SUPERIOR COURT OF THE DISTRICT OF COLUMBIA
CIVIL DIVISION

MICHAEL E. MANN, PH.D., Plaintiff, v. NATIONAL REVIEW, INC., et al., Defendants.	Case No. 2012 CA 008263 B Judge Natalia Combs Greene Next event: Motion Hearing February 19, 2013

REPLY BRIEF IN SUPPORT OF DEFENDANTS COMPETITIVE ENTERPRISE INSTITUTE AND RAND SIMBERG'S MOTIONS TO DISMISS PURSUANT TO THE D.C. ANTI-SLAPP ACT AND TO RULE 12(b)(6)

DAVID B. RIVKIN, JR. (D.C. Bar No. 394446)
BRUCE D. BROWN (D.C. Bar No. 457317)
MARK I. BAILEN (D.C. Bar No. 459623)
ANDREW M. GROSSMAN (D.C. Bar No. 985166)
BAKERHOSTETLER LLP
Washington Square, Suite 1100
1050 Connecticut Avenue, NW
Washington, DC 20036
Tel: (202) 861-1500
Facsimile: (202) 861-1783
drivkin@bakerlaw.com
bbrown@bakerlaw.com
mbailen@bakerlaw.com
agrossman@bakerlaw.com

Counsel for Defendants Competitive Enterprise Institute and Rand Simberg

TABLE OF CONTENTS

Page

TABLE OF CONTENTS
(continued)

Page

TABLE OF AUTHORITIES

Page(s)

CASES

STATUTES

OTHER AUTHORITIES

ABBREVIATIONS

CEI	Defendant Competitive Enterprise Institute
CEI Defendants	Defendants Competitive Enterprise Institute and Rand Simberg
CRU	East Anglia University's Climatic Research Unit
EPA	Environmental Protection Agency
Ex.	CEI Defendants' Exhibits (numbered items are exhibits to the CEI Defendants' Anti-SLAPP Motion, and lettered items are exhibits to this filing)
ICCER	Independent Climate Change E-mails Review
IPCC	Intergovernmental Panel on Climate Change
P. Ex.	Exhibits to the Plaintiffs' Response to CEI's Motions
SAP	The University of East Anglia's Scientific Assessment Panel
SLAPP	Strategic Lawsuit Against Public Participation

INTRODUCTION AND SUMMARY OF ARGUMENT

Plaintiff Michael E. Mann gets the First Amendment's protection of free speech precisely backwards. Mann claims the right to characterize individuals and groups who disagree with him as being engaged in "pure scientific fraud,"[1] as publishing "bogus" research,[2] as "hired assassin[s],"[3] as "deniers,"[4] as "shills for the fossil fuel industry,"[5] as "deeply unethical,"[6] and as perpetuators of a "crime against humanity."[7] But faced with criticism of his own views, he claims that a government agency has decided the matter once and for all in his favor, that any dissent is therefore false and defamatory, and that CEI is "estopped" from arguing otherwise. The First Amendment, of course, rejects that premise in favor of the "bedrock . . . principle . . . that citizens have a right to voice dissent from government policies." *Tobey v. Jones*, --- F.3d ---, 2013 WL 286226, at *8 (4th Cir. Jan. 25, 2013) (citing *Mills v. Alabama*, 384 U.S. 214, 218 (1966)). That principle is not only central to our

---

[1] Environmental Protection Agency, 3 *EPA's Response to the Petitions to Reconsider the Endangerment and Cause or Contribute Findings for Greenhouse Gases under Section 202(a) of the Clean Air Act* (2010) (quoting email from Michael Mann to Andy Revkin, reporter, N.Y. Times (Feb. 8, 2005)) ("The McIntyre and McKitrick paper is pure scientific fraud."), Ex. A, at 73. (Lettered exhibits are those attached to this reply; numbered exhibits are those attached to the CEI Defendants' Anti-SLAPP Motion; and "P. Exs." are those attached to the Plaintiffs' Response in Opposition.)

[2] Michael E. Mann, *The Hockey Stick and the Climate Wars* 141 (2012).

[3] Interview by James Coomarasamy with Michael Mann and Marc Morano, on BBC World Service Newshour (Nov. 30, 2012), https://soundcloud.com/ameliaf/newshour-moran-mann-climatewar. On his Facebook page, Mann describes this as an "interview on attacks on #climate scientists by industry hired guns like Marc Morano." Michael Mann, Facebook Post (Dec. 2, 2012), http://www.facebook.com/MichaelMannScientist/posts/457448474312117.

[4] *E.g.*, *A Climate Scientist Fights Back: Penn State professor discusses his new book on the climate wars*, Green Light (March 21, 2012), Ex. 1, at 2.

[5] Rick Piltz, Michael Mann Interview: Denialists are waging "asymmetric warfare" against climate science, Climate Science Watch (Mar. 10, 2010), Ex. 35.

[6] Adam Forrest, "*We Need to Adapt . . . Changes are Coming no Matter What": Michael Mann, the US scientist caught up in the 'Climategate' controversy, on why a new sense of urgency is needed*, The Big Issue (April 3, 2012), Ex. 3, at 1.

[7] Ex. 3, at 1; Paul Dechene, *I Won, We Lost*, Planets Magazine (July 26, 2012), Ex. 4, at 4.

system of government, but also to the scientific process, which depends on those willing to challenge prevailing wisdom in the never-ending search for enlightenment. A government report, or even a stack of them, does not and cannot mean, as Mann asserts, that a matter of scientific dispute and public debate has been conclusively "put to rest." Plaintiff's Consolidated Memorandum of Points and Authorities in Opposition ("Opp.") at 28. That Mann considers this debate illegitimate and an obstacle to implementing the public policies that he favors does not undermine that conclusion—quite the opposite.

It is telling that Mann does not mention or address his repeated statements that this lawsuit's purpose is to harass and silence his ideological opponents—statements that are proof positive of a classic SLAPP suit. *See* Memorandum of Points and Authorities in Support of Defendants Competitive Enterprise Institute and Rand Simberg's Special Motion to Dismiss Pursuant to the D.C. Anti-SLAPP Act ("CEI Defendants' Anti-SLAPP Motion"), at 2-3, 23-26. He does not deny that he launched this litigation to intimidate "groups seeking to discredit the case for concern over climate change," with the intent to "silence" them.[8] That is the precise kind of abuse of legal process that the D.C. Anti-SLAPP Act was enacted to curtail or, barring that, punish.[9]

Mann's suit should be dismissed at this stage because the statements he challenges are protected expressions of opinion as a matter of law. The principal defect in Mann's reasoning is that he ignores context. In the context of the Blog Post, it is apparent that the statements Mann challenges are expressions of opinion critical of his research, not accusations of unlawful conduct.

---

[8] Michael Mann, Facebook Post (Oct. 23, 2012), Ex. 9; Michael Mann, Facebook Post (May 16, 2012), Ex. 10.

[9] Council of the District of Columbia Committee on Public Safety and the Judiciary Committee Report, Report on Bill 18-893, "Anti-SLAPP Act of 2010" ("Report on Bill 18-893"), Ex. 8, at 1.

And in the context of the heated global warming debate, the statements of which Mann complains are actually quite temperate. Any doubt on that score may be allayed by consulting Mann's routine use of far harsher language (including *express* accusations of "scientific fraud" and "bogus" research) directed at CEI and other "deniers." In this context, name calling is the norm.

Because Mann ignores context, he proffers an interpretation of the Blog Post that could be shared by no reasonable reader. If, as Mann contends, the Blog Post were asserting that he committed criminal fraud or made up data, why would it link to criticisms of his *scientific methodology*? Why would it link to investigation reports that it describes as "declar[ing] him innocent of any wrongdoing" and that Mann claims "exonerate" him? And why would it conclude by calling for "a fresh, truly independent investigation" of Mann's research, rather than simply demand that he be fired? The only reasonable reading is that the Blog Post is a critical commentary on Penn State's "whitewash[ed]" investigation of Mann.

It is therefore protected under the First Amendment as a supportable interpretation of underlying facts and under District of Columbia law as a fair comment. Mann's response to this point is to assert that the usual legal standard in such cases—whether "*no reasonable person could find* that the [defendant's] characterizations were supportable interpretations" of true underlying facts, *Moldea v. New York Times Co.*, 306 U.S. App. D.C. 1,8, 22 F.3d 310, 317 (D.C. Cir. 1994) ("*Moldea II*")—applies only to "evaluations of a literary work," Opp. at 49. That is false. *See, e.g., Guilford Transp. Indus., Inc. v. Wilner*, 760 A.2d 580, 597-600 (D.C. 2000) (newspaper columnist's statements regarding company's hostility to organized labor protected as supportable interpretation).

Mann's emotional distress claim fares no better, given that he fails to engage, or even mention, the Supreme Court's precedent governing application of the First Amendment to such claims. This omission, however, is understandable, given that the statement Mann challenges "could

not reasonably have been interpreted as stating actual facts about the public figure involved" and so is not actionable. *Hustler Magazine, Inc. v. Falwell*, 485 U.S. 46, 50 (1988).

But even if that claim were not dismissed on the merits, it would have to be dismissed under Rule 12(b)(6) for failure to state a claim, because Mann identifies no allegation in his Complaint that supports the required element of actual malice. His libel claims fail on the same ground.

Finally, Mann's request for attorneys' fees and costs lacks any merit, as it simply repeats several of the more overheated charges from elsewhere in his brief. While plainly frivolous on the merits, it does serve as a timely reminder of Mann's admitted aim in this litigation: to silence his critics through the abuse of legal process and risk of liability. The Court should not allow itself to be used to facilitate Mann's attempt to muzzle opposing points of view on an important issue of intense public interest. Instead, it should carry out the purpose of the D.C. Anti-SLAPP Act and the First Amendment by dismissing his claims against the CEI Defendants.

ARGUMENT

I. Rather than "Exonerate" Him, the Reports that Mann Cites Actually Raise Questions Regarding His Scientific Methodology, Supporting the Blog Post's Point

Although the CEI Defendants will not burden the Court with detailed discussion of every error, misstatement, or omission in Mann's factual recitation—these factual disputes are not directly relevant to the issues before the Court at this stage—three points do require response. First, it is simply not the case that Mann's work has been subject to no serious and legitimate criticism in the peer-reviewed literature and elsewhere. Second, Mann's claim that any of the investigations into the Climategate scandal delved into the science or truly "exonerated" him are false. And third, Mann's focus on the proceedings regarding EPA's denial of petitions to reconsider its Greenhouse Gas Endangerment Finding is a red herring, because neither EPA nor the D.C. Circuit issued decisions on any issue even arguably implicated by this litigation. Given this background, the CEI

Defendants' statements are a supportable interpretation of the facts, fair comment on Mann's controversial research methodologies, and therefore a protected expression of opinion.

A. Mann's Work Has Been Met With Controversy and Criticism

Mann's recitation of the factual background confirms one of the CEI Defendants' central points: that Mann's research, while central to the case for man-made global warming, is controversial and has been the target of much criticism over the years. As he concedes, his "hockey stick" research, from the time of its initial publication, has been subject to vigorous debate in scientific, policy, and political circles. Opp. at 11-16. He also concedes that the disclosure of the Climategate emails intensified this debate, with numerous policymakers raising questions about reconstructions of the global temperature record and numerous institutions pressured into conducting investigations of conduct within the field. Opp. at 16-28. Mann may believe that these debates and concerns over his and others' research methodologies are unfounded or counterproductive, but the fact that these things took place demonstrates that many others disagree.

And they had reason to do so. For example, in response to the CEI Defendants' discussion of McIntyre and McKitrick's criticisms of Mann's work, Mann states that "*every* peer-reviewed study that has examined McIntyre and McKitrick's claims has found them to be inaccurate." Opp. at 15 (emphasis added). That is false, as even a single example demonstrates: In a 2011 paper published in the *Annals of Applied Statistics* (a peer-reviewed journal), Blakely McShane (Northwestern University) and Abraham Wyner (University of Pennsylvania) confirmed McIntyre and McKitrick's claims that Mann's statistical methods assume the hockey-stick result and that his temperature proxy data perform worse at temperature estimation than "fake" data run through similar methodologies. Their conclusion: "the long flat handle of the hockey stick is best

understood to be a feature of regression and less a reflection of our knowledge of the truth."[10] "Climate scientists," they say, "have greatly underestimated the uncertainty of proxy-based reconstructions and hence have been overconfident in their models."[11]

And that is far from the only scholarly criticism of Mann's statistical methodology. Mann cites a 2006 report by the National Academies of Science's National Research Council as confirming his work, while omitting its criticisms that "[l]ess confidence can be placed in large-scale surface temperature reconstructions for the period from A.D. 900 to 1600" and that "[v]ery little confidence can be assigned to statements concerning the hemispheric mean or global mean surface temperature prior to about A.D. 900 . . . ."[12] And while Mann attempts to cast doubt on Edward Wegman's critical report to Congress on Mann's statistical methodology, he does not challenge its conclusions (some of which he conceded in congressional testimony). *See* Opp. at 16. Indeed, McShane and Wyner echo Wegman's criticisms of the poor use of advanced statistical methods in climate science and repeat Wegman's lament "that there are very few mainstream statisticians working on climate reconstructions."[13] In fact, they identify Wegman's work as the only published "collaboration with university-level, professional statisticians" on temperature reconstructions prior to their own.[14]

---

[10] Blakeley B. McShane and Abraham J. Wyner, *A Statistical Analysis of Multiple Temperature Proxies: Are Reconstructions of Surface Temperatures Over the Last 1000 Years Reliable?*, 5 Annals of Applied Statistics, no. 1, 2011, Ex. B, at 39.
[11] *Id.* at 40.
[12] National Research Council of The National Academies, *Surface Temperature Reconstructions for the Last 2,000 Years* 3 (2006), http://www.uoguelph.ca/~rmckitri/research/NRCreport.pdf.
[13] Blakeley B. McShane and Abraham J. Wyner, Ex. B, at 6.
[14] *Id.* at 39.

