three principles include (1) a consensus approach to weighing climate science, (2) relying on past confirmatory results to assess particular models, and (3) that the climate scientist be trained specifically in climate science as opposed to other fields of science. *Id.* at 27–31.

23 Fred K. Morrison, Craig Manson & Matthew C. Wickersham, *Climate Change Science and the* Daubert *Standard*, 44 Wm. & Mary Env't. L. & Pol'y Rev. 391 (2020).

24 *Id.* at 407.

25 *Id.* at 408.

26 *Id.* at 410.

27 *Id.* at 412.

28 Engel & Overpeck, *supra* note 21, at 17.

29 Alvaro Hasani, *Forecasting the End of Climate Change Litigation: Why Expert Testimony Based on Climate Models Should Not Be Admissible*, 32 Miss. Coll. L. Rev. 83, 100 (2013) ("[T]he subjectivity involved in the development of models, that is decisions as to which element of a system to integrate into a model or which type of model is best suited for the purpose, further disguise the standards used.").

30 *Id.* at 23 (writing that climate impact projects cannot be tested through randomized control tests that produce a known "rate of error"); *see also id.* at 25 (characterizing climate science as dealing with uncertainties but also a "continuing rapid pace of new discoveries"); *see also id.* at 26 (agreeing that downscaling is more uncertain than the global effects of climate change).

31 *See* Morrison et al., *supra* note 23, at 405 ("With respect to climate science, plaintiffs are likely to rely on the Intergovernmental Panel on Climate Change ('IPCC').").

32 Engel & Overpeck, *supra* note 21, at 18.

33 *See* Morrison et al., *supra* note 23, at 405 & n.88 (writing that "this Article does not address" the arguments for or against the admissibility of the IPCC reports, but instead

This Note similarly treats the IPCC findings and methods as key in describing important aspects of climate science; and, further than Morrison et al., explicitly looks to their admissibility as evidence.

This Note takes on the question of whether climate testimony can survive a *Daubert* motion under Federal Rule of Evidence 702. It finds that such evidence can. While climate science exhibits legal characteristics of both formal science and expert evidence, it is best thought of as hyperqualified technical expertise. The current literature overlooks this way of approaching the debate. Instead of only measuring climate testimony up to the scientific yardstick of *Daubert*, judges should compare climate scientists to other experts who put forward "technical, or other specialized knowledge."[34] A direct comparison to forensic methods bolsters the case for similar treatment. Although climate models do need to prove causation, admissibility addresses whether testimony can "demonstrate . . . that [the experts'] opinions are reliable" where "[t]he evidentiary requirement of reliability is lower than the merits standard of correctness."[35] In contrast, whether they can prove causation concerns the weight a factfinder gives to the evidence.

The Note proceeds as follows. Part I explains the inception of courts' gatekeeping role for expert evidence created by *Daubert* and its progeny. It describes how courts have categorized different kinds of evidence under either *Daubert* or the more permissive *Kumho Tire* standard. Courts have interpreted *Kumho Tire* to hold such evidence to largely in-group comparisons, to other like experts, instead of subjecting the evidence to a purely "scientific" analysis. Part II reveals how this construction operates in the daily admission of forensic evidence in courts across the country. The discussion reveals that courts indeed appear to approach forensic expert testimony under a less strictly scientific approach, especially surrounding the falsifiability factor of the analysis, but still finds such testimony admissible. Part III then explains climate science and analyzes it under the *Daubert* standard. It finds that climate science is not legally "scientific" under *Daubert*. Instead, Part IV argues that climate testimony and the IPCC AR6 should be analyzed under the *Kumho Tire* expert testimony standard. This argument also benefits from the advances made by this most recent report, released in 2021.[36] It compares forensics and climate science directly, showing that climate

"assumes that the [reports] will be recognized and plaintiffs will attempt to use the reports to draw further individual conclusions").

34 Kumho Tire Co. v. Carmichael, 526 U.S. 137, 149 (1999).

35 Fed. R. Evid. 702 advisory committee's note on 2000 amendment to Rule 702 (quoting *In re* Paoli R.R. Yard PCB Litig., 35 F.3d 717, 744 (3d Cir. 1994)).

36 *See Sixth Assessment Report*, IPCC, https://www.ipcc.ch/assessment-report/ar6 [https://perma.cc/4J5H-H9YL].

evidence's claim to admissibility is comparatively much stronger than that of forensics, including the methodologies relied upon by forensics experts. These points together make a strong case that courts should admit this science as sufficiently reliable.

I
The Evolution from *Daubert* to *Kumho Tire*: Scientific Testimony Versus Expert Evidence

This Part gives a brief background of the *Daubert* standard, its incorporation into the Federal Rules of Evidence, and its extension by *Kumho Tire* to the regulation of all expert testimony. Section I.A reviews the sea-change of *Daubert*, while Section I.B introduces how courts have appeared to evaluate experts whose approaches are not purely scientific. This Part finds a difference, especially in more subjective expert methods, between *Daubert* and *Kumho Tire*.

A. A Brief History of Scientific Evidence in the Courtroom

Modern standards governing the admissibility of scientific evidence did not emerge until the early twentieth century with the *Frye* decision, which became the main test for courts until *Daubert*.[37] *Frye* is significant because it was the first step in answering questions especially germane to this Note: How do American courts divide the work of assessing scientific evidence between the judge and the jury, how much discretion should judges have to make that decision, and how should judges (and juries) understand innovations or newer scientific evidence?[38] In essence, how should courts rule in the face of uncertainty?

In *Frye v. United States*, the D.C. Circuit barred the results of an early polygraph device proffered by a criminal defendant to prove his innocence by adopting a "general acceptance" test—essentially punting the question of admissibility to the scientific community.[39] There is an enticing simplicity (perhaps) and practical power in this approach:

---

[37] David L. Faigman, *The* Daubert *Revolution and the Birth of Modernity: Managing Scientific Evidence in the Age of Science*, 46 U.C. Davis L. Rev. 893, 898 (2013) (describing how *Daubert* overturned the prior test governing the admissability of scientific expert testimony, *Frye*).

[38] *See generally* Edward J. Imwinkelried, *Trial Judges — Gatekeepers or Usurpers? Can the Trial Judge Critically Assess the Admissibility of Expert Testimony Without Invading the Jury's Province to Evaluate the Credibility and Weight of the Testimony?*, 84 Marq. L. Rev. 1, 2, 40 (2000) (describing how determinations of the admissability of scientific evidence require the judge to assume the role of the fact finder and the procedural consequences this assumption creates).

[39] 293 F. 1013, 1014 (D.C. Cir. 1923). The court wrote that approval by the "particular field in which it belongs" should allow scientific consensus to decide on admissibility. *Id.*

Judges are experts of law, not science.[40] Its simplicity also reveals its greatest flaw. It is a less flexible test, as the decision itself reads: "Just when a scientific principle or discovery crosses the line between the experimental and demonstrable stages is difficult to define. Somewhere in this twilight zone the evidential force of the principle must be recognized . . . ."[41] Defining that twilight zone is harder to do with newer technologies and scientific approaches.

This "Ancien Régime" gave way to the "revolution"[42] of *Daubert* and its related developments, "the *Daubert* trilogy": the *Daubert* decision, *General Electric Company v. Joiner*,[43] and *Kumho Tire Company v. Carmichael*.[44] Of the trilogy, *Daubert* is the foundation; the other decisions expand upon its innovations and refine its themes. In *Daubert*, plaintiffs brought a toxic tort action against a manufacturer of Bendectin, a drug they alleged caused birth defects.[45] The issue was whether to allow the plaintiffs' expert testimony that concluded the drug caused the defects.[46] The Court made one ruling with three parts: expert testimony must be relevant to the issue in dispute; the expert must be sufficiently qualified; and the testimony "must be supported by appropriate validation."[47] Judges now had the responsibility to assess the validity and reliability of the methods employed by the expert in coming to their conclusion. Judges rely on the factors well known by any law student who has taken evidence: falsifiability,[48] peer review,[49] error rate,[50] methods controlling its operation,[51] and the vestige of *Frye* in a general acceptance factor.[52] However, unlike the strict test in *Frye*, the trial judge can balance these factors flexibly under FRE Rule 104(a), finding that expert evidence is reliable by a preponderance of the evidence.[53]

---

[40] Defining the relevant community can pose problems. *See* Faigman, *supra* note 37 at 899 ("Although the *Frye* test appears straightforward, the simplicity of stating it belies the complexity inherent in applying it to concrete cases.").

[41] *Frye*, 293 F. at 1014.

[42] Faigman's analogy underscores the "sea change" *Daubert* represented to the admissibility of scientific evidence. Faigman, *supra* note 37, at 907.

[43] 522 U.S. 136 (1997).

[44] 526 U.S. 137 (1999).

[45] Daubert v. Merrell Dow. Pharms., Inc., 509 U.S. 579, 582 (1993).

[46] *Id.* at 583–84.

[47] *Id.* at 590.

[48] *Id.* at 593 (quoting Karl R. Popper, Conjectures and Refutations: The Growth of Scientific Knowledge 37 (5th ed. 1989)).

[49] *Id.* at 593.

[50] *Id*. at 594.

[51] *Id.*

[52] *Id.*

[53] *Id.* at 592–93.

These guidelines do not serve as a "holy writ" in the words of Justice Scalia, but instead as guardrails for the trial court's analysis.[54]

Daubert gave judges a new purview and command: "*Daubert* reset courts' gaze onto the inner workings of knowledge gathering" as before courts had "at most inspected merely the outer forms of expert evidence."[55] This solidified the judge's role as a gatekeeper of admissibility, "set[ting] the boundary between the judge's responsibility to determine admissibility and the fact-finder's responsibility to assess weight."[56] The other two sequels further fleshed out the doctrine. *Joiner* protected the trial court's decision, instituting an abuse of discretion review standard.[57] *Kumho Tire* erased the distinction between scientific and nonscientific testimony—every technical expert whose findings were not primarily based on science would also be assessed under the basic guidelines *Daubert* established.[58] However, not all the factors may apply in the same way to real estate appraisal, as to neuroscience, or to a heart surgeon testifying.