B. Not One of the Reports Heeded CEI's Call for an Inquiry into the Science

And that is why CEI and other critics of the "consensus" position on global warming had hoped that the revelations of the Climategate scandal would lead to a critical review of the methodology of the science, particularly the temperature reconstructions underlying the hockey-stick diagram. As Mann concedes, CEI and others repeatedly called for such an investigation, but he points to absolutely nothing to support his assertion that these "calls were heeded." Opp. at 19. Instead, as Mann himself recounts, these investigations focused on such things as whether the scientists who were implicated had made up data. Opp. at 20-28. None, however, addressed concerns that the complicated statistical models contrived by Mann and others were biased or that their output (e.g., the hockey stick figure) had been oversold—which was the Blog Post's entire point.

To be clear, that criticism is true of *every single one* of the reports cited by Mann. The University of East Anglia's Scientific Assessment Panel ("SAP") conceded that "[t]he potential for misleading results arising from selection bias is very great in this area,"[15] but specifically declined to investigate that issue, leading to a rebuke by Members of the British Parliament.[16] The Independent Climate Change E-mails Review ("ICCER") specifically declined to make any "statement regarding the correctness of any of these analyses in representing global temperature trends" or to "address any possible deficiencies of the method" employed by University of East Anglia researchers and

---

[15] University of East Anglia, *Report of the International Panel Set Up by the University of East Anglia to Examine the Research of the Climate Research Unit* (2010), Ex. 24, at 3.
[16] James Randerson, *Oxburgh: UEA vice-chancellor was wrong to tell MPs he would investigate climate research*, The Guardian (Sept. 8, 2010), Ex. 26.

Mann.[17] The House of Commons Science and Technology Committee lamented that it lacked the time to look into the science: "If there had been more time available before the end of this Parliament we would have preferred to carry out a wider inquiry into the science of global warming itself."[18] It specifically stated, "this was not an inquiry into global warming."[19] Similarly, the British Government response to the House of Commons report stated that "[i]t was not our purpose to examine, nor did we seek evidence on, the science produced by CRU."[20] Penn State's investigation disclaimed any intention of wading into a "bona fide scientific disagreement or debate."[21] Rather than defend the science of long-term temperature reconstructions, EPA's decision to deny reconsideration of its Greenhouse Gas Endangerment Finding (discussed further below) disclaimed any substantial reliance on that research.[22] The Department of Commerce's inquiry was limited to investigating any misconduct by National Oceanic and Atmospheric Administration ("NOAA") personnel and "did not assess the validity or reliability of NOAA's or any other entity's climate science work."[23]

---

[17] Sir Muir Russel, et al., *The Independent Climate Change E-mails Review* (July 2010), Ex. 18, at 49.
[18] House of Commons Science and Technology Committee, *The disclosure of climate data from the Climatic Research Unit at the University of East Anglia* (2010), P. Ex. 7, at 9.
[19] *Id.*
[20] Secretary of State for Energy and Climate Change, *Government Response to the House of Commons Science and Technology Committee 8th Report of Session 2009-10: The disclosure of climate data from the Climate Research Unit at the University of East Anglia* (2010), P. Ex. 8, at ¶31.
[21] Inquiry Committee for the Case of Dr. Michael E. Mann, *RA-10 Inquiry Report: Concerning the Allegations of Research Misconduct Against Dr. Michael E. Mann, Department of Meteorology, College of Earth and Mineral Sciences, the Pennsylvania State University* (Feb. 3, 2010), Ex. 22, at 2.
[22] EPA's Denial of the Petitions to Reconsider the Endangerment and Cause or Contribute Findings for Greenhouse Gases Under Section 202(a) of the Clean Air Act, 75 Fed. Reg. 49,556, 49,571/3 (Aug. 13, 2010), P. Ex. 11.
[23] Letter from Todd. J. Zinser to Senator James M. Inhofe (Feb. 18, 2011), P. Ex. 12, at 2.

Finally, the National Science Foundation ("NSF") inquiry, which was undertaken only because the agency found Penn State's inquiry to be insufficient, actually acknowledged "concerns . . . about the quality of the statistical analysis techniques that were used in [Mann's] research" and "concern about how extensively [Mann's] research had influenced debate in the overall research field."[24] But it declined to conduct any analysis of Mann's work, on the basis that it was irrelevant to the object of its investigation: the existence of "research misconduct" as defined in its regulations as "plagiarism, fabrication, and falsification."[25]

In sum, far from CEI's calls for a scientific inquiry being heeded, *not one* of the "investigations" actually *investigated* the methodology of the science.

C. Not One of the Reports "Exonerated" Mann

And as for Mann's supposed "exoneration," only two of the investigations—those conducted by Penn State and NSF—actually focused on Mann's conduct. Penn State's investigation relied almost entirely on evidence provided by Mann[26] and declined to even speak with any experts critical of Mann's work.[27] While NSF did additionally speak with several critics in its inquiry into possible "data fabrication or falsification"—an accusation not raised in the Blog Post—it did not conduct a full investigation of Mann's data practices (e.g., a forensic investigation or an attempt to recreate Mann's datasets) because it determined that "no direct evidence has been presented that

---

[24] National Science Foundation Office of Inspector General Office of Investigations, *Closeout Memorandum Case Number: A09120086*, Ex. 23, at 3.
[25] *Id.* at 2-3.
[26] Office of the Vice President for Research at Penn State, *Investigation of climate scientist at Penn State complete* (June 4, 2010) (listing evidentiary sources), Ex. 6, att. G, at 6.
[27] National Science Foundation Office of Inspector General Office of Investigations, Ex. 23, at 2.

indicates the Subject fabricated the raw data he used for his research or falsified his results."[28] Mann was not "exonerated" following an investigation into the facts; rather, it would be more accurate to say that the inquiries into his conduct were dropped at a preliminary stage.

In fact, the very investigations that Mann cites as "exoneration," Opp. at 28, actually raise questions concerning his research and conduct. As the NSF report explained, the "publicly released emails . . . contained language that reasonably caused individuals, not party to the communications, to suspect some impropriety on the part of the authors," including Mann.[29] That same report, as described above, raised "concerns" regarding Mann's statistical methods and influence on the field. The ICCER report recognized that there are "multiple sources of uncertainty in respect of proxy temperature reconstructions," such as those by Mann, and that these "are the subject of an ongoing and open scientific debate" as to their correctness.[30] Similarly, the SAP report actually identified the potential for bias in the statistical models used for long-term temperature reconstructions and specifically found that some research groups engaged in paleoclimate reconstruction had employed "inappropriate statistical tools with the potential for producing misleading results."[31] The bodies that issued these reports apparently disagree with Mann's bluster that "there is simply no legitimate support for any different conclusion" on the issues raised by the Climategate emails. Opp. at 18.

D. EPA's Reconsideration Proceeding Addressed No Issue Relevant to This Litigation

Finally, Mann makes much of EPA's reconsideration decision. Opp. at 22-25. To begin with, the CEI Defendants are honestly puzzled by Mann's strange accusation that their choice not to

---

[28] *Id.* at 3.
[29] *Id.* at 2-3.
[30] Sir Muir Russel, et al., Ex. 18, at 57.
[31] University of East Anglia, Ex. 24, at 2-3.

discuss an irrelevant administrative proceeding (more on that below) could be construed as "a deliberate attempt to hide information from this Court," Opp. at 24, when that proceeding was prominently disclosed in Mann's Complaint. *See* Compl. ¶22 (stating that CEI filed a petition for reconsideration of EPA's Endangerment Finding).

In any case, there is nothing for the CEI Defendants to hide. On December 15, 2009, EPA issued a finding that "six greenhouse gases taken in combination . . . contribute to the greenhouse gas air pollution that endangers public health and welfare under [Clean Air Act] section 202(a)."[32] After the period for public comment on this Endangerment Finding had closed, the Climategate scandal struck, raising questions regarding some of the scientific research underlying EPA's decision. CEI, joining with two other nonprofit public policy groups, filed a petition for reconsideration of the Endangerment Finding on February 12, 2010, arguing that it was based on "scientifically flawed studies," among them Mann's "hockey stick" research.[33] On August 13, 2010, EPA denied all ten of the petitions for reconsideration that had been filed.[34]

Rather than embrace and defend Mann's research, EPA instead denied that it had relied on it in deciding to issue the Endangerment Finding:

> Petitioners argue that if the current warming is not "unprecedented," our ability to attribute the current warming to greenhouse gases is undermined, and that EPA has not provided "compelling" evidence that the current temperatures are unusual compared to the last 1,000 years. Petitioners misstate EPA's conclusions and overstate the role of this line of evidence. EPA has not claimed that current

---

[32] Endangerment and Cause or Contribute Findings for Greenhouse Gases Under Section 202(a) of the Clean Air Act, 74 Fed. Reg. 66,496, 66,496/1 (Dec. 15, 2009).

[33] *In re Endangerment and Cause or Contribute Findings for Greenhouse Gases Under Section 202(a) of the Clean Air Act*, Petition for Reconsideration of the Nongovernmental International Panel on Climate Change, the Science and Environmental Policy Project, and the Competitive Enterprise Institute (Feb. 12, 2010), P. Ex. 26, at 1.

[34] EPA's Denial of the Petitions to Reconsider, 75 Fed. Reg. at 49,557/1, P. Ex. 11.

> warming is "unprecedented"; the Administrator's Endangerment Finding stated that "The second line of evidence arises from indirect, historical estimates of past climate changes that suggest that the changes in global surface temperature over the last several decades are unusual." EPA found the scientific evidence "supports" this conclusion, not that it compels it, as petitioners incorrectly assert. EPA clearly characterized the uncertainty in this line of the evidence, properly stating that there is significant uncertainty in the temperature record prior to 1600 A.D.[35]

EPA's notice mentions Mann once, in a footnote citation to a 2009 paper.[36]

As to Climategate, EPA does not claim to have conducted any independent investigation, but states only that it "has reviewed all of the CRU emails."[37] According to EPA, "[t]he core defect in petitioners' arguments [regarding Climategate] is that these arguments are not based on consideration of the body of scientific evidence" that the agency says underlies the Endangerment Finding.[38] For that reason, the agency decided that arguments based on the Climategate emails did not require it to reconsider the Endangerment Finding.[39]

EPA also published a "Myths vs. Facts" document on its website, quoted at length by Mann, concerning the denial of the petitions for reconsideration—essentially, a press release. The agency again stated that its investigation consisted of "carefully review[ing] the CRU emails" and that its findings on global warming were based on "multiple lines of evidence" besides those implicated by the Climategate scandal.[40]

---

[35] *Id.* at 49,571/3 (citations omitted).
[36] *Id.* at 49,571/3 n.25.
[37] *Id.* at 49,581/1.
[38] *Id.* at 49,557/3. Although Mann omits it, the quoted sentence is actually the topic sentence of the first paragraph quoted by Mann. *See* Opp. at 23.
[39] *Id.* at 49,557/3.
[40] EPA, *Myths v. Facts: Denial of Petitions for Reconsideration of the Endangerment and Cause or Contribute Findings for Greenhouse Gases under Section 202(a) of the Clean Air Act*, P. Ex. 25.

On June 26, 2012, the D.C. Circuit upheld EPA's decision to deny the petitions for reconsideration.[41] The agency, it explained, was due "an extreme degree of deference" on scientific questions.[42] The court reasoned that EPA had not acted in an arbitrary and capricious fashion because it had relied on the IPCC assessment, which in turn "relied on around 18,000 studies," such that any "inaccurate information" in the assessment "does not appear sufficient to undermine the substantial overall evidentiary support for the Endangerment Finding."[43] The decision does not mention Mann, the "hockey stick" diagram, proxy-based temperature reconstruction, paleoclimate, or really any subject that relates to this litigation.[44]

In sum, when EPA was faced with criticisms of Mann's research, EPA denied that it acted on the basis of that research, rather than defend it, and maintained that it had always recognized doubts about the reliability of paleoclimate reconstructions such as Mann's. The D.C. Circuit, in turn, upheld EPA's decision to deny reconsideration as not arbitrary and capricious because EPA had relied on so many studies other than those implicated by Climategate. As should be apparent, these events are, at most, of indirect relevance to the instant case.

---

[41] *Coalition for Responsible Regulation v. EPA*, 401 U.S. App. D.C. 306, 684 F.3d 102 (D.C. Cir. 2012).
[42] *Id.* at 129.
[43] *Id.* at 125.
[44] The nonbinding statement of issues cited by Mann, Opp. at 2 & n.4 (citing P. Ex. 36), is actually from CEI's challenge to the Endangerment Finding, not its challenge to the EPA's denial of its petition to reconsider the Endangerment Finding. Because the Climategate emails were disclosed months after the close of comments on the Endangerment Finding, arguments regarding Climategate were procedurally barred from that challenge. *See, e.g., National Wildlife Federation v. EPA*, 286 F.3d 554, 562 (D.C. Cir. 2002) ("It is well established that issues not raised in comments before the agency are waived and this Court will not consider them."). In its issue statement in the reconsideration challenge, CEI did not ask the D.C. Circuit to adjudicate any matters at issue in this case but (in relevant question), "[w]hether EPA's treatment of the 'Climategate' documents and of other evidence which developed or came to light after its Endangerment decision is arbitrary, capricious, or otherwise contrary to law." *Competitive Enter. Inst. v. E.P.A.*, Nonbinding Statement of Issues 2 (Nov. 17, 2010), Ex. C.

II. Mann's Claims Fail Under the D.C. Anti-SLAPP Act Because Mann Cannot Show that They Are "Likely To Succeed on the Merits"

In response to the Defendants' Anti-SLAPP Motions, Mann has abandoned his complaints against several of the statements that he had previously alleged to be defamatory, presumably recognizing that those claims would fail as a matter of law. *Compare* Compl. ¶¶26, 28, 32 *with* Opp. at 41. Those claims that remain, however, are equally flawed, because the CEI Defendants' statements are expressions of protected opinion under the First Amendment and D.C. common law. Mann's claims should therefore be dismissed, with prejudice, at this stage of the litigation.

A. Mann Concedes that the D.C. Anti-SLAPP Act Applies to His Claims

"Dr. Mann does not dispute that the Anti-SLAPP statute applies here . . . ." Opp. at 37. That is the end of the matter.

Nonetheless, Mann insists that his lawsuit is "distinguishable from the type of action the District of Columbia had in mind when it enacted [the Act]." Opp. at 34. That claim is belied, first and foremost, by the Act's text. The Anti-SLAPP Act applies to "[a]ny written or oral statement made . . . (ii) In a place open to the public or a public forum in connection with an issue of public interest . . . ." D.C. Code § 16-5501(1)(A). "Issue of public interest" is, in turn, defined as "an issue related to health or safety; environmental, economic, or community well-being; the District government; a public figure; or a good, product, or service in the market place." D.C. Code § 16-5501(3). Both the Blog Post and CEI's press release on Mann's legal threats were written statements accessible to the public, and they both concerned issues of environmental and economic well-being

related to a public figure, Mann.[45] Accordingly, the CEI Defendants' Anti-SLAPP Motion "shall be granted unless the responding party [i.e., Mann] demonstrates that the claim is likely to succeed on the merits . . . ." D.C. Code § 16-5502(b). Whether or not Mann's complaint was "well-pled," *see* Opp. at 34, is irrelevant under the text of the Act. *See* D.C. Code § 16-5502(b). If that were the standard, the core provision of the Act would be superfluous, because a complaint that is not well-pled is already subject to dismissal under Rule 12(b)(6).

Mann's argument that the Act should not apply to suits by a "lone individual" or should apply only to suits against individuals is contrary to the Act's text and purpose. To begin with, Mann overlooks that he has sued two individuals, Rand Simberg and Mark Steyn, seeking to hold them jointly and severally liable for money damages and costs "to the highest extent permitted by law." Compl. ¶¶11-12, 101. In any case, this interpretation of the Act, based on an unpublished federal court order's holding that predates the D.C. Act by nearly a decade and has not been followed by any court, Opp. at 35, would deny application of the Act to media and advocacy organizations, which are among the most likely to face SLAPP suits and require the Act's protection. *Cf. Klayman v. Segal*, 783 A.2d 607, 613 n.5 (D.C. 2001) ("Because prolonged litigation in defamation actions against media defendants may inhibit free speech . . . . summary procedures are essential."). Indeed, Mann ignores that the first dismissal under the Act was of a lawsuit by an individual against a corporation and its employee. Order, *Lehan v. Fox Television Stations, Inc.*, No. 2011 CA 004592 (D.C. Sup. Ct. Dec. 2, 2011) (King, J.); *see also Farah v. Esquire Magazine, Inc.*, 863 F. Supp. 2d 29, 36-

[45] *See* Mann, *The Hockey Stick and the Climate Wars*, at 253 ("I became a public figure . . . ."); Bill Blakemore, *'New McCarthyism' Described by Climate Scientist Michael Mann*, abcnews.com (July 8, 2012), Ex. 2, at 3 (Mann identifies himself as "a public figure in this debate" on global warming); Bill Blakemore, *Climate Denialists Worse than Tobacco CEOs Lying Under Oath, Says Mann*, abcnews.com (July 8, 2012) (Mann states that he has "actually learned to embrace the role"), Ex. 2, at 10.

40 (D.D.C. 2012) (dismissing defamation case against magazine and its employees under D.C. Anti-SLAPP Act). He also ignores the D.C. Council's view that the impact of SLAPP suits "is not limited to named defendants['] willingness to speak out, but prevents others from voicing concerns as well." Council of the District of Columbia Committee on Public Safety and the Judiciary Committee Report, Report on Bill 18-893, "Anti-SLAPP Act of 2010" ("Report on Bill 18-893"), Ex. 8, at 1 (describing the Act's "Background and Need"). In the D.C. Council's view, stopping lawsuits "aimed to punish or prevent the expression of opposing points of view," *id.*, at the earliest possible instant is a public good, the benefit of which accrues to all persons subject to the jurisdiction of this Court. Given the unusual (and unrebutted) fact that Mann has confirmed that his aim in this litigation is to silence opposing points of view on global warming, CEI Defendants' Anti-SLAPP Motion at 1-3, 23-26, it would be difficult to imagine a more appropriate case for dismissal under the Act.

Mann's contention that the CEI Defendants do not yet "show any signs of having their First Amendment rights 'muzzled'" also misses the mark. Opp. at 36. It is well recognized that defendants such as CEI and Simberg are fully entitled to seek the protection of anti-SLAPP statutes notwithstanding their ability to continue to exercise their First Amendment rights. *Lafayette Morehouse, Inc. v. Chronicle Publ'g Co.*, 37 Cal. App. 4th 855, 863 (1995) (plain language of statute applied to libel cases brought against a media outlet for reporting on issues of public concern); *Stern v. Doe*, 806 So. 2d 98, 100 (La. Ct. App. 2001) ("The purpose of [the Louisiana anti-SLAPP statute] is to review frivolous and meritless claims against the media at a very early stage in the legal proceedings."). To hold that the Act applies only when SLAPP defendants have in fact been silenced would defeat its purpose of blocking attempts to litigate parties into silence.

Finally, Mann understandably resists the daunting burden that the Act's language imposes on him. As Mann describes, the Act was modeled on California's anti-SLAPP statute, but instead of

requiring (as under California law) a "probability the plaintiff will succeed on the merits," it requires the plaintiff to show that he is "likely to succeed on the merits." Opp. at 37-38. Mann's assertion that "[t]his is a distinction without a difference," Opp. at 38, denies all credit to the D.C. Council's choice to depart from California's approach in this one, crucial respect. The legislative history shows that the D.C. Council studied the laws of the states and federal law in crafting the Act's text, and then substantially revised that text (including the provision at issue) based on comments by the American Civil Liberties Union ("ACLU") to better address what it recognized to be a substantial problem within the District. Report on Bill 18-893, Ex. 8, at 1, 8.

To that end, it chose a word, "likely," that the same dictionary relied upon by Mann defines as "having a high probability of occurring or being true" and "very probable."[46] Mann offers no explanation for why he quotes, instead, the definition of "likelihood," a different word that is not used in the statute. Opp. at 38. Based on the words the Council chose, it is quite clear that the Council did intend, contrary to Mann's contention, that a court applying the Act "determine whether it is more probable than not that plaintiff will prevail on the claim." Opp. at 38.

B. Mann Identifies No Provably False Statements of Fact Because the CEI Defendants' Characterization of His Research Is Protected Opinion

Mann focuses on the verifiability of statements as the lynchpin of the opinion analysis, Opp. at 43-46, but fails to acknowledge that courts are also required to consider the context in which the statements were published when determining whether they are actionable to begin with. *See Moldea II*, 306 U.S. App. D.C. at 5, 22 F.3d at 314. If the context is one in which a reader expects to be presented with statements of opinion, defendants "must be given some leeway to offer 'rational

---

[46] *Likely*, Merriam-Webster.com, http://www.merriam-webster.com/dictionary/likely.

interpretation"' of the facts. *Id.* at 4, 22 F.3d at 313. In such cases, the "correct measure" of whether a statement is verifiably false is whether "*no reasonable person could find* that the [defendant's] characterizations were supportable interpretations" of true underlying facts disclosed to the reader. *Id.* at 8, 22 F.3d at 317 (emphasis in original). Any statements that fail to satisfy this stringent test are protected opinion. Mann's contention that this analysis applies only to "evaluations of a literary work" has been definitively rejected by, among others, the Court of Appeals, which has also rejected his argument that generalized claims can be subject to the verification necessary to establish a provable falsehood.

1. Mann Ignores the Central Role of Context

Mann is compelled to concede that "context has been a determinative factor for courts in the wake of *Milkovich*," but he errs in his contention that context matters only in "the necessarily subjective theater of artistic commentary and review." Opp. at 47. Contrary to Mann's contention, *Milkovich v. Lorain Journal Co.*, 497 U.S. 1, 19 (1990), while rejecting "an artificial dichotomy between 'opinion' and fact," did not undermine the central role of context in determining whether a statement is actionable. *See* Opp. at 46-47. The case law is clear that context matters in three respects: genre, subject matter, and the work as a whole. Any of these may be decisive.

Indeed, the *Milkovich* court described with approval the Court's earlier decision in *Greenbelt Cooperatuve Publishing Ass'n, Inc. v. Bresler*, 398 U.S. 6 (1970), which recognized the relevance of context in holding that a characterization of a developer's negotiating position as "blackmail" could not support a defamation claim. 497 U.S. at 16-17. The Court recognized that, "as a matter of constitutional law, the word 'blackmail' *in these circumstances* was not slander when spoken," *id.* (internal quotation marks omitted) (emphasis added), due to the context of the heated public debate over the developer's tactics and the statement's placement in an article that provided greater factual context. *Id.* at 17. The *Milkovich* court also cited with approval *Letter Carriers v. Austin*, 418 U.S. 264,

284-86 (1974), a case holding that use of the word "traitor" to describe a union "scab" was not actionable "in the context of this case," which was an article regarding a heated labor dispute published in a pro-union newsletter, because readers would have understood that word "to demonstrate the union's strong disagreement with the views of those workers who oppose unionization." 418 U.S. at 284. The Court recognized that "such exaggerated rhetoric was commonplace in labor disputes" and so was not actionable in that context. *Id.* at 286.

And were there any doubt on the continued relevance of context, the Supreme Court laid it to rest in its recent decision in *Snyder v. Phelps*, 131 S. Ct. 1207, 1219 (2011), which held a church's offensive message on picket signs at a soldier's funeral to be protected speech based on "the whole context of how and where it chose to say it."[47]

Moldea II confirmed that *Milkovich* was not intended, and should not be read, to "sweep away" the established "principle of looking to the context in which speech appears" that was at the heart of *Ollman v. Evans*, 242 U.S. App. D.C. 301, 750 F.2d 970 (D.C. Cir. 1984) (en banc), and other cases. *See Moldea II*, 306 U.S. App. D.C. 1, 5-6, 22 F.2d at 314-15 (discussing *Ollman*). In *Moldea v. New York Times*, 304 U.S. App. D.C. 406, 414-18, 15 F.3d 1137, 1145-49 (D.C. Cir. 1994) ("*Moldea I*"), the D.C. Circuit initially found actionable two passages supporting the newspaper's assessment that Moldea was a "sloppy" journalist: the questioning of his assertion that Joe Namath "guaranteed" a Super Bowl victory "shortly after a sinister meeting in a bar with a member of the opposition" and the criticism of Moldea for the "reviv[al] of the discredited notion" that an owner of the L.A. Rams "who had a penchant for gambling[] met foul play when he drowned in Florida."

---

[47] The Supreme Court's 2010 grant of certiorari in *Snyder*, 130 S. Ct. 1737 (2010), vacated the Fourth Circuit decision cited by Mann. *See* Opp. at 47 & n.87.

The court further thought it "important to make clear that . . . our analysis of this case is not altered by the fact that the challenged statements appeared in a 'book review' rather than in a hard news story." *Id.* at 414-15, 15 F.3d at 1145-46.

But on rehearing, the D.C. Circuit reversed its earlier holding on these two passages for "fail[ing] to take sufficient account" of the context in which the statements appeared. *Moldea II*, 306 U.S. App. D.C. at 2, 22 F.3d at 311. It recognized that *Milkovich* was decided "against the backdrop of th[e] settled principle" that different genres of writing have a different influences on the average reader, and that it had "erred in assuming that *Milkovich* abandoned the principle of looking to the context in which [the statement] appears." *Id.* at 5-6, 22 F.3d at 314-15. Instead of "disavow[ing] the importance of context," the Supreme Court "simply discounted it in the circumstances of that case." *Id.* at 5, 22 F.3d at 314 (internal quotation marks omitted). Applying the correct standard, the D.C. Circuit dismissed the case in its entirety.

And contrary to Mann's characterization of the case, Opp. at 47-48, *Weyrich v. New Republic, Inc.*, 344 U.S. App. D.C. 245, 235 F.3d 617 (D.C. Cir. 2001), similarly accepted the importance of context. Recognizing that *The New Republic* "is itself well-known to be a magazine of political commentary, a self-described 'Weekly Journal of Opinion,'" it held that a reference to the plaintiff's supposed "bouts . . . of paranoia" was not actionable because it was "[p]resented in such a loose manner, in such a well-understood context . . . ." As the court explained, although if "looking at these statements in isolation, a reasonable reader *might* interpret them to attribute a diagnosable and debilitating mental affliction to appellant . . . the First Amendment demands that we place these references in their proper context." *Id.* at 253, 235 F.3d at 625. Also contrary to Mann's description, the court did not hold that certain isolated words or statements regarding the plaintiff were actionable, *see* Opp. at 48, but allowed to proceed claims regarding "a number of [allegedly] false anecdotes, suggesting to the average reader that appellant is not only a political reactionary, but

emotionally volatile, perhaps even mentally unsound, and otherwise unfit for his profession." 344 U.S. App. D.C. at 255, 325 F.3d at 627. Unlike the "loose" characterizations of Weyrich the court held to be protected—references to "bouts of pessimism and paranoia," "habits of suspicion, pessimism, and antagonism," and the fact that other conservatives have acted "as nutty as Weyrich"—these anecdotes were "historical vignettes," some "utiliz[ing] quotations, some purportedly from appellant, to further reinforce the impression that the stories are in fact true." *Id.* at 254, 325 F.3d at 626. The difference was that these anecdotes were sufficiently detailed and specific to overcome any presumption that, based on their appearance in a political magazine, they were expressions of opinion. *Id.*

And Mann ignores entirely that the D.C. Court of Appeals has been especially sensitive to context. The Court of Appeals recognized in *Guilford Transp. Indus., Inc. v. Wilner*, 760 A.2d 580, 582-83 (2000), that the context-sensitive approach of *Ollman* remains good law and that it applied to an op-ed column concerning a labor dispute. The plaintiffs, including a railroad company, claimed that the column falsely portrayed them as antagonistic to labor and implicitly accused them of violating federal labor statutes. *Id.* at 585. The court, however, found the challenged statements non-actionable, based on three contextual factors. First, it was "critical . . . that the allegedly defamatory utterances in this case appeared in an Op–Ed column in which Wilner [the defendant] was commenting on matters of substantial public concern." *Id.* at 597.