Academics and practitioners have debated whether this revolution aided the admissibility of evidence in federal courts. The general consensus is that *Daubert* restricted the admissibility of plaintiff evidence in civil cases,[59] while it has done little to stop the flow of forensic evidence proffered by the government in criminal cases.[60] Commentators disagree on whether this development charts the course set out by the justices in *Daubert*, although any singular strand of thought might be hard to untangle.[61] It is crucial, however, to understand how *Daubert* interacts with *Kumho Tire*, and how judges have interpreted each. Because climate science has elements of subjectivity, as well as more formal scientific rigor, a judge must understand both.

---

54 Kumho Tire Co. v. Carmichael, 526 U.S. 137, 159 (1999) (Scalia, J., concurring).

55 Faigman, *supra* note 37, at 907.

56 *Id.* at 909.

57 Gen. Elec. Co. v. Joiner, 522 U.S. 136, 139 (1997).

58 *Kumho Tire*, 526 U.S. at 148.

59 *See, e.g.*, Andrew Jurs & Scott DeVito, *The Stricter Standard: An Empirical Assessment of* Daubert*'s Effect on Civil Defendants*, 62 Cath. U. L. Rev. 675 (2013) (finding that, by measuring a decrease in removals from state to federal courts, "civil defendants believe the *Daubert* standard is more restrictive to expert testimony and act accordingly").

60 *See* Elizabeth L. DeCoux, *The Admission of Unreliable Expert Testimony Offered by the Prosecution: What's Wrong with* Daubert *and How to Make It Right*, 2007 Utah L. Rev. 131, 132–33 (2007); *see also* D. Michael Risinger, *Navigating Expert Reliability: Are Criminal Standards of Certainty Being Left on the Dock?*, 64 Ala. L. Rev. 99, 109–12 (2000) (compiling district court cases where *Daubert* challenges against civil plaintiffs often succeeded, while criminal defendant challenges to prosecution methods often failed).

61 *See* Faigman, *supra* note 37, at 911 (blaming confusion about whether *Frye* or *Daubert* is more restrictive on the justices themselves, characterizing *Daubert* as "a cornucopia of confused messages").

B. Daubert *vs.* Kumho Tire*: Determining the Line Between "Scientific" and "Technical" Knowledge*

Kumho Tire seemed to push back on the notion that expert methodologies needed to meet the exact test enunciated in *Daubert*, but the decision itself is unclear. While the Court re-emphasized that the *Daubert* factors should be utilized, the Court did not "rule out, nor rule in, for all cases . . . the applicability of the factors mentioned in *Daubert* . . . ."[62] The majority did state affirmatively that "it would prove difficult, if not impossible, for judges to administer evidentiary rules under which a gatekeeping obligation depended upon a distinction between 'scientific' knowledge and 'technical' or 'other specialized' knowledge. There is no clear line that divides the one from the others."[63] Further, the full list of factors may not "appl[y] to all experts in every case."[64] This begs the question of whether courts should hold non-scientific expert evidence to a standard of scientific proof, or something else.

Judge Pollack in the noteworthy *Llera Plaza* case seemed to initially conflate the "scientific" *Daubert* factors with the *Kumho Tire* test.[65] The court in that case initially applied a "scientific" label to the *Daubert* factor test. In his first formulation of the test, the judge categorized fingerprinting as a technical rather than scientific discipline under *Kumho Tire*.[66] However, the court judged the methodology using the "scientific" standards of *Daubert*, comparing the method to other scientific disciplines.[67] He found that the testimony did not live up these standards and could only be admitted without the experts' ability to render an opinion as to a particular fingerprint match.[68] However, in reversing himself shortly thereafter, the judge walked back an application of a "scientific" *Daubert* standard to each factor.[69] Indeed, upon reconsideration, he found that fingerprinting was sufficiently rigorous under a peer review standard within the specialist community, instead of adjudicating it through what the *Daubert* Court would consider a "scientific community."[70] He made

---

62 *Kumho Tire*, 526 U.S. at 150–52.

63 *Id.* at 148.

64 *Id.* at 141.

65 United States v. Llera Plaza (*Llera Plaza I*), 179 F. Supp. 2d 492, 508 (E.D. Pa. 2001), *vacated and superseded on reconsideration*, 188 F. Supp. 2d 549 (E.D. Pa. 2002).

66 *Id.* at 516.

67 *Id.* ("[T]he court finds that ACE-V does not adequately satisfy the 'scientific' criterion of testing . . . or the 'scientific' criterion of peer review . . . or . . . ACE-V's 'scientific' rate of error . . . and that, at the critical evaluation stage, ACE-V does not operate under uniformly accepted 'scientific' standards . . . .").

68 *Id.*

69 United States v. Llera Plaza (*Llera Plaza II*), 188 F. Supp. 2d 549, 563 (E.D. Pa. 2002).

70 *Id.* at 564–65.

the same move with testing as well as rate of error.[71] This view is consistent with how other courts treat technician expert evidence[72]—it is not judged according to what would be acceptable in a more rigorous "scientific" context, but instead leaves the "scientific" adjudication for those disciplines falling within *Daubert*.

The move from judging non-scientific experts on a purely scientific standard to a more flexible technician standard is crucial to understand. By judging expert evidence by reference to others in that field, such experts appear to be subjected to a different level of scrutiny.[73] Such an insight is not meant necessarily to cast doubt on the reliability of such methods, nor on Judge Pollack's eventual approach, which represents the almost-unanimous position of federal courts.[74] It is merely descriptive and offered to suggest that such a move was not inevitable. There may exist good rationales for permitting such self-referential evidence, as often rigorous scientific evidence is either unavailable or would be impossible to create. It also allows for the admission of potentially probative evidence, following the Court's acknowledgment that "[t]oo much depends upon the particular circumstances of the particular case at issue" to have too strict a rule for all of the factors listed in *Daubert*.[75] These concerns do not mean technical evidence is faulty or should not be admitted—it should be a question of weight for the finder of fact after they are presented with the adversarial interrogation of such evidence. The next Part of this Note details how courts follow this approach with two types of important forensic evidence, representative of courts' general interpretation of *Kumho Tire*: fingerprinting and ballistics.

II
Courts Find Forensic Methods Admissible Under *Kumho Tire*

This Part discusses and compares the admissibility of two main forms of forensic evidence that are routinely allowed in criminal prosecution cases under *Kumho Tire*: ballistics and fingerprinting techniques. By analyzing both under every *Daubert* factor, the Part reveals how courts apply an approach that looks more to the reliability of the expert

---

71 *Id.*

72 *See generally* discussion *infra* Part II.

73 Courts employing this approach do seem to pledge loyalty to a line in *Kumho Tire*, that: "It is to make certain that an expert, whether basing testimony on professional studies or personal experience, employs in the courtroom the same level of intellectual rigor that characterizes the practice of an expert in the relevant field." Kumho Tire Co. v. Carmichael, 526 U.S. 137, 151 (1999).

74 *See generally* discussion *infra* Part II.

75 *Kumho Tire*, 526 U.S. at 150.

evidence rather than a strictly scientific qua the "scientific method" approach as laid out in *Daubert*. Especially divergent from the professed strictures of *Daubert* are those factors most impacted by the subjective nature of the forensic approaches: falsifiability, peer review, and controlling methods. By showing how courts understand such evidence as sufficiently reliable, this Part lays the groundwork for comparison to the climate science presented later in Part IV.

A. Falsifiability

The first *Daubert* factor judges must consider to determine whether a sufficient basis exists for a scientific proposition is its "falsifiability, or refutability, or testability"; essentially, whether the inquiry involves "generating hypotheses and testing them to see if they can be falsified."[76] Implicit in falsifiability is the idea of reproducibility—that one must be able to test a hypothesis *again* to falsify it.[77] Many forensic methods depend on the subjective value judgment of a single expert, compromising any true reproducibility and hence falsifiability.[78] Thus, such evidence sits uneasily with this *Daubert* factor, although courts regularly admit forensic evidence despite a lack of strict falsifiability.[79] And further, the tides of change can swell quickly to admit new evidence that was previously considered unreliable. In 1923, the Illinois Supreme Court rejected a ballistics report as "not only impossible, but

---

[76] Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579, 593 (1993). The Court partially relied on the work of philosopher of science Karl Popper for this factor in their test, crediting his idea that the line between science and non-science was the ability to test and prove false a scientific proposition. *See id.* (citing Karl Popper, Conjectures and Refutations: The Growth of Scientific Knowledge 37 (5th ed. 1989)). Commentators have criticized the Court's use of Popper as the basis for their test of scientific validity, especially surrounding falsifiability. *See* Sean O'Connor, *The Supreme Court's Philosophy of Science: Will the Real Karl Popper Please Stand Up?*, 35 Jurimetrics 263, 276 (1995) ("Popper's system is an odd choice for the model of good science."); Ronald J. Allen, *Expertise and the* Daubert *Decision*, 84 J. Crim. L. & Criminology 1157, 1168–69 (1994) ("The court simply replaced a judicial anachronism with a philosophical one."). Further, Popper rejected verificationism, the inductive clustering of results that tend to prove a larger theory, positing instead that "there can be no reason to believe that a theory that passed a certain test today would pass the same test tomorrow." Susan Haack, *Federal Philosophy of Science: A Deconstruction—And a Reconstruction*, 5 N.Y.U. J.L. & Lib. 394, 404 (2010).

[77] Karl Popper, The Logic of Scientific Discovery 66 (1992) ("[N]on-reproducible single occurrences are of no significance to science.").

[78] *See infra* note 82.

[79] It is unclear how much of this is due to true judicial determinations of falsifiability, or general confusion. *See* Sophia I. Gatowski, Shirley A. Dobbin, James T. Richardson, Gerald P. Ginsburg, Mara L. Merlino & Veronica Dahir, *Asking the Gatekeepers: A National Survey of Judges on Judging Expert Evidence in a Post-*Daubert *World*, 25 Law & Hum. Behav. 433, 433 (2001) (finding in telephonic interviews with state trial court judges only five percent of 400 gave correct answers about the definition of falsifiability—although these were state court judges, not all of whom have ever dealt with *Daubert*).

'preposterous.'"[80] Merely seven years later that same court admitted firearms identification evidence as part of a movement of "widespread judicial acceptance."[81] Firearms identification seeks to confirm that certain ammunition was fired from a particular firearm, based on toolmarks bored into the gun's barrel.[82] The Association of Firearm and Tool Mark Examiners (AFTE), an industry group that supports the use and diffusion of ballistics expertise, has called this traditional pattern recognition method "subjective in nature, founded on scientific principles and based on the examiner's training and experience"[83] and said that "it remains for the examiner to determine for [themselves] the modicum of proof necessary to arrive at a definitive opinion."[84] Ballistics experts often get a second opinion, although this process can subject the second examiners to confirmation bias.[85] So for falsifiability, ballistics is plagued with subjectivity concerns, as only a second subjective (and potentially biased) participant can verify the process.

Fingerprinting is another discipline that relies on subjective expert understanding. The method of fingerprinting most commonly employed today in the United States originates from the United Kingdom.[86] Researchers there developed a method of matching ridges on human fingertips with prints found on another surface, to link a person to a location or object.[87] The first homicide conviction in the United States to rely on fingerprint evidence was in 1911 (again in Illinois).[88] Even putting aside problems with the main assumption of the uniqueness

---

[80] Paul Giannelli & Edward Imwinkelried, Scientific Evidence 548 (4th ed. 2007) (citing People v. Berkman, 139 N.E. 91, 94 (Ill. 1923)).

[81] *Id.*; People v. Fisher, 172 N.E. 743, 753 (Ill. 1930).

[82] This Section refers regularly to the 2016 report of the President's Council of Advisors on Science & Technology: *Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods.* This landmark report on forensic methods evaluated the reliability of these tools, and recommended ways in which such evidence could be strengthened and appropriately used in courtrooms. President's Council of Advisors on Sci. & Tech., Exec. Office of the President, Forensic Science in Criminal Courts: Ensuring Scientific Validity of Feature-Comparison Methods 104 (2016), https://obamawhitehouse.archives.gov/sites/default/files/microsites/ostp/PCAST/pcast_forensic_science_report_final.pdf [https://perma.cc/3LJQ-UXPT] [hereinafter PCAST 2016].

[83] *Theory of Identification as It Relates to Toolmarks*, 30 AFTE J. 86, 86 (1998).

[84] Joseph L. Peterson, Ellen L. Fabricant & Kenneth S. Field, Crime Laboratory Proficiency Testing Research Program 207 (1978).

[85] For a recent example, see Keith L. Alexander, *Ballistics Work at D.C.'s Crime Lab Criticized by Forensic Experts*, Wash. Post (Mar. 26, 2021), https://www.washingtonpost.com/local/public-safety/dc-crime-lab-ballistics-mistake/2021/03/26/42e992aa-8c0e-11eb-a730-1b4ed9656258_story.html [https://perma.cc/Y222-AZV2] ("The auditors also said management pressured an analyst into changing their second finding.").

[86] Jessica M. Sombat, *Latent Justice:* Daubert*'s Impact on the Evaluation of Fingerprint Identification Testimony*, 70 Fordham L. Rev. 2819, 2828 (2002).

[87] *Id.*

[88] People v. Jennings, 96 N.E. 1077 (Ill. 1911).

and permanence of fingerprints,[89] the fingerprinting method itself (the Analysis, Comparison, Evaluation, and Verification (ACE-V) method) relies on subjective value judgments. Examiners look at three levels: general ridge patterns, ridge characteristics, and microscopic ridge attributes.[90] There is no agreement about the "minimum number of points of similarity [between fingerprints] required before a conclusion of identity may be reached."[91] Compare this to ballistics, and the widely-followed AFTE method: "Toolmark analysis does not follow an objective standard requiring, say, a certain percentage of marks to match."[92] Just like ballistics, other reviewers often look to "test" the first opinion, but often confirmation bias interferes and leads to the confirmation of incorrect results.[93] Thus, fingerprinting suffers from similar problems about testability, as a second look is *also* subjective, and without proper safeguards (blind examination) it can lead to motivated (read: nonscientific) results.

Falsifiability as a touchstone has not kept other types of forensic specialists off the stand. Psychologists routinely testify in trials, assessing defendants or theoretical individuals in the exact circumstances of the defendant. Investigating scientists in that field undertook a series of verification studies attempting to reproduce many foundational studies.[94] In many cases, they could not.[95] Few would say that psychology is not a rigorous discipline, although clearly its basic testability suffers from serious problems. Fewer still have advocated for keeping psychologists off the stand. Therefore, it appears that courts do not find a lack of scientific testability to be fatal to expert testimony admissibility.

B. Peer Review

Courts often do not hold forensic methods to the same standards of peer review as other areas of science. The *Daubert* Court wrote that a

---

[89] *See* Sandy L. Zabell, *Fingerprint Evidence*, 13 J.L. & Pol'y 143, 164 (2005) ("Although there is a substantial literature on the uniqueness of fingerprints, it is surprising how little true scientific support for the proposition exists.").

[90] Paul C. Giannelli, Daubert *Challenges to Fingerprints*, 42 Crim. L. Bull. 624, 627 (2006).

[91] *Id.* at 627–28.

[92] United States v. Monteiro, 407 F. Supp. 2d 351, 370 (D. Mass. 2006).

[93] For a high-profile example, see the FBI's misidentification of a fingerprint left at the 2004 Madrid train bombings. Sarah Kershaw, Eric Lichtblau, Dale Fuchs & Lowell Bergman, *Spain and U.S. at Odds on Mistaken Terror Arrest*, N.Y. Times (June 5, 2004), https://www.nytimes.com/2004/06/05/us/spain-and-us-at-odds-on-mistaken-terror-arrest.html [https://perma.cc/9ZZQ-AQML].

[94] Open Science Collaboration, *Estimating the Reproducibility of Psychological Science*, 349 Sci. 943 (2015).

[95] *Id.*

court should consider whether a method is published in an appropriate journal to find peer review: "The fact of publication (or lack thereof) in a peer reviewed journal thus will be a relevant, though not dispositive, consideration."[96] Firearm analysis methods regularly call the review of a second examiner "peer review." In the scientific peer review method envisioned by *Daubert*, "publication in a peer-reviewed journal is a relevant, though not dispositive, consideration in assessing the validity of a particular technique."[97] The AFTE's journal and its associated technique were sufficient for Judge Saris in *United States v. Monteiro*, although she noted "other peer reviewed articles have not been universally laudatory of the current technique of identification."[98] Indeed, the President's Council of Advisors on Science & Technology 2016 report warned that judges should exercise caution when using industry journals as their own form of peer review. Such "journals" would not be acceptable in other published areas of science, as "papers on the foundational validity of forensic feature-comparison methods [should be] published in leading scientific journals rather than in forensic-science journals, where, owing to weaknesses in the research culture of the forensic science community . . . the standards for peer review are less rigorous."[99] To call this "peer-reviewed" to the scientific standard set out in *Daubert* is perhaps a stretch.[100]

Judges have understood fingerprint peer review in a similar way. One court found that "peer review is the standard operating procedure among latent print examiners,"[101] another indication that *Daubert* "peer review" is more akin to an industry-referential reliability analysis. As with ballistics, judges do not appear to consider the publications of technicians "scientific" in the same way as other peer-reviewed journals. As highlighted in the initial discussion of technician expertise, Judge Pollack concluded that although fingerprint specialists were not scientists and their forensic journals were not "scientific journals in *Daubert*'s peer review sense,"[102] these factors did not "militate against the utility of the [fingerprint] identification procedures."[103]

---

96 Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579, 593–94 (1993).

97 United States v. Monteiro, 407 F. Supp. 2d 351, 366 (D. Mass. 2006) (quoting *Daubert*, 509 U.S. at 594).

98 *Id.* at 366–67.

99 PCAST 2016, *supra* note 82, at 125.

100 See the amici brief from the New England Journal of Medicine in *Daubert* itself. Brief of the New England Journal of Medicine, Journal of the American Medical Association, and Annals of Internal Medicine as Amici Curiae in Support of Respondent at 2, Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579 (1993).

101 United States v. Havvard, 117 F. Supp. 2d 848, 854 (7th Cir. 2001).

102 *Llera Plaza II*, 188 F. Supp. 2d 549, 563 (E.D. Pa. 2002).

103 *Id.*

C. Error Rate

The *Daubert* ruling charged judges with looking at the "known or potential rate of error" when assessing expert evidence.[104] While the opinion does not lay out a specific threshold for the error rate, simply calling on judges to ascertain *whether* one exists, the usual legal thresholds for rates of error are demanding.[105] Neither ballistics nor fingerprinting have sufficiently scientific error rates, largely because of the lack of large-scale studies or meta-analyses. There are very few credible research studies on ballistic error rates. The most reliable, the Ames Laboratory Study, tasked examiners to make independent comparisons between samples.[106] In the most positive light, the examiners only found false positives (matches where there were none present) one in sixty-six times, or at the highest confidence interval bound, one in forty-six times.[107] On the other hand, 33.7% of the samples could not be identified—although the study decided that these should not be considered "errors" as there might just have been insufficient groove marks to make a call.[108] This study suggests relatively low error rates, but this is functionally the only proper study of ballistics accuracy. Concluding anything from a single study, instead of hundreds or thousands, is antithetical to the probabilistic nature of scientific inquiry.

Fingerprinting has an even less convincing error rate, based on completed studies. The PCAST 2016 Report summarized all the major fingerprinting analyses through 2016 and concluded that "the studies collectively demonstrate that many examiners can, under *some* circumstances, produce correct answers at *some* level of accuracy."[109] Encouragingly, the FBI published one black-box study in a peer-reviewed journal, following blind verification techniques.[110] However, other cases show far less reliability; for example, in the *Llera Plaza* case, the FBI

---

104 Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579, 594 (1993).

105 The usual acceptable link is 90–95% confidence intervals. *See* Daniel L. Rubinfeld, *Econometrics in the Courtroom*, 85 Colum. L. Rev. 1048, 1058 n.33 (1985) (summarizing the consensus approach of a two-standard-deviation approach); *see also* EEOC v. Mavis Discount Tire, Inc., 129 F. Supp. 3d 90, 110 (S.D.N.Y. 2015) (writing that while courts have rejected a "formal" litmus test for Title VII claims, two or three standard deviations "can be highly probative").

106 David P. Baldwin, Stanley J. Bajic, Max Morris & Daniel Zamzow, A Study of False-Positive and False-Negative Error Rates in Cartridge Case Comparisons (Ames Laboratory, U.S. Dep't of Energy Technical Report #IS-5207, 2014), https://www.ojp.gov/ncjrs/virtual-library/abstracts/study-false-positive-and-false-negative-error-rates-cartridge-case [https://perma.cc/PH5K-W783].