Second, the statements were "made in the context of a labor dispute," such that statements "which on their face resemble statements of fact, may, depending on the circumstances, be treated as statements of opinion not subject to an action for libel" because such disputes "normally involve considerable differences of opinion and vehement adherence to one side or the other." *Id.* at 597-98 (citation and internal quotation marks omitted). In that context, the Court held, even some "provably false" statements could not support a defamation claim:

> [W]e do not believe that those statements in the column which the plaintiffs have characterized as "provably false" are of the genre which would support a defamation case against the author of a column on the opinion page of a newspaper. The plaintiffs' focus has been on Wilner's allegation that Guilford [the main plaintiff] "bolted" from national wage and benefit negotiations. According to Professor Northrup, Guilford did not bolt; rather, it declined to "opt into" national "handling." Either way, Guilford negotiated locally and not nationally. Even assuming that Wilner's use of the verb "bolted" reflects lack of precision, and treats the plaintiffs with undeserved asperity, the challenged language surely pales in comparison to "blackmail," *Greenbelt Publ'g Ass'n*, 398 U.S. at 11-14, or "traitor" or "scab," *Austin*, 418 U.S. at 282-87. If we were to adopt a rule of law which sustains the plaintiffs' position on this issue, then authors of every sort would be forced to provide only dry, colorless descriptions of facts, bereft of analysis or insight.

Id. at 598-99 (some citations and internal quotation marks omitted).

And third, the Court considered the work as a whole, particularly the column's acknowledgment that one deal that it criticized had been upheld by a federal agency. *Id.* at 599; *see also Klayman*, 783 A.2d at 616 ("[T]he publication must be considered as a whole . . . ."). Thus, "[a]ny reasonable reader of the column would understand that Guilford took certain actions, that Wilner was apparently unenthusiastic about those actions, and that the ICC [Interstate Commerce Commission] basically sustained them." 760 A.2d at 599. "This is not the stuff of which successful libel suits are made," the Court concluded. *Id.*

2. Taken in Context, the Challenged Statements Are Protected Opinion

Although acknowledging that context may be "determinative," Mann makes no attempt to apply the law to the facts of this case other than to state that it is not a "liability shield" in every instance. *See* Opp. at 47. But all three contextual factors of *Guilford*, which is binding on this Court, demonstrate that the challenged statements are protected expressions of opinion.

First is genre. Mann concedes that the Blog Post and Lowry's statement (which Mann contends CEI republished) "were published on websites that . . . often offer opinion commentary." Opp. at 47. *Moldea II*, *Weyrich*, and *Guilford* each recognize the significance of this contextual factor, with *Guilford* declaring it to be "critical." 760 A.2d at 597. Indeed, the Blog Post employs the kind

of "'strong statements, sometimes phrased in a polemical manner that would hardly be considered balanced or fair elsewhere in the newspaper,'" that would tip off a reasonable reader that its contents are not "hard news" but expressions of non-verifiable opinion. 760 A.2d at 583 (quoting *Ollman*, 242 U.S. App. D.C. at 317, 750 F.2d at 986).

Second is the broader context of the public debate over global warming. Just as the Court of Appeals recognized regarding labor disputes, statements made in the context of the debate over global warming "normally involve considerable differences of opinion and vehement adherence to one side or the other." 760 A.2d at 598. In fact, Mann's recent book fairly well chronicles the heated debate over global warming, often describing it as a "war" or a "battle."[48] In this context, forceful, highly opinionated language and hyperbole are expected, from advocates on both sides. Mann uses precisely this type of language when he describes CEI as dishonest, accuses it of being an "industry front group," and characterizes its work as "fraudulent." *See* CEI Defendants' Anti-SLAPP Motion at 23-26. In such a heated context, the Court of Appeals held, statements "which on their face resemble statements of fact, may, depending on the circumstances, be treated as statements of opinion not subject to an action for libel." 760 A.2d at 597. Were the law otherwise, it would sweep up too much speech on matters of public interest, stifling free and open debate. *Id.*

The statements at issue here are not distinguishable from those at issue in *Guilford* in terms of their vehemence or general implication of disapproval. The article in *Guilford* stated that the plaintiffs had "ignited" a "bitter labor-management conflict," "bolted from traditional national wage and benefits negotiations," engaged in "chaotic legal fisticuffs," and employed questionable legal tactics later blessed by a federal agency "with a zealous pro-management bias." *Id.* at 584-85. Here,

---

[48] *See, e.g.*, Mann, *The Hockey Stick and the Climate Wars*, at 233.

Mann complains of statements that his research is "intellectually bogus" and relies on "data manipulation" to reach certain conclusions, that he is the "posterboy" of a "corrupt and disgraced" field, and that claims against him have been dismissed in investigations by a public university whose willingness to pursue "academic and scientific misconduct" the CEI Defendants (and many others) doubt. Opp. at 41; Compl., Exs. A, C. There is no relevant distinction, and Mann does not suggest one. *See also infra* § II.B.3 (discussing related hyperbole inquiry).

Third is the question of how those statements are situated in the work as a whole. Just as the column at issue in *Guilford* reported that a federal agency had approved the plaintiff's actions, the Blog Post reports that Mann was "declared innocent of any wrongdoing" by Penn State and was also cleared by the NSF's investigation. The Blog Post does not call for Mann to be fired—a call that would naturally follow an accusation of fraud—but for the university to commission " a fresh, truly independent investigation" of his research. As in *Guilford*, any reasonable reader would understand that Mann took certain actions, that the CEI Defendants were unenthusiastic about those actions, and that Penn State and the NSF found Mann guilty of no wrongdoing. *See* 760 A.2d at 599. That "is not the stuff of which successful libel suits are made." *Id.*

3. Taken in Context, the Challenged Statements Are Rhetorical Hyperbole

Mann ignores the fact that the challenged statements are, in the context of the climate-change debate, clearly rhetorical hyperbole, phrased in colorful language, and not actionable assertions of fact. Indeed, the entirety of Mann's response consists of citations to comments left on CEI's website by unknown third parties and to statements by Mann's supporters and allies

professing *shock* that anyone would compare Penn State's handling of Mann to its handling of Jerry Sandusky.[49] *See* Opp. at 52-55.

Had Mann addressed the case law—this section of his brief discusses none and does not even attempt to apply the law to the facts of the case—he would know these citations are irrelevant:

> [T]he inquiry into whether a statement should be viewed as one of fact or one of opinion must be made from the perspective of *an 'ordinary reader'* of the statement. It is also clear that the determination of whether a statement is opinion or rhetorical hyperbole as opposed to a factual representation *is a question of law for the court.*

Mr. Chow of N.Y. v. Ste. Jour Azur S.A., 759 F.2d 219, 224 (2d Cir. 1985) (citation omitted) (emphasis added). Yet faced with a "question of law," Mann ignores the law entirely, choosing instead to focus on a carefully curated selection of unrepresentative comments and articles that support his interpretation of the challenged statements. If the law actually held that the proper standard is a battle of citations, the CEI Defendants would proffer their own list, and the Court could determine which list is longer. But because the law does not hold that, the CEI Defendants will instead address the governing precedents.

Under those precedents, when the literal or factual nature of a statement is challenged, the Court should play it down the middle. Thus, "[i]n determining whether a statement is fact or opinion, a court is, of course, trying to assess the average reader's view of the statement rather than that of either the most skeptical or most credulous reader." *Ollman*, 242 U.S. App. D.C. at 310 n.16, 750 F.2d at 979 n.16. "The court should not . . . indulge far-fetched interpretations of the challenged publication. The statements at issue should not be interpreted by extremes, but should be construed as the average or common mind would naturally understand them." *Guilford*, 760 A.2d

---

[49] This is despite the fact that Mann drops his defamation claim regarding that statement in the same section of his brief. *See* Opp. at 52; Compl. ¶26 (listing challenged statements).

at 594-95 (internal quotation marks omitted). Once again, context is key. *Dilworth v. Dudley*, 75 F.3d 307, 310 (7th Cir. 1996) (explaining that hyperbole "cannot be determined without consideration of context"); *Letter Carriers*, 418 U.S. at 286-87; *Greenbelt*, 398 U.S. at 14.

Hyperbolic rhetoric is not actionable because it "cannot reasonably [be] interpreted as stating actual facts . . . ." *Weyrich*, 344 U.S. App. D.C. at 252, 235 F.3d at 624 (internal quotation marks omitted). For example, when Mann asserts on a radio broadcast that one of his critics is a "hired assassin," no reasonable listener takes that statement to mean that the critic has, in fact, been hired to murder Mann or anyone else. *See* CEI Defendants' Anti-SLAPP Motion at 55 (analyzing other instances of Mann's hyperbole).

The language at issue here is similarly hyperbolic and would be recognized as such by any ordinary reader, who would be attuned to the heated rhetoric typically employed in the public debate over global warming. For example, Mr. Lowry's statement that Mann's work is "intellectually bogus" would not, to a reasonable reader, mean or imply criminal fraud any more than the statement the term "intellectually bankrupt" implies insolvency. A word like "bogus" is precisely the kind of word used to express outrage and disagreement, as opposed to stating cold, hard facts. On that basis, an Ohio court rejected the claim that use of the word "bogus" to describe legal claims implied fraud, finding instead that the word "suggests opinion." *Cooke v. United Dairy Farmers, Inc.*, No. 04AP-817, 2005 WL 736246, at *6 (Ohio Ct. App. Mar. 31, 2005).

And on that same basis, many courts have held that even use of the word "fraud" or "fraudulent" was, in context, only hyperbole, not an assertion of fact. *E.g., Coghlan v. Beck*, --- N.E.2d ---, 2013 WL 240421, at *11 (Ill. Ct. App. Jan. 22, 2013) (defendants' description of plaintiffs' enterprise as a "fraud machine" not actionable "in the overall context" of publication criticizing plaintiff); *Art of Living Found. v. Does*, No. 10-CV-05022, 2011 WL 2441898, at *7-8 (N.D. Cal. 2011) (defendants' statement that plaintiffs engaged in "fraud" and "obtained money from

participants on false, deceitful declarations" not actionable in the context of "obviously critical blogs . . . with heated discussion and criticism" of defendants and in the "specific context" of a heated Internet debate); *Nicosia v. De Rooy*, 72 F. Supp. 2d 1093, 1104 (N.D. Cal. 1999) (defendant's statement that the plaintiff was a "self-serving fraud," a "criminal" and acted "illegally" not actionable due to the "controversial subject matter of the debate" and use of language "too loose and hyperbolic to be susceptible of being proved true or false"); *Beattie v. Fleet Nat. Bank*, 746 A.2d 717, 727 (R.I. 2000) (defendant's statement that plaintiff's appraisal was so misleading "as to be considered fraudulent" not actionable where context of deal indicated it "amounted to a hyperbolic . . . opinion"); *Phantom Touring, Inc. v. Affiliated Publ'n*, 953 F.2d 724, 729 (1st Cir. 1992) (newspaper's description of a theatrical production as "a rip-off, a fraud, a scandal, a snake-oil job" not actionable due to use of "hyperbolic" language admitting "numerous interpretations" that could not be verified); *600 W. 115th Street Corp. v. Von Gutfeld*, 603 N.E.2d 930, 937 (N.Y. 1992) (defendant's statement that plaintiff's permit application was "as fraudulent as you can get and it smells of bribery and corruption" not actionable in the context of a heated debate among ordinary citizens who could not be supposed to be in possession of undisclosed facts); *Henry v. Halliburton*, 690 S.W.2d 775, 778 (Mo. 1985) (defendant's statement that insurance agent is "'a fraud and a twister'" did not suggest that agent committed specific crime where it was clear from the context that the defendant was expressing only his opinion and not alleging any "specific crime"); *Stuart v. Gambling Times, Inc.*, 534 F. Supp. 170, 172 (D.N.J. 1982) (defendants' description of plaintiffs' book as "the # 1 fraud ever perpetrated upon the gambling reader" not actionable where it was clear from the context that this was only the reviewer's opinion and not a suggestion that any specific "acts occurred which would be criminally punishable").

The other statements of which Mann complains are of a similar vein: strong but "imaginative expression," *Coles v. Washington Free Weekly, Inc.*, 881 F. Supp. 26, 32 (D.D.C. 1995), that

is meant to signal disagreement and disdain, not any accusation of criminal fraud. And it is precisely in that kind of context that courts have routinely held words bearing connotations of both fraud and disdain to be "mere hyperbole rather than falsifiable assertions of discreditable fact," when taken in that kind of context. *Dilworth*, 75 F.3d at 310 ("'scab,' 'traitor,' 'amoral,' 'scam,' 'fake,' 'phony,' 'a snake-oil job,' 'he's dealing with half a deck,' and 'lazy, stupid, crap-shooting, chicken-stealing idiot.'"). Mann offers no reason why the result here should be any different.

4. Mann's Assertion that the CEI Defendants' Statements Are Verifiable Ignores Binding Case Law

Mann asks the Court to pass judgment on the veracity of the exact kind of "general characterizations" that the Court of Appeals has held are not "concrete enough to reveal 'objectively verifiable' falsehoods" that could possibly be the subject of a defamation claim. *Rosen v. AIPAC, Inc.*, 41 A.3d 1250, 1259 (D.C. 2012) (footnote omitted); *see* Opp. at 43-46. In *Rosen*, the Court of Appeals held that an employer's statement to the *New York Times* that an employee had been dismissed for actions regarding the use of classified information that differed from "the standards that AIPAC expects and requires of its employees" was not actionable because the employer lacked any specific written standards. 41 A.2d at 1260. While it expected employees to adhere to such standards as obeying the law and following counsel's advice, that "was too subjective, too amorphous, too susceptible of multiple interpretations . . . to make any of them susceptible to proof of particular, articulable content." *Id.* Because "standards" was "a word of aggregation" at a "high[] level of generality" and "could have meant many things, none self-evident," any statement that the plaintiff had not followed those standards could not be "provably false" and therefore was not actionable. *Id.* at 1260-61.

Rosen, in turn, relied on and approved of two cases demarking the limits of verifiability. *Id.* at 1258-59. *McClure v. Am. Family Mut. Ins. Co.*, 223 F.3d 845, 856 (8th Cir. 2000), concerned press statements made by an employer regarding two insurance agents it had fired for lobbying. The

company had stated that the agents "engaged in 'disloyal and disruptive activity,'" had not understood the "'value of loyalty and keeping promises,'" had acted "'against the best interests of the insurance buying public,'" "'were in direct violation of their agreements,'" and had engaged in "'conduct unacceptable by any business standard.'" *Id.* at 853. As the Court of Appeals noted with approval, the *McClure* court concluded that these "remarks on a subject lending itself to multiple interpretations cannot be the basis of a successful defamation action because as a matter of law no threshold showing of 'falsity' is possible in such circumstances." *Id.* (internal quotation marks omitted). Similarly, in *Gibson v. Boy Scouts of Am.*, 360 F. Supp. 2d 776, 781 (E.D. Va. 2005), the Boy Scouts published a statement that an individual was "unfit to be a Scoutmaster and in Scouts." As the Court of Appeals again noted with approval, the *Gibson* court held that the words were too general to "contain a provably false factual connotation" and so were "merely the expression of the speaker's opinion." *Id.*; *see Rosen*, 41 A.2d at 1259.