107 PCAST 2016, *supra* note 82, at 11 (summarizing the Ames Laboratory Study).

108 *Id.* at 110–11 (referring to the Ames Laboratory Study).

109 *Id.* at 95.

110 Bradford T. Ulery, R. Austin Hicklin, JoAnn Buscaglia & Maria A. Roberts, *Accuracy and Reliability of Forensic Latent Fingerprint Decisions*, 108 PNAS 7733 (2011).

sent the defendant's fingerprints out to several state labs for testing: 9 out of 39 did not correctly identify both latent prints.[111] Thus, courts do not consider the lack of known error rates of a technique to be a sine qua non for the reliability of this evidence.

D. Methods Controlling Operation

Courts must also look to "the existence and maintenance of standards controlling the technique's operation."[112] These forensic methods follow very loose controlling standards. While there is a basic ballistic method, the AFTE toolmark theory, it is completely up to the examiner to determine what constitutes "sufficient agreement" between the grooves on the barrel and on the ammunition.[113] It is "tautological" in that the identification "is not based on any qualitative standard for how many striations or marks need to match or line up" but "is based on a holistic assessment of what the examiner sees."[114] An examiner finds a match when they find a match. Further, "most examiners, many of whom are state and local law enforcement agents, go through no formal training program, certification or annual testing . . . ."[115] For those who are tested, there are further problems with what are deemed type I and type II tests: type I tests are indicators of competency, designed for internal purposes, while type II proficiency tests assess an examiner's conclusions in the real world.[116] Indeed, a high pass rate for an internal test may be indicative of an easy test that sandbags results.

Fingerprinting mirrors this lack of standards. As with tool marking, there is no agreed-upon number of points between fingerprints to consider them a match. The number of matching ridge points (deemed Galton points after an early proponent in England)[117] in the UK is sixteen, in Australia twelve; in the U.S. the FBI looks to a twelve point "quality assurance" standard, but individual states have their own standards.[118] Like ballistics, there are also questions about the regulation

---

111 *Llera Plaza II*, 188 F. Supp. 2d 549, 559 (E.D. Pa. 2002).

112 Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579, 594 (1993) (citation omitted) (giving the example of a "professional organization's standard governing spectrographic analysis").

113 United States v. Monteiro, 407 F. Supp. 2d 351, 363–64 (2006).

114 *Id.*

115 Joan Griffin & David J. LaMagna, Daubert *Challenges to Forensic Evidence: Ballistics Next on the Firing Line*, Champion, Sept.–Oct. 2002, at 60.

116 Daniel J. Capra, *Symposium on Forensic Expert Testimony,* Daubert*, and Rule 702*, 86 Fordham L. Rev. 1463, 1509 (2018) (quoting Professor Jay Koehler).

117 *See* Sombat, *supra* note 86 at 2827–28 ("[T]he work of Galton truly substantiated fingerprinting as a science.").

118 *Llera Plaza I*, 179 F. Supp. 2d 492, 513 (E.D. Pa. 2001), *vacated and superseded on reconsideration*, 188 F. Supp. 2d 549 (E.D. Pa. 2002).

of the examiners themselves, and indications that sandbagging may be prevalent. When a former fingerprint analyst at Scotland Yard assessed the FBI's internal proficiency testing, he found it so easy that "[i]f I gave my experts these tests . . . they'd fall about laughing."[119] If this is the FBI's proficiency test, likely the best-equipped and most rigorous program in the country, it is hard to imagine that state forensic labs are living up to a sufficient level of proficiency testing.

E. General Acceptance

The general acceptance prong, which looks to the support of a theory within the "relevant scientific community," has become largely self-referential.[120] While forensic science may not have acceptance in the broader scientific community, courts have not found this a significant impediment to admissibility. The forensic ballistics community agrees with their methodology—unsurprisingly—although there are some dissenters for more objective methods.[121] Courts have found such agreement amongst technicians persuasive.[122] For fingerprinting, in *Llera Plaza II*, Judge Pollack compared fingerprinting experts to "accountants, vocational experts, [and] accident-reconstruction experts" who only had "technical, or other specialized knowledge."[123] This represents a different conception from a "scientific community," but is consistent with what other courts have held when it comes to fingerprint analysis.[124]

F. Summary: Expert Forensic Methods Are Not Judged on Purely Scientific Grounds

This Part reveals that courts deviate from what might be expected under the Court's ruling in *Daubert* when evaluating the admissibility of forensic methods. It supports this Note's contention that under *Kumho Tire* courts assess this evidence under a spirit of acceptance of expert testimony even when such evidence is not strictly scientific. The

---

119 *Llera Plaza II*, 188 F. Supp. 2d 549, 565 (E.D. Pa. 2002).

120 Daubert v. Merrell Dow Pharms., Inc., 509 U.S. 579, 594 (1993).

121 Alfred A. Biasotti & John E. Murdock, *The Scientific Basis of Firearms and Toolmark Identification*, *in* 4 MODERN SCIENTIFIC EVIDENCE: THE LAW AND SCIENCE OF EXPERT TESTIMONY 990, 999 (David L. Faigman et al. eds., 2002).

122 *See, e.g.*, Commonwealth v. Meeks, Nos. 2002-10961, 2003-10575, 2006 WL 2819423 (Mass. Sup. Ct. Sept. 28, 2006).

123 188 F. Supp. 2d at 563–64 (citations and footnote omitted).

124 *See* United States v. Collins, 340 F.3d 672, 682 (8th Cir. 2003) ("Fingerprint evidence and analysis is generally accepted."). *But see* United States v. Crisp, 324 F.3d 261, 276 (4th Cir. 2003) (Michael, J., dissenting) ("Nothing in the record in this case shows that the fingerprint examination community has challenged itself sufficiently or has been challenged in any real sense by outside scientists.").

next Part argues that climate science shares characteristics with forensics, but also that on almost every *Daubert* factor, it has a stronger claim to admissibility.

III
Climate Testimony Is Legally Cognizable Expert Testimony

Climate science does not fit neatly into the science category under *Daubert*, but instead shares sufficient characteristics with technical experts under *Kumho Tire*. Section III.A explains the scientific inquiry into climate science; Section III.B assesses climate science under the *Daubert-Kumho Tire* approach, revealing similarities with the forensic experts explained in the previous Part. Additionally, by reviewing the strengths of the discipline, it lays a foundation for Part IV to compare climate experts to forensic experts. Many of the factors that make climate science close to legally recognized science for admissibility purposes make a strong case for finding reliability.

A. Climate Science for Dummies

Climate change science is a rapidly developing field of scientific inquiry, but its roots reach back much further into several long-established disciplines. Climate science represents the intertwining of different fields of inquiry, including "the likes of meteorology, oceanography, glaciology, some aspects of geography, and distinctive categories of earth sciences."[125] As the public grew more interested in climate change in the 1970s and 80s, the fields converged as a way of both mapping the environment as it existed and understanding the impacts of human activity.[126] What was once an ad hoc convergence of studies has evolved into a body of knowledge with "the organizational features of a scientific discipline, such as peer-reviewed journals, departments found at top research universities, and advanced degrees."[127] Climate science seeks to understand the earth's climate in two steps. It first seeks to uncover the climate conditions of the natural world; second, it looks to gauge how "atmospheric concentrations of greenhouse gases . . . will affect climate change" by "isolat[ing] climate impacts attributable to such higher elevations from those attributable to natural variability in the

125 Dennis Bray & Hans von Storch, *Climate Science: An Empirical Example of Postnormal Science*, 80 Bull. Am. Meteorological Soc'y 439, 439 (1999).

126 *Id.* at 440.

127 Engel & Overpeck, *supra* note 21, at 13.

climate system."[128] Climate scientists use models to understand the different impacts of inputs like "solar radiation, volcanic material in the atmosphere, and the atmospheric concentration of greenhouse gases" on four main interacting components of the system: atmosphere, land, surface, and ocean and sea ice.[129] The lack of a second Earth precludes natural experiments, to discern whether an effect would have occurred in the absence of human activity. Therefore, much climate science focuses on running experiments with models to forecast future climate scenarios.[130] To validate their findings, scientists apply their models backwards in time, or "hindcast," to give them confidence in their future projections.[131] Scientists gather data about these prior states through direct physical evidence. Ice cores from icebergs, fossilized trees, sediment and cave formations, and other evidence of the paleoclimate serve as critical components in ensuring the validity of the models.[132] Understanding how the system works holistically is also crucial, as different parts of the climate can create "positive feedback" loops that can hasten climate change.[133] For example, increasing global temperatures accelerate the melting of the ice sheets, which in turn decreases the amount of reflective surface on the planet that can redirect solar radiation back into space.[134]

An important point of contestability, especially in the litigation context, is how accurately these methods can attribute specific climate events or impacts to specific actors. Termed "downscaling," the process challenges climate scientists to understand the effects of climate change on particular regions or even particular severe weather events.[135] This method in particular has drawn fire from climate skeptics as lacking rigor, as it is in fact more difficult to attribute discrete weather events or regional effects to causes in the changing global climate.[136]

However, climate science has rapidly advanced. Different models and methods are reflected in the compendium report, the IPCC

---

128 *Id.*

129 *Id.* at 15.

130 David A. Randall & Richard A. Wood, *Climate Models and Their Evaluation*, *in* Climate Change 2007: The Physical Science Basis 589, 594 (Susan Solomon et al. eds., 2007).

131 *Id.* at 601.

132 Engel & Overpeck, *supra* note 21, at 16.

133 *Id.*

134 *Id.*

135 *Id.* at 17.

136 Morrison et al., *supra* note 23, at 412–13 (quoting *Downscaling*, Climate Change Austl., https://www.climatechangeinaustralia.gov.au/en/climate-campus/modelling-and-projections/climate-models/downscaling [https:/perma.cc/KZ9Q-9YJR]) (explaining that despite the increased prevalence in downscaled datasets, the process of applying global trends to specific regions produces information that is "*weakly coordinated*" and "can have *significant errors*").

Assessment Report 6 (AR6). The creation of the IPCC in 1988 marked the first serious scientific collaboration and international institutionalization of climate studies.[137] This Note focuses on the IPCC reports, as they are the most comprehensive compendia of climate science, representative of the wide array of models and findings gathered by the scientific community studying this issue. It is also likely that their findings and methods will be immediately challenged as insufficiently reliable in future litigation, as the literature suggests.[138] More research has turned to specific event attribution in the last seven years, as the AR6 was published in 2021 (AR5 came out in 2014).[139] The IPCC finalized the first part of this report, the Working Group I (WGI) contribution, on August 9, 2021.[140] The WGI functions as a meta-analysis of existing climate research. Two hundred thirty-four experts from sixty-four countries served as lead authors, who coordinated with teams of other authors—another thirty-six review editors contributed to distributing and evaluating the 78,000 expert and governmental comments to the report.[141] The authors and author teams prepared a first draft, for which other expert reviewers (who were not part of the preparing author teams and were able to self-declare their expertise) submitted substantive comments.[142] Therefore, the report represents a collaborative and iterative process of knowledge gathering, synthesis, and communication.

To make assertions about anthropogenic impact on climate change, the author teams evaluated each claim and expressed both a measure of confidence and a probability. Confidence is "a qualitative measure of the validity of a finding, based on the type, amount, quality, and consistency of evidence . . . and the degree of agreement."[143] The teams start by evaluating the available science and agreement, and describe the type and quality of data as limited, medium, or robust, with the degree of

---

137 *Id.*

138 *Id.*; *see also* Hasani, *supra* note 29, at 84 (arguing that climate "testimony cannot meet the legal standard for admissibility" due to uncertainty surrounding the climate data and models).

139 Paola A. Arias et al., *Technical Summary*, *in* Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change 41 (Valérie Masson-Delmotte et al. eds., 2021) [hereinafter *AR6 Technical Summary*].

140 *Sixth Assessment Report*, *supra* note 36.

141 *Authors*, IPCC, https://www.ipcc.ch/report/ar6/wg1/about/authors [https://perma.cc/XJ6S-ERZZ].

142 IPCC, Appendix A to the Principles Governing IPCC Work: Procedures for the Preparation, Review, Acceptance, Adoption, Approval, and Publication of IPCC Reports (2013).

143 *AR6 Technical Summary*, *supra* note 139, at 38.

agreement as low, medium, and high.[144] If they found a sufficient basis to conclude there was scientific agreement on a proposition, they then assigned a confidence level of very low, low, medium, high, or very high.[145] Likelihood is a "quantitative measure of uncertainty in a finding" often based on probabilistic expressions in quantitative model results or surveys of expert judgment.[146]

The WGI Technical Summary proceeds through many different human impacts on the environment and rates each with these different criteria. It also (helpfully) discusses the changes from the AR5 report from 2014. Many ratings have now been upgraded, due to better modeling and more robust data sets, to "high confidence" or even "established fact."[147] Examples include a range of phenomena, from the broader claims about human impact on the climate system globally, to more granular claims about extreme climate events.[148] In addition to greater confidence in global anthropogenic warming attribution, the report now labels "high confidence" levels for regional effect attribution for hot extremes, and "medium confidence" for heavier precipitation and agricultural and ecological drought.[149]

However, even with such strong claims, should climate science be afforded the same status as other sciences under the law?

B. Climate Science Exhibits Many Similarities to Other Technical Methods

This Note argues that the forward-looking nature of climate modeling likely makes it insufficient to pass the "scientific" *Daubert* standard,

---

144 Deliang Chen, Maisa Rojas & Bjørn H. Samset, *Framing, Context, and Methods*, *in* Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change 169 (Valérie Masson-Delmotte et al. eds., 2021) [hereinafter *AR6 Framing*].

145 *Id.*

146 *Id.* These were rated "virtually certain: 99–100%, very likely: 90–100%, likely: 66–90%, about as likely as not: 33–66%, unlikely: 0–33%, very unlikely: 0–10% . . . ." Further, these measures can interact. For example, "the probability of low-likelihood, high impact outcomes increase with higher global warming levels (high confidence)" or "[b]ased on multiple lines of evidence, the *very likely* range of equilibrium climate sensitivity is between 2 degrees Celsius (*high confidence*) and 5 degrees Celsius (*medium confidence*)." *Id.* at 170.

147 *AR6 Technical Summary*, *supra* note 139, at 41 (noting the change from "human influence on the climate is clear" in AR5 to "human influence on the climate system is now an established fact").

148 *Id.* at 42 (While AR5 "assessed that human influence had been detected in changes in some climate extremes," an entire chapter in AR6 "concludes that it is now an established fact that human-induced greenhouse gas emissions have led to an increased frequency and/or intensity of some weather and climate extremes since 1850, in particular for temperature extremes.").

149 *Id.* at 108–09.

and that climate scientists are better thought of as expert technicians and modelers. This Section deals mainly with the IPCC AR6, while also evaluating the underlying models and work that comprise it. However, it is important to note that the IPCC methodology itself tends to alleviate some concerns around key factors like subjectivity (as meta-analysis can address problems like bias although judges must still scrutinize the inputs of the data. The factors follow, with discussions and critiques.[150]

1. Falsifiability

Climate science is not falsifiable in the same way as other scientific disciplines. Critics persuasively argue that climate evidence is not strictly "scientific," in that climate models merely replicate complex systems in nature and cannot be validated.[151] They also argue that such modeling is not reproducible because of the inherently subjective nature of the process.[152] This latter argument is the stronger of the two, and while such characterizations are exaggerated, climate science does suffer from pure testability and reproducibility.

First, critics claim that climate science is merely based on projections into the future that solely model the world, not actually observe it.[153] They contend that these predictions can never really be tested. Critics cite cases where courts excluded expert witness testimony where the experts "failed to validate [hypotheses] . . . with testing."[154] They also reject hindcasting as a metric of reliability—because, the line goes, even when climate models can duplicate the past, this is no assurance they can accurately predict the future.[155]

Such criticism is not completely misplaced. Modeling is not literally descriptive of the real world—experts create models *because* it would be too difficult to isolate certain variables to find causal connections. Understanding a but-for world is therefore the purview of models introduced in litigation all the time. Models are necessarily subjective because in attempting to isolate cause and effect, a modeler must make assumptions. Whether it is understanding whether New York City police tactics racially profiled citizens of that city,[156] or whether appraisers over-estimated the value of residential properties preceding the housing

---

150 Anna-Bettina Haidich, *Meta-Analysis in Medical Research*, 14 Hippokratia 29, 30 (2010).

151 Hasani, *supra* note 29, at 25.

152 Morrison et al., *supra* note 23, at 418.

153 Hasani, *supra* note 29, at 25.

154 Morrison et al., *supra* note 23, at 417 n.148 (quoting Nease v. Ford Motor Co., 848 F.3d 219, 232 (4th Cir. 2017)).

155 *Id.* at 418.

156 Floyd v. City of New York, 959 F. Supp. 2d 540 (S.D.N.Y. 2013).

bubble collapse,[157] or whether Ivy League universities discriminated in their admissions processes against Asian-American students,[158] all of these models relied on isolating effects by considering but-for worlds: if minority citizens were only stopped according to their proportion of the population, if the homes were valued according to certain criteria and neighborhood valuations, or if admissions were only dependent on a single factor like grades. Thus, in this respect, climate experts could be considered more akin to statisticians or econometricians, who model the world using assumptions, and whose testimony is regularly admitted under *Kumho Tire*.

The rejection of backward-looking validation seems to cede that climate science is indeed falsifiable à la *Daubert*. It also underscores another oversight: Climate modeling can prove that past events occurred because of climate change. This past versus future distinction is important, as models that seek to prove past effects are even closer to legal science. Describing events in the past through models and confirming with hindcasting removes speculation about the future. For predicting future events, the case does become weaker. Still, such a myopia argues only the future results can confirm current climate models.[159] This overlooks the perhaps obvious fact that climate modeling can validate results from models already developed and verified. Indeed, validating different predictions in mean global temperature increases since the first IPCC AR1 in 1990 works to validate prior observations. Scientists re-analyze past predictions with data from after the projection was made,[160] and these prior models are "generally able to project actual future global warming."[161]

Lastly on this first critique, the rejection of hindcasting is the weakest criticism, inaccurately trying to distinguish climate modeling from other regularly-accepted models that use the technique. Indeed, financial models use the same technique (termed backtesting) for all

---

157 Fed. Hous. Fin. Agency v. Nomura Holding Am., Inc., 104 F. Supp. 3d 441 (S.D.N.Y. 2015).

158 Students for Fair Admissions, Inc. v. President and Fellows of Harvard Coll., 397 F. Supp. 3d 126 (D. Mass. 2019).

159 *See* Hasani, *supra* note 29, at 98–99.

160 *See* Zeke Hausfather, Henri F. Drake, Tristan Abbott & Galvin A. Schmidt, *Evaluating the Performance of Past Climate Model Projections*, 47 Geophysical Rsch. Letters (2020); *see also* Ulrich Cubasch & David Wuebbles, *Introduction*, *in* Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change (Thomas F. Stocker et al. eds., 2013).

161 *AR6 Framing*, *supra* note 144, at 186.

sorts of models, from understanding credit risk[162] to creating trading portfolios.[163] Further, one of the many reasons the accuracy of climate science has increased (besides the ability for computers to analyze massive amounts of complex geospatial data) is that models are constantly updated, with old assumptions giving way to new methods.[164] The uncovering of more paleoclimate evidence that can create a better picture of what the past climate looked like, allowing for more accurate hindcasting, has led to a much richer picture of the past. In the first IPCC assessment, AR1 in 1990, scientists only had access to five million years of temperature data, five million years of sea change data, and only 160,000 years of CO_2 data—now, they can work with sixty-five million years, fifty million years, and four hundred fifty million years of data for each category respectively.[165] This evidence enables climate scientists to more accurately understand future scenarios based on confirmatory evidence from Earth's Paleogenic past.