So too here. This is apparent from a review of the statements. CEI's statements that Mann "has molested and tortured data in the service of politicized science that could have dire economic consequences for the nation and planet" and "had been engaging in data manipulation" are not, as Mann claims, "plainly factual and verifiable."[50] *See* Opp. at 43. These are general terms subject to multiple meanings, some technical, some critical, many benign. *See* CEI Defendants' Anti-SLAPP Motion, at 41-3. They state nothing that is provably false; in other words, they state an opinion.

Mann's challenge to Lowry's statement that his work is "intellectually bogus" also fails. This statement—which CEI never made, in any case, *see infra* § II.C—is plainly the kind of "general

[50] Although Mann mentions this statement, Opp. at 43, he has abandoned any claim regarding it and no longer considers it to be among those statements he challenges in this litigation, Opp. at 41.

characterization[]" that the Court of Appeals held is not "concrete enough to reveal 'objectively verifiable' falsehoods" that could support a defamation claim. *Rosen*, 41 A.3d at 1259. So unspecific, it would not be viewed by an ordinary reader as making an assertion of fact.

The same is true of the statement that "Mann has become the posterboy of the corrupt and disgraced climate science echo chamber." To begin with, it is not clear to what verifiable facts this statement could refer, because it is a characterization of a field of research and its political supporters (the "corrupt and disgraced *climate science echo chamber*") and of how others view Mann within that field ("the *posterboy*"). Mann asserts that this "statement explicitly accuses Dr. Mann of corruption," as if using the words "Mann" and "corrupt" in the same sentence was itself unlawful, but the statement neither sets forth nor implies any particular facts. Opp. at 43.

Rinaldi v Holt, Rinehart & Winston, Inc., 366 N.E.2d 1299, 1307-08 (N.Y. 1977), actually cuts strongly against Mann, offering a sharp contrast to his hazy claims. *See* Opp. at 43-44. It upheld dismissal of a plaintiff judge's libel action regarding a book that had deemed him "probably corrupt," but initially found the statement actionable because it was "not used merely in a 'loose, figurative sense'" to demonstrate general disagreement. 366 N.E. at 1307. Instead, it was a conclusion based on the book's detailed accounts of "illustrative cases" before the judge—accounts that the judge, in his complaint, alleged to be materially false. *See id.* at 1303. For that reason, "[t]he ordinary and average reader would likely understand the use of these words, in the context of the entire article, as meaning that plaintiff had committed illegal and unethical actions." *Id.* at 1307.[51] So while the plaintiff in *Rinaldi* pointed to factual statements so concrete that no reader could take the

[51] The court nonetheless affirmed dismissal of that claim because the plaintiff was unable to "set forth sufficient evidentiary facts to generate a triable issue of fact as to the falsity and actual maliciousness of the accusations of criminal conduct." 366 N.E.2d at 1307.

charge of corruption merely as an expression of opinion, Mann points only to a word, "corrupt," standing alone and unsupported by any detail—something that any reader would recognize as an epithet, not a factual conclusion.

Finally, Mann never explains how a question directed as criticism at Penn State's investigation of Mann actually states any fact regarding Mann. He merely asserts as much, Opp. at 45, missing the whole point of the question as challenging the university's motivations and diligence. But his interpretation is implausible: why, after all, would one who is asserting that Mann committed fraud call for "a fresh, truly independent investigation" in the very next sentence, rather than simply demand that he be fired? Instead, a reasonable reader would see that the surrounding text is critical of the university, not Mann, and questions Penn State's motives in both the Sandusky and Mann affairs. That reader would take it as restating the central premise of the Blog Post: that Penn State puts its own interests ahead of ferreting out inconvenient truths. Phrased as a question, it leads the reader to that conclusion.[52]

But even assuming Mann's tortured interpretation, "academic and scientific misconduct" is no more concrete than the statements held not actionable in *Rosen* (employee violated his employer's "standards"), *McClure* (employees acted "against the best interests of the insurance buying public" and had engaged in "conduct unacceptable by any business standard"), and *Gibson* (plaintiff was "unfit to be a Scoutmaster and in the Scouts"). Like those statements, it is simply "too subjective,

---

[52] Mann's citation, Opp. at 51, of *Afro-Am. Publ'g Co. v. Jaffe*, 366 F.2d 649, 653-55 (D.C. Cir. 1966), regarding rhetorical questions is inapposite, because that case concerned defamatory meaning—i.e., whether a statement casts its subject into disrepute—and not opinion. *See id.* at 654 (considering whether the statements at issue "tend[ed] to bring the plaintiff into contempt, ridicule and disgrace in the community in which he operated his business").

too amorphous, too susceptible of multiple interpretations," 41 A.3d at 1260, to suggest any verifiable fact. Instead, what it suggests is opinion.

5. Mann Misstates the Law on the "Supportable Interpretation" Standard and "Fair Comment" Privilege

Mann's contention that only statements that are "evaluations of a literary work" may be protected as a "supportable interpretation" of facts, *see* Opp. at 49, has been rejected by, among others, the Court of Appeals. *Guildford*, 760 A.2d at 597 (discussing the application of the "supportable interpretation" standard in case challenging statements in an op-ed column regarding a labor dispute); *id.* at 601 (applying the standard because the article left the reader "free to draw his or her own conclusions regarding whether the plaintiffs acted wrongfully"); *see also Dodds v. Am. Broad. Co.*, 145 F.3d 1053, 1067-68 (9th Cir. 1998) ("Prime Time Live" report on alleged misconduct by judge); *Partington v. Bugliosi*, 56 F.3d 1147, 1156-57 (9th Cir. 1995) (book and television docudrama that impugned attorney's competence and performance in murder trial); *Washington v. Smith*, 317 U.S. App. D.C. 79, 81, 80 F.3d 555, 557 (D.C. Cir 1996) (magazine article impugning competence and performance of basketball coach); *Hunter v. Hartman*, 545 N.W.2d 699, 706 (Minn. Ct. App. 1996) (statement by talk show host impugning competence and performance of sports orthopedist); *Fasi v. Gannett Co., Inc.*, 930 F. Supp. 1403, 1409-10 (D. Haw. 1995) (newspaper editorial that described mayor's actions as "legalized blackmail"). In fact, the CEI Defendants are aware of no case that has held that the "supportable interpretation" standard is limited to book reviews.

And Mann's assertion that *Milkovich*, and not *Moldea II*, "governs here" is nonsensical, because the two cases are not at all inconsistent. *See Moldea II*, 306 U.S. App. D.C. at 5-6, 22 F.3d at 314-15 (discussing *Milkovich*); *Guilford*, 760 A.2d at 597 (discussing *Milkovich* and *Moldea II*). *Moldea II* follows *Milkovich's* holding that "statements of opinion can be actionable if they imply a provably false fact, or rely upon stated facts that are provably false," 306 U.S. App. D.C. at 4, 22 F.3d at 313,

and also recognizes that "a supportable interpretation . . . does not present a verifiable issue of fact that can be actionable in defamation," *id.*

As the Anti-SLAPP Motion describes in detail, the Blog Post links to a wealth of factual materials that provide a basis for its commentary, and each of the challenged statements, in turn, is commentary on those disclosed facts and other facts readily available to the public. *See* CEI Defendants' Anti-SLAPP Motion at 48-51. Rather than address this point, Mann asserts that all of the challenged statements are allegations of fraud and are therefore unsupported, because none of the linked or publicly available materials "sets forth a scintilla of evidence" that would support such an opinion. Opp. at 49. But he does not dispute, nor could he, that the Blog Post prominently links to and describes both the Penn State and NSF reports, stating that the former "declared him innocent of any wrongdoing" and that the latter did so as well, and links to other materials that are critical of Mann's research methodology.[53] *See* Ex. 6 and attachments. This is indistinguishable from *Guilford*, in which the challenged column was harshly critical of the plaintiffs' stance toward organized labor and described several of plaintiffs' run-ins with labor unions in critical language but also accurately "disclose[d] that the ICC ruled in Guilford's favor on some issues and that Guilford had engaged in contested litigation with the union and with Amtrak." 760 A.2d at 601. In this instance, as in *Guilford*, "the reader is therefore free to draw his or her own conclusions regarding whether the plaintiffs acted wrongfully." *Id.* And while that reader might perceive that the CEI Defendants are not sympathetic to Mann, "that surely does not render the column defamatory." *Id.*

[53] Mann's statement that some of the "disclosed facts . . . are authored by Mr. Simberg himself" is plainly false. *See* Opp. at 49. A person cannot "author[]" a fact, only report it. The two articles by Simberg linked in the Blog Post report facts from the scientific literature, news reports, the Climategate emails, and other sources. *See* Ex. 6, attachs. B, C. Notably, Mann makes no attempt to challenge the facts reported in those articles. *See* Opp. at 49-50.

Accordingly, Mann's burden is to show that each challenged "statement is 'so obviously false' that 'no reasonable person could find that [its] characterizations were supportable interpretations' of the underlying facts." *Washington*, 317 U.S. App. D.C. at 81, 80 F.3d at 557 (quoting *Moldea II*, 306 U.S. App. D.C. at 8, 22 F.3d at 317). Mann makes no attempt to do so. *See* Opp. at 48-50.

Mann's argument regarding application of the District of Columbia's "fair comment" privilege is incoherent. Mann argues that, even if the challenged statements were expressions of opinion, they would still not be protected by the privilege because "the law protects only opinions based on true facts, accurately disclosed." Opp. at 50 (internal quotation marks omitted). But after stating that, Mann points to no misstatement of facts or failure to disclose that would defeat application of the privilege, instead simply concluding that "Defendants' statements do not offer an opinion regarding Dr. Mann . . . ." Opp. at 51.

Indeed, the leading case cited by Mann supports application of the privilege. *Fisher v. Washington Post Co.*, 212 A.2d 335, 337 (D.C. 1965), makes clear that, "[s]o long as the comment is the speaker's actual opinion, based on fact, about a matter of public interest, the words are protected . . . ." On that basis, it dismissed claims regarding the statement that a gallery's show was "badly hung," recognizing that "opinions could differ on such matters." *Id.* In dismissing the lawsuit, it also rejected the argument "for opinion to be protected by the fair comment doctrine, the facts upon which it is based must be stated or referred to so that the reader might draw his own conclusions." *Id.* at 338. It is enough that "the facts are available to the public . . . ." *Id.* In this instance, the challenged statements were highly critical of Mann's research, a factual predicate that is

clearly available to the public.[54] But rather than leave readers to find Mann's research on their own, the Blog Post links to Mann's webpage and to a wealth of commentary on Mann's research and conduct. *See* Ex. 6 and attachments. Accordingly, the fair comment privilege applies here, providing an additional ground for dismissal of Mann's claims.

6. Mann Abandons His Emotional Distress Claim

Mann fails to address or even mention *Hustler Magazine, Inc. v. Falwell*, 485 U.S. 46, 50 (1988), which spells out the First Amendment's limitation of "a State's authority to protect its citizens from the intentional infliction of emotional distress." As described in the CEI Defendants' Anti-SLAPP Motion at 57-58, the plaintiff in that case, a well-known minister, claimed that the defendant, a pornographic magazine, had intentionally subjected him to emotional distress by publishing an article purporting to be an interview with him "in which he states that his 'first time' was during a drunken incestuous rendezvous with his mother in an outhouse." *Id.* at 48. Because "that speech could not reasonably have been interpreted as stating actual facts about the public figure involved," the Supreme Court held that the First Amendment precluded liability, even for "speech that is patently offensive and is intended to inflict emotional injury." *Id.* at 50.

Hustler compels the same result here. Mann identifies no "actual fact" that can be discerned from the CEI Defendant's statement that "Mann could be said to be the Jerry Sandusky of climate science." *See* Opp. at 57-58. Nor could he: no "actual fact" regarding Mann is apparent, particularly given that the Blog Post expressly stated that Mann was not engaged in "molesting children." *See*

---

[54] *Michael E. Mann: Research*, Penn State University Department of Meteorology, http://www.meteo.psu.edu/holocene/public_html/Mann/research/research.php.

Compl. ¶26. The CEI Defendants' comparison is not actionable because it is a statement of pure opinion and hyperbole, not a false or even verifiable assertion of fact.

In that light, the cases cited by Mann are irrelevant. Two do not involve expressive conduct at all and therefore do not implicate the protections of the First Amendment. *Kotsch v. District of Columbia*, 924 A.2d 1040, 1046 (D.C. 2007) (an arrest); *Muratore v. M/S Scotia Prince*, 845 F.2d 347, 352-53 (1st Cir. 1988) (stalking and other harassment by photographers on a cruise ship). One predates *Hustler* by 18 years and raises no First Amendment issue. *Moore v. Greene*, 431 F.2d 584, 591 (9th Cir. 1970). And the last involved statements that plainly did state "actual facts." *Kolegas v. Heftel Broad. Corp.*, 607 N.E.2d 201, 212 (Ill. 1992). Specifically, the defendants, a radio station and its hosts, had broadcast that the plaintiff's wife, afflicted with neurofibromatosis, "was so hideous that no one would marry her except under duress" and that the plaintiff's "wife and five-year-old child," also afflicted with neurofibromatosis, "had deformed heads." *Id.* Nonetheless, the court's opinion does not address any First Amendment defense to the plaintiff's emotional distress claim.

Mann does not even attempt to establish that his emotional distress claim can survive the CEI Defendants' First Amendment defense. It must therefore be dismissed.

C. Mann Makes No Real Attempt to Distinguish Case Law Holding that a Hyperlink Is Not Republication

Mann's argument that CEI can be liable for *National Review* editor Rich Lowry's characterization of Mann's research as "intellectually bogus" fails to seriously address the consistent line of cases holding that a party cannot be liable for hyperlinking to allegedly defamatory statements

so long as it does not itself publish those statements.[55] *See* CEI Defendants' Anti-SLAPP Motion at 56-57.