The second critique about the subjectivity of modeling is likewise persuasive under *Daubert*. Modelers make independent decisions on the proper scope and object of their analysis[166] and aspects of their model like parameters and assumptions about the data.[167] They must choose between different models,[168] and then "calibrate" the data by adding coefficients and other mathematical terms to further refine the

---

162 *See, e.g.*, Hangyu Ma & Lei Yue, *Credit Analytics Statistical Models' Backtesting and Recalibration: A Primer*, S&P Global (Feb. 7, 2020), https://www.spglobal.com/marketintelligence/en/news-insights/research/demystifying-credit-risk-models-backtesting-and-recalibration [https://perma.cc/F89D-2SDS] ("Model backtesting and recalibration are important and natural stages in the lifecycle of any statistical model and should be performed on an annual basis.").

163 CFI Team, *Backtesting: Applying a Strategy or Predictive Model to Historical Data to Determine Its Accuracy*, Corp. Fin. Inst. (Dec. 28, 2022), https://corporatefinanceinstitute.com/resources/data-science/backtesting [https://perma.cc/G6VD-QHGC] ("Analysts use backtesting as a way to test and compare various trading techniques without risking money. The theory is that if their strategy performed poorly in the past, it is unlikely to perform well in the future (and vice versa).").

164 *AR6 Framing*, *supra* note 144, at 184.

165 *Id.* at 245.

166 Arthur F. Lutz, Herbert W. ter Maat, Hester Biemans, Arun B. Shrestha, Philippus Wester & Walter W. Immerzeel, *Selecting Representative Climate Models for Climate Change Impact Studies: An Advanced Envelope-Based Selection Approach*, 36 Int'l J. Climatology 3988, 3989 (2016) (describing how the variables considered in a model can shift depending on the "character and goals" of the study).

167 *Id.*

168 *Id.*

analysis.[169] Critics charge that this allows for motivated results,[170] but this same criticism can be levied against any model, especially those generated for the purposes of litigation.[171] Issues with reproducibility could be softened by the fact that peers can understand a model if it conforms to those used by other scientists in the field (part of the general acceptance test). If it is not merely a black box, the model is acceptable. Thus, the lack of reproducibility might not be fatal, as other climate scientists or opposing experts could go back and try to see why older models did not work as well as predicted.[172]

However, overall, it is true that climate science is more akin to other disciplines that model, instead of, say, a chemist testing different substances in a closed laboratory. Indeed, an important EPA guidance document on environmental modeling discusses how the "subjective evaluations of experts may be needed to determine appropriate values for model parameters and inputs that cannot be directly observed or measured," and such evaluations constitute an exercise of "expert judgment."[173] Thus, under the existing *Daubert* regime, falsifiability would likely preclude such subjective modeling.[174]

2. Peer Review

This factor does not need as much analysis, as both the IPCC and individual model methods rely heavily on iterative peer review to validate findings. The IPCC and other compendium reports evaluate propositions under a "weight of the evidence" approach, "falsifying hypotheses through multiple pathways, including verification of model

---

169 Orrin H. Pilkey & Linda Pilkey-Jarvis, Useless Arithmetic: Why Environmental Scientists Can't Predict the Future 43 (2007) ("Modeling equations are sometimes modified and altered, until the model correctly 'predicts' an already known natural event.").

170 Matthew W. Swinehart, *Remedying* Daubert*'s Inadequacy in Evaluating the Admissibility of Scientific Models Used in Environmental-Tort Litigation*, 86 Tex. L. Rev. 1281, 1292 (2008) ("Model critics mockingly refer to such coefficients and constants as 'fudge factors.'").

171 Fed. Hous. Fin. Agency v. Nomura Holding America, Inc., No. 1:11-cv-6201–DLC, slip. op at 112 (S.D.N.Y. May 11, 2015), ECF No. 1686 (considering, and ultimately rejecting, suggestions of bias in an expert report).

172 *See, e.g.*, Hausfather, *supra* note 160, at 7–8.

173 EPA, Off. of the Sci. Advisor, EPA/100/K-09/003, Guidance on the Development, Evaluation, and Application of Environmental Models 46 (2009).

174 There is a strong argument that *Daubert* does a poor job of handling environmental modeling—especially with falsifiability. *See generally* Swinehart, *supra* note 170, at 1301–05. An overemphasis on reproducibility and testability attempt to remove uncertainty that cannot be removed, misunderstanding the nature of the practice. Courts do not throw out other modeling because of a modeler's ability to empirically remove all uncertainty. *See, e.g.*, Floyd v. City of New York, 959 F. Supp. 2d 540, 583 (S.D.N.Y. 2013) (approving of a conservative statistical assumption, where the number of Fourth Amendment violations would "almost certainly never be known").

projections through observed phenomen[a]."[175] The expansion of peer-reviewed literature has also accelerated: "Recently, scientific climate change research has doubled in output every 5-6 years; the majority of publications deal with issues related to the physical climate system."[176] The most important models, like the Climate Model Intercomparison Project (CMIP) 6, which is a sort of super model of one hundred models from fifty modelling centers,[177] has been subject to intense scrutiny by the scientific community.[178] Therefore, peer review of these models happens all the time. This level of peer review would lean heavily towards admissibility even under a scientific *Daubert* standard.

3. Error Rate

A court's assessment of error rate would follow a similar line as falsifiability because the ability to test is integral to the ability to prove hypotheses true or false. Climate scientists utilize modeling because there is no second Earth from which one could create a randomized experiment. It is therefore difficult to describe the amount of times that an individual model creates accurate predictions or not.[179] However, climate scientists constantly look to previous projections to assess accuracy.[180] These exercises show that there is an error rate associated with certain models. As noted with testability, scientists can test and find the error rate of earlier predictions from the twentieth century. In one such analysis, researchers reviewed models from the 1970s to the late 2000s and found that "climate models published over the past five decades were generally quite accurate in predicting global warming."[181]

---

175 Engel & Overpeck, *supra* note 21, at 25.

176 *AR6 Framing*, *supra* note 144, at 243.

177 *Latest Projections of Future Climate Now Available*, Copernicus Climate Change Serv. (Mar. 23, 2021), https://climate.copernicus.eu/latest-projections-future-climate-now-available [https://perma.cc/W2TR-YB87].

178 *See generally* Zeke Hausfather, Kate Marvel, Gavin A. Schmidt, John W. Nielsen-Gammon & Mark Zelinka, *Climate Simulations: Recognize the 'Hot Model' Problem*, Nature (May 4, 2022), https://www.nature.com/articles/d41586-022-01192-2 [https://perma.cc/CM2U-NWGQ] ("Numerous studies have found that these high-sensitivity models do a poor job of reproducing historical temperatures over time and in simulating the climates of the distant past"); *see id.* ("A 2020 community review (that four of us co-authored) combined lines of evidence from palaeoclimate, observations of surface temperatures and ocean heat content, and models of physical processes. It concluded that the equilibrium climate sensitivity (ECS) is likely (with a 66% chance) to be in the range of 2.6–3.9 °C, and very likely (with a 90% chance) to lie between 2.3 and 4.7 °C.").

179 Encyclopedia of Epidemiology 360 (Sarah Boslaugh ed., 2008) ("Error rate suggests quantifiable errors such as the p value or measures of Type 1 and Type 2 errors.").

180 *AR6 Framing*, *supra* note 144, at 184.

181 *See* Hausfather et al., *supra* note 160, at 1.

Perhaps conceding falsifiability, critics acknowledge that error rates do exist, and critique the AR5's findings that 111 of 114 models reported greater warming than observed from 1998 to 2012, a "97 percent error rate."[182] Attacks like this cherry-pick the data, selecting a very narrow timeframe where warming did slow, due to natural variability.[183] Moreover, overshooting a projected estimate within a certain range does not show the predictions were "incorrect." The study fails to account for whether the true value fell within a range of error.[184] In this regard, critics unfairly raise the bar against climate science, arguing somewhat incomprehensibly that climate model predictions "have no known error rate because what they are being used for—to predict anthropogenic global warming—is a one-of-a-kind event."[185] Taken at its broadest and most generous interpretation, that climate models must "project how average weather will change decades in advance,"[186] this claim reveals a misunderstanding of the work climate scientists do and the results they find. Evidence suggests with a high degree of confidence that climate change is *already* occurring.[187] As a result, climate change is not a one-of-a-kind event but instead a multi-faceted, multi-impact, multi-year phenomenon that scientists currently observe. Critics also claim that "the range of projected warming rates has not been reduced in over twenty years of climate model development" and that this "is indicative of our lack of understanding."[188] While this is factually correct, increases in understanding from multiple data streams over the past fifteen years have also increased confidence in that range.[189]

Climate modelers can also find error rates due to the increasing standardization of model methodology and datasets. Voluminous reports with these types of assumptions are standard for *Daubert* hearings that introduce any type of modeling.[190] Most generally, there are only three main types of climate models: energy balance models, intermediate

---

182 Morrison et al., *supra* note 23, at 419.

183 Intergovernmental Panel on Climate Change, Climate Change 2014: Synthesis Report 43 (2015), https://ar5-syr.ipcc.ch/ipcc/ipcc/resources/pdf/IPCC_SynthesisReport.pdf [https://perma.cc/2ATY-48GW].

184 Morrison et al., *supra* note 23, at 419.

185 Brooks E. Harlow & Roy W. Spencer, *An Inconvenient Burden of Proof? CO2 Nuisance Plaintiffs Will Face Challenges in Meeting the* Daubert *Standard*, 32 Energy L.J. 459, 482 (2011).

186 *Id.*

187 *See infra* notes 202–04 and accompanying text.

188 *See infra* note 139, at 483.

189 *AR6 Framing*, *supra* note 144, at 183 tbl 1.2. Note especially the assessed range of 1.5–4.5 degrees Celsius (*likely*) in the AR6 (2013) to the 2.5–4.0 (*likely*), 2.0–5.0 (*very likely*). While the "central estimate" of ECS has not changed (but fluctuated) from 1979 studies, the confidence in that range has increased.

190 *See supra* notes 150–52.

complexity models, and general circulation models.[191] Each of these types of models are refined by datasets widely used by large groups of practitioners, allowing scientists to situate and understand a model's framework and assumptions.[192] There are several important intergovernmental and nongovernmental initiatives that work to standardize and allow for cross-model comparison—the most well-known and widely utilized is the World Climate Research Programme's Coupled Model Intercomparison Project (CMIP). The goal of this initiative is "to generate a set of standard simulations that each model will run . . . allow[ing] results to be directly comparable across different models."[193] Such standardizing efforts allow for easier comparisons, even when scientists run different experiments. This project seeks to increase transparency, as the simulations are openly accessible for any scientist who wants to utilize them.[194] Accordingly, 45% of climate research papers published in 2016 cited the CMIP Phase 5 (the fifth generation of the project).[195] In sum, the charge that "there is no way of allowing for model error by sampling the space of all possible models in a representative way, because distance within this space is undefinable"[196] does not substantively engage with the existing climate science methodology.