To begin with, Mann provides no support for his contention that CEI's comment that Lowry "expertly summed up the matter" converted its hyperlink into republication. Opp. at 55-56. In *U.S. ex rel. Klein v. Omeros Corp.*, No. C09-1342-JCC, 2012 WL 4874031, at *10 (W.D. Wash. Oct. 15, 2012), the court recognized as black-letter law that, "[u]nder traditional principles of republication, a mere reference to an article, regardless how favorable it is as long as it does not restate the defamatory material, does not republish the material." *Id.* (quoting *In re Phila. Newspapers, LLC*, 690 F.3d 161, 175 (3d Cir. 2012)). The key to a traditional republication, it explained, "is that it *presents the material*, *in its entirety*, before a new audience. A mere reference to a previously published article does not do that. While it may call the *existence* of the article to the attention of a new audience, it does not present the *defamatory* contents of the article to that audience." *Id.* at *11 (quoting *Salyer v. S. Poverty Law Ctr., Inc.*, 701 F. Supp. 2d 912, 916 (W.D. Ky. 2009)) (emphasis in original). This case is indistinguishable from *Klein*, in that CEI did not restate Lowry's allegedly defamatory statement. *See* Compl., Ex. D.

Second, Mann simply asserts, again without any support or even any reasoning, that cases concerning the effect of republication on the running of limitations periods are not relevant here. But the threshold inquiry in each instance is *whether a new publication occurred at all*, the same matter at issue here.[56] This is because, "under the single publication rule, the statement is considered

---

[55] Although he uses the plural term "defamatory statements," Opp. at 55, Mann elsewhere clarifies that he claims only one statement to be defamatory: "intellectually bogus," Opp. at 41.

[56] Mann's Complaint alleges that CEI's hyperlink "adopted *and republished* Mr. Lowry's defamatory statement." Compl. ¶84 (emphasis added).

published and the statute of limitations runs as soon as the communication enters the stream of commerce." *Salyer*, 701 F. Supp. 2d at 915 (internal quotation marks omitted). Republication, however, is an exception to that rule: "Republishing material in a new edition, editing and republishing it, or placing it in a new form resets the statute of limitations." *Id.* at 914 (internal quotation marks omitted). Regardless, *Klein,* which did not involve the running of a limitations period, relied on the analysis of cases that do in support of its holding that a hyperlink, without the restatement of the link's allegedly defamatory contents, is not actionable. 2012 WL 4874031, at *11.

Third and finally, Mann presents no support for his argument that to "endorse" allegedly defamatory speech, without ever repeating it, is itself defamation. It is not, because an endorsement lacks the central element of a defamation claim: publication of "a false and defamatory statement concerning the plaintiff." *LeFande v. District of Columbia*, 864 F. Supp. 2d 44, 51 (D.D.C. 2012); *see also* Restatement (Second) of Torts § 558(a) (1977).

Lowry's characterization of Mann's research as "intellectually bogus" is an expression of pure opinion entitled to the First Amendment's strongest protections. But it is also a statement that CEI never made or republished and for which it therefore could not be liable.

D. Mann's Collateral Estoppel Argument Is A Red Herring Because Neither the EPA Nor D.C. Circuit Resolved Any Matter at Issue in This Litigation

Mann attempts to make an end-run around one of the key elements of libel: the question of truth or falsity of the challenged statements. He purports to "believe[] that Defendants will concede that their statements were false (especially in light of the fact that they have not argued to the contrary in their briefs)." Opp. at 41 n.78. But if not, he claims, "CEI will be collaterally estopped from asserting that its statements are true based upon its participation in the EPA proceedings and subsequent appeal to the District of Columbia Circuit." *Id.* Both of these contentions are false.

The CEI Defendants have never conceded that the challenged statements are false and at no time—neither now nor at the moment of publication—have believed them to be. Mann's assertion

that the CEI Defendants "have not argued to the contrary" in their prior briefs is irrelevant. *See* Opp. at 41 n.78. It is well established that a motion to dismiss "tests the legal sufficiency of a complaint." *Winston v. Clough*, 712 F. Supp. 2d 1, 5 (D.D.C. 2010) (quoting *Smith-Thompson v. District of Columbia,* 657 F. Supp. 2d 123, 129 (D.D.C. 2009). The CEI Defendants' briefs thus far have focused on the *legal issues* currently before this Court, and not (as Mann would apparently require) the *factual issues* that will be litigated if his claims are not dismissed at this stage.

Mann's argument, confined to a footnote, Opp. at 41 n.78, that CEI is collaterally estopped from asserting the truth of the challenged statements is equally mistaken. That footnote references the EPA litigation and sets forth the four-prong test for collateral estoppel. Without actually applying that test to the facts of this case, Mann simply asserts that "all of the necessary elements for collateral estoppel are present." *Id.*

As an initial matter, the values of the First Amendment make estoppel especially inappropriate here. The heavy burden that a libel plaintiff must bear—particularly in cases (like this one) involving a public figure and issues of public concern—provides the "breathing space" that is required for freedom of expression to survive. *See New York Times Co. v. Sullivan*, 376 U.S. 254, 271-72 (1964). For that reason, it is "a constitutional requirement that the plaintiff bear the burden of showing falsity," even though that rule may result in the dismissal of some meritorious claims. *Phila. Newspapers, Inc. v. Hepps*, 475 US 767, 776 (1986). Mann seeks not only to stifle public debate through this lawsuit, but to do so while avoiding the traditional safeguards of free debate that apply to every other plaintiff in every other lawsuit implicating First Amendment freedoms.

On the merits, Mann's estoppel argument fails in at least three ways. Estoppel has four "foundational requirements": "(1) the issue [was] actually litigated and (2) determined by a valid, final judgment on the merits; (3) after a full and fair opportunity for litigation by the parties or their privies; (4) under circumstances where the determination was essential to the judgment, and not

merely dictum." *Modiri v. 1342 Rest. Group, Inc.*, 904 A.2d 391, 395-96 (D.C. 2006) (quoting *Davis v. Davis*, 663 A.2d 499, 501 (D.C. 1995)).

First, the truth or falsity of the CEI Defendants' statements was never "actually litigated," nor was any claim regarding Mann's research or conduct. As discussed in *supra* § I.D, CEI filed a petition to reconsider of EPA's Endangerment Finding in 2010, raising issues relating to the Climategate scandal, and the EPA denied it and other such petitions on the basis that the research implicated by Climategate was irrelevant to its Endangerment Finding because it had relied on so many other studies. 75 Fed. Reg. 49,556 (Aug. 13, 2010). The D.C. Circuit upheld that reasoning in *Coalition for Responsible Regulation v. EPA*, 401 U.S. App. D.C. 306, 684 F.3d 102 (D.C. Cir. 2012).

Mann argues that the D.C. Circuit's decision forecloses any consideration of the truth or falsity of statements made nearly two years after EPA's petition denial, because EPA's denial specifically disproves any "data manipulation" on his part. *See* Opp. at 41 n.78. But even if a general claim like "data manipulation" was the kind of thing that could be proven true or false, *see supra* § II.B.4, those proceedings did not "actually litigate" the issue or determine it "on the merits." *Compare Modiri*, 904 A.2d at 395-96; *Davis*, 663 A.2d at 501. EPA's notice mentions Mann only once, in a footnote citation to a 2009 paper. 75 Fed. Reg. at 49,571/3, P. Ex. 11. EPA does not claim that it conducted any independent investigation of Climategate, but only that it "has reviewed all of the CRU emails." *Id.* at 49,581/1. EPA's notice does not claim to have made any decision regarding Mann or his research. Instead, EPA determined that Climategate implicated only a small portion of the research on which its Endangerment Finding relied, such that it had no reason to reconsider the Endangerment Finding. *Id.* at 49,571/3.

Because the D.C. Circuit upheld EPA's reasoning, its decision has nothing to say about Mann's conduct or research. CEI and other petitioners challenged the EPA's denial of their petitions for reconsideration as "arbitrary, capricious, or otherwise contrary to law." *Competitive*

Enter. Inst. v. EPA, Nonbinding Statement of Issues (Nov. 17, 2010), Ex. C. The D.C. Circuit disagreed. *See Coalition for Responsible Regulation v. EPA*, 401 U.S. App. D.C. 306, 684 F.3d 102 (D.C. Cir. 2012). It reasoned that the agency had not acted in an arbitrary and capricious fashion when it had relied on the IPCC assessment, which in turn "relied on around 18,000 studies that were peer-reviewed." *Id.* at 329, 684 F.3d at 125. In the Court's view, as in EPA's, Mann's research and his conduct were irrelevant to the question before it: whether EPA had acted irrationally in denying the petitions for reconsideration. Accordingly, the D.C. Circuit's decision does not mention Mann, the "hockey stick" diagram, or any other subject that would bear on the truth or falsity of the CEI Defendants' statements.

In short, the parties did not actually litigate the question of "data manipulation" by Mann, and the court did not decide the issue. Collateral estoppel cannot apply where "it is not clear" that the issue a party seeks to preclude "was actually determined." *Bobby v. Bies*, 556 U.S. 825, 834 (2009). Here, it is clear that the issue was not determined at all. Moreover, the arbitrary and capricious standard applies in that case is nothing like the burden a libel plaintiff bears, as a constitutional requirement, to *prove* falsity. Even if the D.C. Circuit had ruled that EPA did not act irrationally by relying on Mann's research, that ruling would not have any estoppel effect here. *Patton v. Klein*, 746 A.2d 866, 871 (D.C. 1999) ("Collateral estoppel does not apply if the issues are not identical, even if the issues are similar."); *Hutchinson v. D.C. Office of Emp. Appeals*, 710 A.2d 227, 236 (D.C. 1998) (no estoppel "when the issues in the prior and current litigation are not identical").

Second, because the truth or falsity of the CEI Defendants' statements was never litigated, collateral estoppel fails because that question was not "determined by a valid, final judgment on the merits." And, third, because that determination was never made, it was certainly not essential to any judgment. Indeed, the EPA's and D.C. Circuit's precise reasoning was that determination of any such issue was completely irrelevant.

III. Mann's Claims Fail Under Rule 12(b)(6) Because Mann Did Not Plausibly Allege that the CEI Defendants Acted with Actual Malice

Mann's claims must be dismissed because he fails to plausibly allege facts that would establish that the CEI Defendants acted with actual malice. *See* Rule 12(b)(6) Motion at 9-15. Rather than confront this argument, Mann parrots back the legal standards and asserts that he satisfies them. This does not render plausible any claim that the CEI Defendants "must have made the false publication with a high degree of awareness of probable falsity, or must have entertained serious doubts as to the truth of [their] publication." *Harte-Hanks Commc'ns, Inc. v. Connaughton*, 491 U.S. 657, 667 (1989) (internal quotation marks and citations omitted). At the same time, Mann has abandoned any attempt to show actual malice in support of his emotional distress claim.

A. Mann's Libel Claims Should Be Dismissed Because He Fails To Plausibly Allege Actual Malice

The CEI Defendants' Rule 12(b)(6) Motion listed each and every factual allegation contained in Mann's Complaint that might be thought to support his legal allegation that the CEI Defendants acted with actual malice. Rule 12(b)(6) Motion at 3-4. It then proceeded to address each allegation. *Id.* at 10-14. As it showed, the bulk of those allegations are plainly conclusory and therefore must be discarded under the first step of *Iqbal* and *Twombly*. *Id.* at 11. The two that remained, while pleading facts, were not "enough to raise a right to relief above the speculative level" under *Iqbal* and *Twombly's* second step. *Parisi v. Sinclair*, 845 F. Supp. 2d 215, 217-18 (D.D.C. 2012).

Mann now abandons all of his allegations of actual malice but one: that the CEI Defendants read certain investigations that "found that there was no evidence of any fraud, data falsification,

statistical manipulation, or misconduct." *See* Opp. at 59-60 (citing Compl. ¶24).[57] Mann then concludes that the CEI Defendants must have acted with actual malice, because there is "simply no way anyone could read the litany of inquiries . . . without coming to the conclusion that Dr. Mann was not guilty of fraud, misconduct, or data manipulation." Opp. at 60.

That is wrong, in four respects. First, this is nothing more than a conclusory allegation that a statement was made "with knowledge that it was false." In every single case where actual malice is at issue, the plaintiff could simply state that some book or website contradicts the challenged statement and the defendant was surely aware of that book or website. Such an approach is incompatible with the Court of Appeal's admonition that, "in the First Amendment area, summary procedures are essential." *Klayman v. Segal*, 783 A.2d 607, 613 n.5 (D.C. 2001). It also falls far short of *Iqbal's* requirement that the plaintiff must "plead[] factual content that allows the court to draw the reasonable inference that the defendant is liable for the misconduct alleged." *Ashcroft v. Iqbal*, 556 U.S. 662, 678 (2009).

Second, none of the challenged statements accuse Mann of criminal misconduct, data falsification, or anything of the sort, but of conducting research that is biased or has been oversold. None of the reports cited in the Complaint addressed these issues, *see supra* § I.B, and so they are irrelevant as to whether the CEI Defendants acted with actual malice. The CEI Defendants, of course, had ample grounds to believe that Mann's science was shoddy, and there can therefore be no question that they acted without actual malice. *See, e.g., supra* § I.A (discussing concerns raised regarding Mann's research and statistical methodology); CEI Defendants' Anti-SLAPP Motion at 9-

[57] Because Mann defends only this one allegation as supporting actual malice, all of his claims necessarily fail if the Court finds that allegation to be insufficient.

12 (discussing criticisms of Mann's research); Blakeley B. McShane and Abraham J. Wyner, Ex. B (peer-reviewed article calling Mann's research into question).

Third, even accepting Mann's implausible interpretation of the challenged statements for the sake of argument, the supposedly exonerating reports that he cites do not contradict them. Mann himself acknowledges that the investigations only "found that there was no evidence" of fraud by Mann, not that he was determined to be innocent of it. Compl. ¶24; Opp. at 59. Accordingly, the reports do not show that the CEI Defendants' statements were made with the knowledge that they were false.

Fourth, Mann's allegation regarding the reports speaks, if at all, to falsity, not actual malice. Recognizing that distinction, the Supreme Court held that a plaintiff had failed to prove actual malice based on knowledge of falsity even where newspaper had a "reasonable doubt" as to falsity. *New York Times*, 367 U.S. at 286-87. Even if Mann could show that the challenged statements are false—which he cannot—he cannot show and does not plausibly allege that "the defendant[s] in fact entertained serious doubts as to the truth" of their publication. *St. Amant v. Thompson*, 390 U.S. 727, 731 (1968).