Further assuaging doubts, the IPCC AR6 includes confidence intervals regularly accepted by courts. It combines the confidence intervals of the data and experiments it surveys to create its own metrics. When results are deemed "very likely," they fall within a 90% confidence interval.[197] This means that there is a 90% chance that the true mean is within a certain range (*not* that 90% of the data falls within that range). This represents an important advance over the science in the AR5. The AR5 could only produce 60% confidence intervals for many propositions; the 90% in AR6 comes within an acceptable range of the 95% confidence level of two standard deviations, accepted by many courts and scientists.[198] Indeed,

---

191 Andreas Schmittner, Introduction to Climate Science 133 (2018).

192 *Id.*

193 Zeke Hausfather, *CMIP6: The Next Generation of Climate Models Explained*, Carbon Brief (Feb. 12, 2019, 8:00AM), https://www.carbonbrief.org/cmip6-the-next-generation-of-climate-models-explained [https://perma.cc/N68J-GMJB].

194 *Id.*

195 David Carlson, Veronika Eyring, Narelle van der Wel & Gaby Langendijk, *WCRP's Coupled Model Intercomparison Project: A Remarkable Contribution to Climate Science*, Euro. Geosciences Union (July 4, 2017), https://www.egu.eu/news/highlight-articles/586/wcrps-coupled-model-intercomparison-project-a-remarkable-contribution-to-climate-science [https://perma.cc/27G8-SV88].

196 Morrison et al., *supra* note 23, at 418 (citing Myles Allen, Jamie Kettleborough & David Stainforth, *Model Error in Weather and Climate Forecasting*, 8 Nonlinear Processes Geophysics 275, 275 (2001)).

197 *AR6 Framing*, *supra* note 144, at 170.

198 *See supra* notes 20–23 and accompanying text.

new criticism argues that such an obsession with the 95% standard is undue, and encourages the rejection of the "dichotomania" between results that lie on either side of the 95% line.[199] Context is key, and therefore "[a] different level can be justified, depending on the application."[200] Accordingly, many claims in the AR6, especially about human influence on the climate system, are now arguably made with sufficient empirical confidence. Courts should be flexible in understanding confidence intervals in the way scientists have recently reevaluated them.

Arguments about the inaccuracy of "downscaling" must be addressed, especially for individual models. In litigation alleging causal relationships between emissions and damages to particular regions or areas, the science attributing fault to a particular actor can present more uncertainty.[201] However, in AR6, there is *high confidence* that human influence has contributed directly to extreme precipitation, droughts, tropical cyclones, and compound extremes, and "some recent hot extreme events would have been extremely unlikely to occur without human influence on the climate system."[202] There is also now more research that can attribute to specific events; for instance, the AR6 found that there is *high confidence* that "anthropogenic climate change contributed to extreme rainfall amounts during Hurricane Harvey and other intense tropical cyclones."[203] Many phenomena still remain understudied—there is, for example, *low confidence* that human influence has affected meteorological droughts (although *medium confidence* that human-induced climate change contributed to the probability or

---

199 Valentin Amrhein, Sander Greenland & Blake McShane, *Scientists Rise Up Against Statistical Significance*, 567 Nature 305, 306 (2019), https://www.nature.com/articles/d41586-019-00857-9 [https://perma.cc/2FHG-6XDL] (writing for over 800 co-signers "call[ing] for the entire concept of statistical significance to be abandoned"). Some courts have accepted 90% confidence intervals to show statistical significance. *See, e.g.*, Burst v. Shell Oil Co., Civ. Action No. 14-109, 2015 WL 3755953, at *6 n.16 (E.D. La. June 16, 2015) (noting that epidemiological studies may be statistically significant when "employ[ing] a 90% confidence interval"), *aff'd*, 650 F. App'x 170 (5th Cir. 2016); Dyson, Inc. v. Euro-Pro Operating LLC, No. 14 C 9442, 2015 WL 1120006, at *3 (N.D. Ill. Mar. 10, 2015) (accepting expert testimony of statistical significance for vacuum-cleaning ability effectiveness at a 90% confidence interval).

200 Amrhein et al., *supra* note 199, at 307.

201 The IPCC in AR5 specifically said information around downscaling was "weakly coordinated, and current results indicate that high-resolution downscaled reconstructions of the current climate can have significant errors. The increase in downscaled data sets has not narrowed the uncertainty range." *Part B: Regional Aspects*, *in* Climate Change 2014: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change 1137–38 (Vicente R. Barros et al. eds., 2014), https://www.ipcc.ch/site/assets/uploads/2018/02/WGIIAR5-PartB_FINAL.pdf [https://perma.cc/UU9H-RDYB].

202 *AR6 Technical Summary*, *supra* note 139, at 108.

203 *Id*. Although, tracing those impacts to a single company is a different challenge on which courts have not yet succeeded.

intensity of recent agricultural and ecological droughts).[204] Consensus has emerged around certain events for which there are multiple lines of evidence—therefore, "many highly impactful extreme weather events have not been studied in the event attribution framework, particularly in the developing world where studies are generally lacking."[205] With advances in climate science over the past seven years, AR6 and more ground-level studies have worked to tackle the problem of attribution.

Overall, then, for future-looking error rates, climate science is vulnerable to the attacks of its critics. However, the advances represented by the AR6 present legally adequate confidence intervals, even down to event attribution and regional temperature differences. For individual models, an error rate may be harder to ascertain, although by relying on compendium projects like CMIP 6, courts may find satisfactory methodology to substantiate the error rate figures. This factor therefore might tilt against individual models depending on the model (especially if other pathway evidence did nothing to substantiate its findings).

4. Methods Controlling Operation

As climate modeling continues its rapid advance, researchers have also continued to standardize approaches to climate modeling. As noted, the CMIP project aims to "generate a set of standard simulations that each model will run," therefore "allowing results to be directly comparable across different models, to see where models agree and disagree on future changes."[206] To simplify these complex models, there are two large, standardized inputs into this project and the AR6 more broadly: Shared Socioeconomic Pathways (SSPs) and Representative Concentration Pathways (RCPs). SSPs attempt to capture possible emissions scenarios, contingent on different societal development and energy usage projections.[207] RCPs present "sample trajectories" of how the greenhouse gas effect might play out—and by marrying this data with the SSPs, CMIP 6 created more robust datasets and modeling of a warming world.[208] Climate scientists are free to create their own models from scratch, but the AR6's

---

204 *Id*.

205 *Id*.

206 Hausfather, *CMIP 6*, *supra* note 193.

207 *See* Keywan Riahi et al., *The Shared Socioeconomic Pathways and Their Energy, Land Use, and Greenhouse Gas Emissions Implications: An Overview*, 42 Glob. Env't Change 153 (Jan. 2017), https://www.sciencedirect.com/science/article/pii/S0959378016300681 [https://perma.cc/TKG5-BX6F].

208 *See* Colin Gannon & Gain Boonvanich, *IPCC's Use of Latest Generation Models Will Help Better Assess Real-World Impact of Climate Change*, Moody's ESG (Sept. 1, 2021), https://esg.moodys.io/insights-analysis-reports/ipccs-use-of-latest-generation-models-will-help-better-assess-real-world-impact-of-climate-change [https://perma.cc/M4V5-VKZ3].

reliance on the CMIP 6 for its findings undergirds its importance.[209] So, in short, there are methods and standards to which most climate scientists look, and on which they rely. Beyond that, there is peer review, and with enough visibility into the three different types of models there appears to be sufficient standards to understand how a model works. A survey of 280 papers based on CMIP 5 published between 2012 and 2018 in six leading climate journals found that while 125 of the papers looked to project "future climate change and associated uncertainties," 86 sought to validate CMIP simulations with observations, and 98 to interpret model results.[210] Therefore the methods that are commonly accepted and utilized in the field are subject to scrutiny and standards to help control climate scientists' work. This tilts toward scientific admissibility.

5. General Acceptance

This should be the least worrisome *Daubert* factor. Ninety-seven percent of actively publishing climate scientists believe humans are causing global warming.[211] While the subject has become a political football and increasingly polarized, the scientific community that actively devotes their research hours to the topic is not—they are unanimous.[212] To reach this conclusion, they rely on the modeling described in detail thus far. This is also accepted by the broader scientific community. It would be a vast overstatement that every prediction rises to the level of scientific fact (an extremely likely occurrence, in probabilistic terms) and the term is only used twice in the AR6 WGI Technical Summary: "Human influence on the climate system is now an established *fact*"[213] and "it is now an established *fact* that human-induced greenhouse gas emissions have led to an increased frequency and/or intensity of some weather and climate extremes since 1850."[214] The report goes through pains to couch its findings in confidence and likelihood terms, immediately next to the proposition made.[215] Climate scientists take appropriate

---

209 *Id*.

210 Ludovic Touzé-Peiffer, Anouk Barberousse & Hervé Le Treut, *The Coupled Model Intercomparison Project – History, Uses, and Structural Effects on Climate Research*, 11 Wiley Interdisc. Revs.: Climate Change 4 (2020).

211 *Do Scientists Agree on Climate Change?*, NASA (July 13, 2023), https://climate.nasa.gov/faq/17/do-scientists-agree-on-climate-change [https://perma.cc/4RWE-KXRJ].

212 *See, e.g.*, J. S. Carlton, Rebecca Perry-Hill, Matthew Huber & Linda S. Prokopy, *The Climate Change Consensus Extends Beyond Climate Scientists*, 10 Env't Rsch. Letters 1, 3 (2015) (finding among surveyed non-climate scientists that "91.9% . . . believed in anthropogenic climate change"); *see* Sedona Chinn, P. Sol Hart & Stuart Soroka, *Politicization and Polarization in Climate Change News Content, 1985-2017*, 42 Sci. Commc'n 112, 113 (2020).

213 *AR6 Technical Summary*, *supra* note 139, at 41 (emphasis added).

214 *Id*. at 42 (emphasis added).