Finally, Mann quibbles over the import of a single Sixth Circuit decision, *Cobb v. Time*, 278 F.3d 629 (6th Cir. 2002). *See* Opp. at 60 (claiming that Defendants "misleadingly" cite *Cobb*). *Cobb* speaks for itself. The Sixth Circuit reversed a finding of actual malice because the defendant magazine attempted to corroborate a story that ultimately turned out to be false. *See* 278 F.3d at 640. The court noted that the reporters had obtained information "from at least one independent source," and found that the record did not "support the conclusion that [the defendant] intentionally avoided learning the truth . . . ." *Id.* Here, the CEI Defendants went above and beyond what the law requires by performing an investigation, *see Harte-Hanks Commc'ns,* 491 U.S. at 688 (explaining that "failure to investigate" is itself "not sufficient to establish reckless disregard"), and they found

materials that "bristled with ambiguities," *Time, Inc. v. Pape*, 401 U.S. 279, 290 (1971), regarding Mann's research. *See* Rule 12(b)(6) Motion at 13-14. That, in turn, precludes as a matter of law any finding of actual malice. *Pape*, 401 U.S. at 290.

Because Mann "pleads facts that are merely consistent with a defendant's liability," his allegations "do not permit the court to infer more than the mere possibility of misconduct." *Iqbal*, 556 U.S. at 678-79 (internal quotation marks omitted). His libel claims should be dismissed.

B. Mann Abandons His Emotional Distress Claim

The one allegation regarding actual malice that Mann continues to defend, Opp. at 59-60, provides no support to Mann's emotional distress claim, which must therefore be dismissed.

Mann's burden at this stage is to show that CEI made the challenged statement—"Mann could be said to be the Jerry Sandusky of climate science"—with actual malice. This is a requirement of the Supreme Court's decision in *Hustler*, which held that "public figures and public officials may not recover for the tort of intentional infliction of emotional distress by reason of publications such as the one here at issue without showing in addition that the publication contains a false statement of fact which was made with 'actual malice,' i.e., with knowledge that the statement was false or with reckless disregard as to whether or not it was true." 485 U.S. at 56.

Yet the one allegation regarding actual malice that Mann has not abandoned provides no indication of the truth or falsity of this comparison. That allegation states, in its entirety:

> All of the above investigations found that there was no evidence of any fraud, data falsification, statistical manipulation, or misconduct of any kind by Dr. Mann. All of the above reports and publications were widely available and commented upon in the national and international media. All were read by the Defendants. To the extent there was ever any question regarding the propriety of Dr. Mann's research, it was laid to rest as a result of these investigations.

Compl. ¶24. This, of course, says nothing about Jerry Sandusky or whether and how a comparison of Mann and Sandusky might be true or false, let alone whether the CEI Defendants knew such a comparison to be false. Because Mann has abandoned any attempt to show that the CEI

Defendants acted with actual malice when they published this statement, his emotional distress claim must be dismissed.

IV. Mann Is Not Entitled to Attorney's Fees and Costs

If the D.C. Anti-SLAPP Act awarded fees and costs for chutzpah, Mann would have a strong case. He began this case by filing a Complaint containing outright falsehoods regarding his status as a Nobel Laureate and seeking legal sanctions for the same kind of heated rhetoric in which he often traffics. While many suspected that this lawsuit was intended principally to harass and silence his critics, Mann stepped forward to confirm that this was so, telling the *Atlantic* that this case "is about saying 'enough is enough'" and harassing those who "want to attack this iconic graph." He posted a message for his Facebook followers describing this lawsuit as part of a "larger battle" against "groups seeking to discredit the case for concern over climate change," and expressing his hope that such groups will be "silenced." Ex. 9; Ex. 10. And after he saw that the CEI Defendants were monitoring his public Facebook page, in December he began systematically deleting public posts and comments relating to the subject matter of this litigation, including a comment published there for months stating that those who disagree with the theory of man-made global warming "are comparable to Jerry Sandusky." Ex. D. And now, in the same filing in which he abandons his emotional distress claim, *see supra* §§ II.B.6, III.B, and silently abandons his challenge to several statements mentioned in his Complaint, *see supra* § II, he seeks attorneys' fees and costs from the CEI Defendants, claiming that their attempt to take advantage of the D.C. Anti-SLAPP Act, *in a lawsuit that he concedes is subject to the Act*, is somehow illegitimate. Opp. at 61-62.

Unfortunately for Mann, chutzpah is not the governing standard. Mann's burden is to show that the CEI Defendants' motion "is frivolous or is solely intended to cause unnecessary delay." D.C. Code § 16-5504(b). It is plainly not, given that Mann concedes that the CEI Defendants have made a *prima facie* showing that the Act applies, Opp. at 37, and given that, in response to the CEI

Defendants' motion, Mann narrowed his claims. Mann's principal argument to the contrary is that he is not only right on the merits, but that the merit of his claims is "abundantly clear." Opp. at 61. The CEI Defendants respectfully disagree and believe that the law and facts are on their side on that point. *See supra* §§ I-III. As for Mann's assertion that the CEI Defendants "deliberately misled the Court, mischaracterized the facts underlying the lawsuits, and . . . simply ignored highly material facts," it is offensive and incorrect, but it is also of a piece with the remainder of his brief, which is long on rhetoric and short on legal argument and detail. The CEI Defendants are honestly puzzled by Mann's strange fixation on the EPA's decision to deny reconsideration of its Endangerment Finding, given that the agency did not purport to pass on any issue now raised in this litigation and given that, faced with challenges to Mann's research, the agency chose to throw it under the bus rather than to defend it. *See supra* §§ I.D, II.D.

Finally, there is nothing "cynical," Opp. at 62, about the CEI Defendants' attempt to defend themselves against Mann's attempt to silence his critics through abuse of legal process. The debate over global warming is vigorous, it accommodates many disparate views, and it is vitally important to the choices that our Nation will make in the years ahead. This lawsuit seeks to stifle that debate. The CEI Defendants seek to protect their own free speech rights and those of others—whether or not they agree or disagree with CEI—to speak freely on this issue without fear of being sued. Their belief that this kind of uninhibited debate is necessary to our system of self-government is not cynical but heartfelt. *Cf.* U.S. Const., amend. I.

CONCLUSION

For the reasons stated here, in the CEI Defendants' Anti-SLAPP Motion, and in the CEI Defendants' Rule 12(b)(6) Motion, the CEI Defendants respectfully request that the Court dismiss, with prejudice, all of Mann's claims against them.

Dated: February 1, 2013

Respectfully submitted,

By: /s/ David B. Rivkin, Jr.

David B. Rivkin, Jr. (D.C. Bar No. 394446)
Bruce D. Brown (D.C. Bar No.457317)
Mark I. Bailen (D.C. Bar No. 459623)
Andrew M. Grossman (D.C. Bar No. 985166)
BakerHostetler LLP
Washington Square, Suite 1100
1050 Connecticut Avenue, NW
Washington, DC 20036
Tel: (202) 861-1500
Facsimile: (202) 861-1783
drivkin@bakerlaw.com
bbrown@bakerlaw.com
mbailen@bakerlaw.com
agrossman@bakerlaw.com

Counsel for Defendants Competitive Enterprise Institute and Rand Simberg

SUPERIOR COURT OF THE DISTRICT OF COLUMBIA
CIVIL DIVISION

MICHAEL E. MANN, PH.D.,	)	
Plaintiff,	)	Case No. 2012 CA 008263 B
v.	)	Judge Natalia Combs Greene
NATIONAL REVIEW, INC., et al.,	)	
Defendants.	)	

DECLARATION OF ANDREW M. GROSSMAN IN SUPPORT OF DEFENDANTS COMPETITIVE ENTERPRISE INSTITUTE AND RAND SIMBERG'S REPLY BRIEF

Pursuant to Rule 43(e), Superior Court Rules of Civil Procedure, I, Andrew M. Grossman, declare as follows:

1. I am counsel in this matter for Defendants for Competitive Enterprise Institute and Rand Simberg (collectively, "the CEI Defendants"). I submit this Declaration in Support of the CEI Defendants' Special Motion to Dismiss Pursuant to the D.C. Anti-SLAPP Act ("Motion"). I have personal knowledge of the facts stated herein, and if called as a witness, I could and would competently testify thereto.

2. Exhibit A attached to the Reply is a true and correct copy of page 73 of the third volume of the U.S. Environmental Protection Agency's "Response to the Petitions to Reconsider the Endangerment and Cause or Contribute Findings for Greenhouse Gases under Section 202(a) of the Clean Air Act," published in 2010.

3. Exhibit B attached to the Reply is a true and correct copy of the article "A Statistical Analysis of Multiple Temperature Proxies: Are Reconstructions of Surface Temperatures Over the Last 1000 Years Reliable?" by Blakely B. McShane and Abraham J. Wyner, as published in volume 5, number 1, of *The Annals of Applied Statistics*.

4. Exhibit C attached to the Reply is a true and correct copy of the "Nonbinding Statement of Issues" filed by the Defendant Competitive Enterprise Institute in D.C. Circuit case no. 10-1318 on November 17, 2010.

5. Exhibit D attached to the Reply is a true and correct copy of a comment thread captured from Plaintiff Michael E. Mann's Facebook page in December 2012.

Executed this 1st day of February, 2013, in Washington, D.C.

/s/ Andrew M. Grossman
Andrew M. Grossman

CERTIFICATE OF SERVICE

I hereby certify that on February 1, 2013, I caused a copy of the foregoing Motion to be served by CaseFileXpress on the following:

John B. Williams
Bernard S. Grimm
Catherine Rosato Reilly
Cozen O'Connor
1627 I Street, N.W., Suite 1100
Washington, D.C. 20006
(202) 912-4800
Fax: (877) 260-9435
jbwilliams@cozen.com

Peter J. Fontaine
Cozen O'Connor
1900 Market Street
Philadelphia, PA 19103
(856) 910-5043
Fax: (866) 850-7491
pfontaine@cozen.com

Counsel for Plaintiff

Shannen W. Coffin
Steptoe & Johnson LLP
1330 Connecticut Avenue, N.W.
Washington, D.C. 20036
(202) 429-6255
scoffin@steptoe.com

Counsel for Defendants National Review and Mark Steyn

By: /s/ Andrew M. Grossman
Andrew M. Grossman

Exhibit A

EPA's Response to the Petitions to Reconsider the Endangerment and Cause or Contribute Findings for Greenhouse Gases under Section 202(a) of the Clean Air Act

Volume 3: Process Issues Raised by Petitioners

U.S. Environmental Protection Agency
Office of Atmospheric Programs
Climate Change Division
Washington, D.C.

Comment (3-34):
Peabody Energy claims that Michael Mann and Tom Wigley in particular "resorted to admonishment of those they disagreed with." Peabody Energy cites the example of Mexican climate researcher Jorge Sánchez-Sesma, who, in an e-mail to Phil Jones sent December 3, 2004, said that he met Mann at a conference. According to Peabody Energy, "at first Mann was 'very kind,' but when Mann found out Sánchez-Sesma's work ran counter to some of Mann's conclusions, Mann 'changed his attitude.'"[64]

In addition, Peabody Energy claims that Michael Mann "continued his battle" against climate change "skeptics" like Stephen McIntyre and Ross McKitrick in the press, quoting an e-mail from Mann to New York Times reporter Andy Revkin, sent on February 8, 2005:

> The McIntyre and McKitrick paper is pure scientific fraud. I think you'll find this reinforced by just about any legitimate scientist in our field you discuss this with. To recap, I hope you don't mention MM [McIntyre and McKitrick] at all. It really doesn't deserve any additional publicity.[65]

Peabody Energy also claims that Tom Wigley engaged in character assassination of two highly credentialed scientists—John Christy and Chris de Freitas—in an attempt to ruin their careers. As evidence, Peabody Energy quotes the following e-mail from Wigley, sent on August 19, 2003:

> Jim Titus mentioned to me that in the legal profession here people are disbarred for behavior like that of De Freitas (and even John Christy -- although this is a more subtle case). We cannot do that of course, but we can alert the community of honest scientists to such behavior and formally discredit these people.[66]

Peabody Energy concludes that the goal of the CRU e-mail authors' "intimidation/bad-mouthing [was] to influence scientific development" and effectively prevent "dissenting scientific voices from being heard" at conferences or in the climate science literature. Peabody Energy argues that the Endangerment Finding should be reconsidered because it relies on peer-reviewed literature, which is biased in favor of views supporting anthropogenic climate change and does not reflect the true breadth of the science.

Response (3-34):
The petitioner uses language such as "admonished," "battled," "engaged in character assassination," and "intimidation/bad mouthing" in an attempt to demonstrate that two scientists, Michael Mann and Tom Wigley, improperly biased the peer reviewed literature, making the body of peer reviewed literature unreliable. Examination of the petitioner's argument does not support their claims. With respect to interaction between Mann and Jorge Sánchez-Sesma, we

[64] E-mail file 1079384474.txt (March 15, 2004), page 661, line 45 of the PDF version entitled: CRU Emails 1996-2009.pdf.
[65] E-mail file 1107899057.txt (February 8, 2005), page 825, line 31 of the PDF version entitled: CRU Emails 1996-2009.pdf.
[66] E-mail file 1061298033.txt (August 19, 2003), page 566, line 15 of the PDF version entitled: CRU Emails 1996-2009.pdf.

Exhibit B

The Annals of Applied Statistics
2011, Vol. 5, No. 1, 5–44
DOI: 10.1214/10-AOAS398

A STATISTICAL ANALYSIS OF MULTIPLE TEMPERATURE PROXIES: ARE RECONSTRUCTIONS OF SURFACE TEMPERATURES OVER THE LAST 1000 YEARS RELIABLE?[1]

By Blakeley B. McShane and Abraham J. Wyner

Northwestern University and the University of Pennsylvania

Predicting historic temperatures based on tree rings, ice cores, and other natural proxies is a difficult endeavor. The relationship between proxies and temperature is weak and the number of proxies is far larger than the number of target data points. Furthermore, the data contain complex spatial and temporal dependence structures which are not easily captured with simple models.