215 *Id*.

steps to ensure their methods, and their conclusions, are not overbroad or exaggerate the state of the science—but they unequivocally accept modeling as a valid way of understanding these changes.

6. *Summary: Climate Science Is Likely Not Sufficiently "Scientific" Under* Daubert

What should a court conclude from the foregoing discussion? The modeling approach climate scientists take is more akin to other technical experts like statisticians that attempt to establish causal connections using but-for worlds—climate scientists cannot run double-blind studies in a lab. The peer review community is critical to the success of climate science and leans more towards a true scientific process. For error rates, while there may be some solid evidence for the past accuracy of certain models, and legally recognized confidence intervals for the IPCC, overall, there are not error rates that can be quantified in a sufficiently rigorous scientific way. Standardizing projects look to make models more accurate and easily understandable, and most scientists generally accept climate modeling as a powerful and appropriate method of scientific investigation. So where does this leave climate science?

IV
Comparing Forensics to Climate Science

This Part concludes the Note by charting a path out of the difficulties posed by the strictures of *Daubert*. It proposes that instead of solely looking at climate experts as scientists, courts should assess their admissibility as technical witnesses with specialized or other knowledge. On a strict reading of *Daubert*, climate science is appreciably stronger than forensics—thus, under the more flexible addition of *Kumho Tire*, impliedly utilized by courts for forensics, climate science should a fortiori find admissibility. This last Part bolsters this conclusion by directly comparing the two types of evidence assessed in Parts II and III. The following discussion is summarized in Table I appended to this Part.

A. *Climate Expert Testimony and the IPCC AR6 Should Be Considered Technical Evidence Under* Kumho Tire

This Note takes the position that climate scientists should be legally considered *almost* scientists—in effect, hyperqualified technicians. The lack of sufficient scientific (legally defined) rigor leads critics to find that courts should preclude climate evidence from admissibility. However, this conclusion is not ineluctable. Instead, the evidence presents a choice to a court: either *still* admit it under *Daubert* as a scientific method,

emphasizing the peer review, controlling methods, and general acceptance prongs; or if a court finds its postulates too subjective and more akin to a technician, it should consider the evidence under *Kumho Tire*.

Opponents of admissibility (and even advocates) do not consider this important move, focusing only on the scientific *Daubert* standard.[216] This fails to comprehend that climate modelers can constitute a technical community. Indeed, even philosophers of climate science consider the role of a climate scientist akin to that of a doctor, with "expertise in the area in which the person is formally trained and continues to practice" with an "ability to address the issue of climate and climate change in accord with formal training and a special mode of discourse."[217]

Importantly, Rule 702 authorizes courts the discretion to make such a determination. *Kumho Tire* itself ruled that Rule 702 did not "segregate[] expertise by type while mapping certain kinds of questions to certain kinds of experts."[218] A judge can admit climate evidence as a technical discipline. And, because many of the elements of climate science would likely be admissible under a strictly *scientific* comparison, climate experts' testimony is even *more* compelling than other groups of technical experts. Under this standard, their testimony should easily make it to the finder of fact.

B. Directly Comparing Forensics to Climate Evidence Provides a Strong Basis for Admissibility

This Section directly compares forensic methods and climate science, cementing the case that judges should admit such evidence.

Falsifiability. Climate science modeling presents some of the same issues with testability as forensics. Individuals create models and bring their subjective biases (conscious or not) to assessing evidence. However, climate models use verifiable datasets and models, as well as employ measures to check their results against past conditions, like other models regularly allowed into evidence.[219] Forensic experts can only check their results with another set of eyes. Thus, climate science has a slight edge over forensics, tilting towards its admissibility.

Peer Review. Peer review for ballistic and fingerprinting journals is much less rigorous than climate science, which is regularly published in some of the most prestigious journals in the scientific world, with some

---

216 *See, e.g.*, Morrison et al., *supra* note 23 (*Kumho Tire* is not mentioned at all in the entire article).

217 *See* Bray & van Storch, *supra* note 125, at 441.

218 *Llera Plaza II*, 188 F. Supp. 2d 549, 563 (E.D. Pa. 2002) (citing Kumho Tire Co. v. Carmichael, 526 U.S. 137, 151 (1999)).

219 *See supra* notes 141–44 and accompanying text.

of the most demanding publication standards in academia.[220] If courts can rely on these much less rigorous trade journals as a reason to admit forensics, climate science should also easily pass this hurdle.

Error Rate. As a result of the subjective identification process, forensic methods have very few large-scale studies that suggest reliable accuracy numbers. While the Ames Laboratory Study and the FBI's black-box study are encouraging data points, individual studies do not seem to create the kind of reliability the Court prescribed in *Daubert*. Other individual studies show much lower and worrisome error rates, like the nearly quarter of state labs that could not identify latent fingerprints from the defendant in *Llera Plaza II*. Compare these rates to that of the Hausfather study of the top global warming models since the 1970s: an 82% accuracy rate.[221] Hundreds of climate science papers are published annually, allowing for a much more robust set of observations and models. Indeed, many claims in the IPCC are made with higher confidence thresholds than ballistics and fingerprinting analysis (90% probabilities).[222] This also is a stronger showing for climate expertise.

Methods Controlling Operation. For ballistics, the AFTE toolmark theory allows the examiner to determine what constitutes "sufficient agreement"; likewise, fingerprinting does not follow a standard number of Galton points to confirm a match. Climate science has compendium reports that are pored over by hundreds of scientists, and findings are qualified with legally-recognized thresholds. For forensics, many examiners do not go through a formal training program, and there is no standardization of training programs that do exist. The high pass rates of exams given to examiners likely also means that such tests are not sufficiently rigorous as to truly test the proficiency of these experts. In contrast, climate scientists undergo significantly more training (PhD-level degrees take years to complete).[223] As described in Part III, there are large models and datasets that are increasingly integrated to standardize methods and increase transparency.[224] While the creation of models is subjective, models can be explained in terms of assumptions, statistical methods, and outcomes. The rigor of peer review also serves as a check on bad modeling and unfounded assumptions.[225] This factor is at least, if not more, tilted in favor of climate scientists.

---

220 *See* Engel & Overpeck, *supra* note 21, at 13.
221 *See* Hausfather et al., *supra* note 160.
222 *AR6 Framing*, *supra* note 144, at 170.
223 *See* Engel & Overpeck, *supra* note 21, at 13.
224 Hausfather et al., *Climate Simulations*, *supra* note 178.
225 *AR6 Framing*, *supra* note 144, at 243.

General Acceptance. Judges should consider the forensic community a group of experts who have "technical, or other specialized knowledge" in the words of Judge Pollack.[226] Thus, under *Kumho Tire*, their expertise can be sufficiently reliable to admit into evidence. Climate science is more akin to a scientific field of inquiry than simply a collection of experts. The field has acceptance not only by reference to its own scientists, but by the broader scientific community.[227] In an interdisciplinary recognition of the rigor and value of the discipline, "climate scientists are the frequent recipients of science's most distinguished awards."[228]

TABLE I. FORENSIC METHODS VS. CLIMATE SCIENCE UNDER *KUMHO TIRE*

	Forensic Methods	**Climate Science**	**Stronger Claim of Admissibility**
Falsifiability	Features: • Dependent on a single examiner—cannot be reproduced but only "confirmed" by a second examiner. • Often not emphasized by courts.	Features: • Models can be reproduced and verified by other scientists; hindcasting and models from the last few decades also can show confirmatory results.	At least even due to subjectivity concerns; reproducibility may give climate science the edge.
Peer Review	Features: • Peer-reviewed journals not considered by scientists as rigorous—more akin to trade journals.	Features: • Peer-reviewed in some of the most prestigious scientific journals in the world. A whole scientific community looks at and attempts to reproduce studies.	Climate Science.
Error Rate	Features: • Few studies, so hard to make larger inferences about accuracy. • Ballistics Ames Study found 1.5% false positive rate; fingerprinting likewise has few studies and in *Llera–Plaza*, about a quarter of state labs could not identify both prints.	Features: • Top global warming models show an 82% accuracy rate since the 1970s. • Meta-analysis from IPCC AR6 finds 90% confidence thresholds for many claims, including specific weather events.	Likely Climate Science, due to greater number of studies and models.
Methods Controlling Operation	Features: • Experts trained relatively quickly • Nothing standardized, and even the most "rigorous" tests are likely far too easy like the FBI proficiency testing (see *Llera Plaza II*).	Features: • Climate scientists must undergo years of schooling (PhDs). • Standardized data sets and model types that allow for rigorous peer review.	Climate Science.
General Acceptance	Features: • Accepted by expert community but not broader scientific community.	Features: • Accepted by the broader scientific community.	Climate Science.

226 *Llera Plaza II*, 188 F. Supp. 2d 549, 563–64 (E.D. Pa. 2002).
227 Engel & Overpeck, *supra* note 21, at 14.
228 *Id.* at 14.

Conclusion

"All models are wrong, but some are useful."
—George Box, 2018[229]

This Note has argued that climate science is sufficiently reliable to pass a FRE 702 *Daubert* test under *Kumho Tire*. While the latest IPCC AR6 report should have little trouble finding its way to the jury, more specific downscaling reports might face greater challenges—but these should be considered much in the same way that courts assess other statistical models. Indeed, *Kumho Tire* leans more towards the reliability of a method, de-emphasizing falsifiability in the analysis. Courts must remember the "liberal thrust"[230] of the Federal Rules of Evidence with respect to expert evidence, and that admissibility hearings are about the reliability of method, not whether at that stage of litigation it is "correct."[231] Especially when cast in the light of other expert testimony admitted under *Kumho Tire*, courts should easily find climate scientists sufficiently qualified and their methods reliable.

[229] George Box, statistician, *quoted in* Guillem Barroso, *"All Models Are Wrong, but Some Are Useful." George E. P. Box*, AdMoRe (May 3, 2018), https://www.lacan.upc.edu/admoreWeb/2018/05/all-models-are-wrong-but-some-are-useful-george-e-p-box [https://perma.cc/C4YH-XWSP].

[230] Daubert v. Merrell Dow. Pharms., Inc., 509 U.S. 579, 588 (1993) (quoting Beech Aircraft Corp. v. Rainey, 488 U.S. 153, 169 (1988)).

[231] Fed. R. Evid. 702 advisory committee's note.