In this paper, we assess the reliability of such reconstructions and their statistical significance against various null models. We find that the proxies do not predict temperature significantly better than random series generated independently of temperature. Furthermore, various model specifications that perform similarly at predicting temperature produce extremely different historical backcasts. Finally, the proxies seem unable to forecast the high levels of and sharp run-up in temperature in the 1990s either in-sample or from contiguous holdout blocks, thus casting doubt on their ability to predict such phenomena if in fact they occurred several hundred years ago.

We propose our own reconstruction of Northern Hemisphere average annual land temperature over the last millennium, assess its reliability, and compare it to those from the climate science literature. Our model provides a similar reconstruction but has much wider standard errors, reflecting the weak signal and large uncertainty encountered in this setting.

1. Introduction. Paleoclimatology is the study of climate and climate change over the scale of the entire history of earth. A particular area of focus is

Received January 2010; revised August 2010.

[1]Discussed in 10.1214/10-AOAS398I, 10.1214/10-AOAS398M, 10.1214/10-AOAS398C, 10.1214/10-AOAS398L, 10.1214/10-AOAS398G, 10.1214/10-AOAS398D, 10.1214/10-AOAS398H, 10.1214/10-AOAS398B, 10.1214/10-AOAS398K, 10.1214/10-AOAS398E, 10.1214/10-AOAS398F, 10.1214/10-AOAS398J, 10.1214/10-AOAS409; rejoinder at 10.1214/10-AOAS398REJ.

Key words and phrases. Climate change, global warming, paleoclimatology, temperature reconstruction, model validation, cross-validation, time series.

temperature. Since reliable temperature records typically exist for only the last 150 years or fewer, paleoclimatologists use measurements from tree rings, ice sheets, and other natural phenomena to estimate past temperature. The key idea is to use various artifacts of historical periods which were strongly influenced by temperature and which survive to the present. For example, Antarctic ice cores contain ancient bubbles of air which can be dated quite accurately. The temperature of that air can be approximated by measuring the ratio of major ions and isotopes of oxygen and hydrogen. Similarly, tree rings measured from old growth forests can be dated to annual resolution, and features can be extracted which are known to be related to temperature.

The "proxy record" is comprised of these and many other types of data, including boreholes, corals, speleothems, and lake sediments [see Bradley (1999) for detailed descriptions]. The basic statistical problem is quite easy to explain. Scientists extract, scale, and calibrate the data. Then, a training set consisting of the part of the proxy record which overlaps the modern instrumental period (i.e., the past 150 years) is constructed and used to build a model. Finally, the model, which maps the proxy record to a surface temperature, is used to backcast or "reconstruct" historical temperatures.

This effort to reconstruct our planet's climate history has become linked to the topic of Anthropogenic Global Warming (AGW). On the one hand, this is peculiar since paleoclimatological reconstructions can provide evidence only for the *detection* of global warming and even then they constitute only one such source of evidence. The principal sources of evidence for the detection of global warming and in particular the *attribution* of it to anthropogenic factors come from basic science as well as General Circulation Models (GCMs) that have been tuned to data accumulated during the instrumental period [IPCC (2007)]. These models show that carbon dioxide, when released into the atmosphere in sufficient concentration, can force temperature increases.

On the other hand, the effort of world governments to pass legislation to cut carbon to pre-industrial levels cannot proceed without the consent of the governed and historical reconstructions from paleoclimatological models have indeed proven persuasive and effective at winning the hearts and minds of the populace. Consider Figure 1 which was featured prominently in the Intergovernmental Panel on Climate Change report [IPCC (2001)] in the summary for policy makers.[1] The sharp upward slope of the graph in the

[1]Figure 1 appeared in IPCC (2001) and is due to Mann, Bradley and Hughes (1999) which is in turn based on the analysis of multiple proxies pioneered by Mann, Bradley and Hughes (1998). Figure 2 is a "spaghetti graph" of multiple reconstructions appearing in Mann et al. (2008). Figure 3 appeared in NRC (2006).

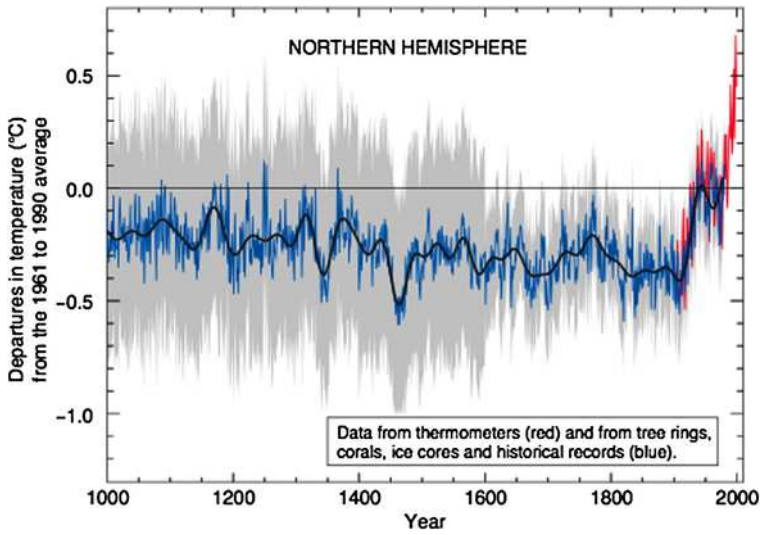

FIG. 1. *Multiproxy reconstruction of Northern Hemisphere surface temperature variations over the past millennium (blue), along with* 40*-year average (black), a measure of the statistical uncertainty associated with the reconstruction (gray), and instrumental surface temperature (red), based on the work by Mann, Bradley and Hughes (1999). This figure has sometimes been referred to as the "hockey stick." Source: IPCC (2001).*

late 20th century is visually striking, easy to comprehend, and likely to alarm. The IPCC report goes even further:

> Uncertainties increase in more distant times and are always much larger than in the instrumental record due to the use of relatively sparse proxy data. Nevertheless the rate and duration of warming of the 20th century has been much greater than in any of the previous nine centuries. Similarly, **it is likely that the 1990s have been the warmest decade and 1998 the warmest year of the millennium.** [Emphasis added]

Quotations like the above and graphs like those in Figures 1–3 are featured prominently not only in official documents like the IPCC report but also in widely viewed television programs [BBC (2008)], in film [Gore (2006)], and in museum expositions [Rothstein (2008)], alarming both the populace and policy makers.

It is not necessary to know very much about the underlying methods to see that graphs such as Figure 1 are problematic as descriptive devices. First, the superposition of the instrumental record (red) creates a strong but entirely misleading contrast. The blue historical reconstruction is necessarily smoother with less overall variation than the red instrumental record since the reconstruction is, in a broad sense, a *weighted average* of *all* global temperature histories conditional on the observed proxy record. Second, the blue curve closely matches the red curve during the period 1902 AD to 1980

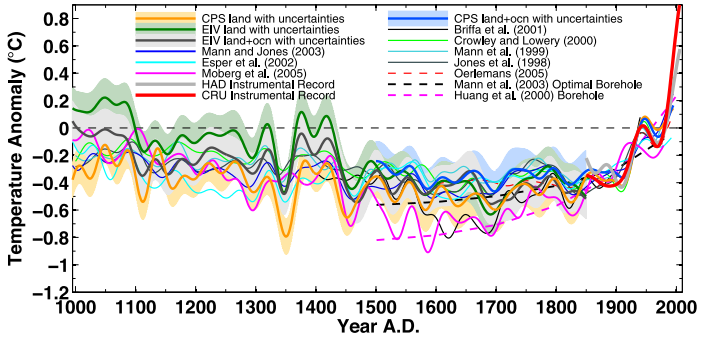

FIG. 2. *Various reconstructions of Northern Hemisphere temperatures over the last* 1000 *years with* 95% *confidence intervals. Source: Mann et al. (2008).*

AD because this period has served as the training data and therefore the blue curve is calibrated to the red during it (note also the red curve is plotted from 1902 AD to 1998 AD). This sets up the erroneous visual expectation that the reconstructions are more accurate than they really are. A careful viewer would know to temper such expectations by paying close attention to the reconstruction error bars given by the wide gray regions. However, even these are misleading because these are, in fact, *pointwise* confidence intervals and not confidence curves for the entire *sample path* of surface

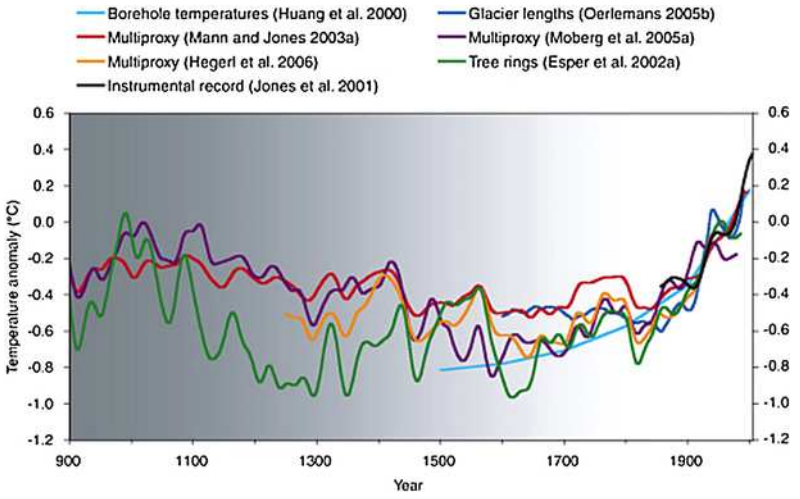

FIG. 3. *Smoothed reconstructions of large-scale (Northern Hemisphere mean or global mean) surface temperature variations from six different research teams are shown along with the instrumental record of global mean surface temperature. Each curve portrays a somewhat different history of temperature variations and is subject to a somewhat different set of uncertainties that generally increase going backward in time (as indicated by the gray shading). Source: NRC (2006).*

temperature. Furthermore, the gray regions themselves fail to account for model uncertainty.

2. Controversy. With so much at stake both financially and ecologically, it is not surprising that these analyses have provoked several controversies. While some have recently erupted in the popular press [Jolis (2009), Johnson (2009), Johnson and Naik (2009)], we root our discussion of these controversies and their history as they unfolded in the academic and scientific literature.

The first major controversy erupted when McIntyre and McKitrick (M&M) successfully replicated the Mann, Bradley and Hughes (1998) study [McIntyre and McKitrick (2003, 2005a, 2005b)]. M&M observed that the original Mann, Bradley and Hughes (1998) study (i) used only one principal component of the proxy record and (ii) calculated the principal components in a "skew"-centered fashion such that they were centered by the mean of the proxy data over the instrumental period (instead of the more standard technique of centering by the mean of the entire data record). Given that the proxy series is itself auto-correlated, this scaling has the effect of producing a first principal component which is hockey-stick shaped [McIntyre and McKitrick (2003)] and, thus, hockey-stick shaped temperature reconstructions. That is, the very *method* used in Mann, Bradley and Hughes (1998) guarantees the shape of Figure 1. M&M made a further contribution by applying the Mann, Bradley and Hughes (1998) reconstruction methodology to principal components computed in the standard fashion. The resulting reconstruction showed a rise in temperature in the medieval period, thus eliminating the hockey stick shape.

Mann and his colleagues vigorously responded to M&M to justify the hockey stick [Mann, Bradley and Hughes (2004)]. They argued that one should not limit oneself to a single principal component as in Mann, Bradley and Hughes (1998), but, rather, one should select the number of principal components retained through cross-validation on *two blocks* of heldout instrumental temperature records (i.e., the first 50 years of the instrumental period and the last 50 years). When this procedure is followed, four principal components are retained, and the hockey stick re-emerges even when the PCs are calculated in the standard fashion. Since the hockey stick is the shape selected by validation, climate scientists argue it is therefore the correct one.[2]

The furor reached such a level that Congress took up the matter in 2006.

[2]Climate scientists call such reconstructions "more skilled." Statisticians would say they have lower out-of-sample root mean square error. We take up this subject in detail in Section 3.

The Chairman of the Committee on Energy and Commerce and that of the Subcommittee on Oversight and Investigations formed an *ad hoc* committee of statisticians to review the findings of M&M. Their Congressional report [Wegman, Scott and Said (2006)] confirmed M&M's finding regarding skew-centered principal components (this finding was yet again confirmed by the National Research Council [NRC (2006)]).

In his Congressional testimony [Wegman (2006)], committee chairman Edward Wegman excoriated Mann, Bradley and Hughes (2004) for use of additional principal components beyond the first after it was shown that their method led to spurious results:

> In the MBH original, the hockey stick emerged in PC1 from the bristlecone/foxtail pines. If one centers the data properly the hockey stick does not emerge until PC4. Thus, a substantial change in strategy is required in the MBH reconstruction in order to achieve the hockey stick, a strategy which was specifically eschewed in MBH... a cardinal rule of statistical inference is that the method of analysis must be decided before looking at the data. The rules and strategy of analysis cannot be changed in order to obtain the desired result. Such a strategy carries no statistical integrity and cannot be used as a basis for drawing sound inferential conclusions.

Michael Mann, in his rebuttal testimony before Congress, admitted to having made some questionable choices in his early work. But, he strongly asserted that none of these earlier problems are still relevant because his original findings have been confirmed again and again in subsequent peer reviewed literature by large numbers of highly qualified climate scientists using vastly expanded data records [e.g., Mann and Rutherford (2002), Luterbacher et al. (2004), Mann et al. (2005, 2007, 2008), Rutherford et al. (2005), Wahl and Amman (2006), Wahl, Ritson and Amman (2006), Li, Nychka and Amman (2007)] even if criticisms do exist [e.g., von Storch et al. (2004)].

The degree of controversy associated with this endeavor can perhaps be better understood by recalling Wegman's assertion that there are very few mainstream statisticians working on climate reconstructions [Wegman, Scott and Said (2006)]. This is particularly surprising not only because the task is highly statistical but also because it is extremely difficult. The data is spatially and temporally autocorrelated. It is massively incomplete. It is not easily or accurately modeled by simple autoregressive processes. The signal is very weak and the number of covariates greatly outnumbers the number of independent observations of instrumental temperature. Much of the analysis in this paper explores some of the difficulties associated with model selection and prediction in just such contexts. We are not interested at this stage in engaging the issues of data quality. To wit, henceforth and

for the remainder of the paper, we work entirely with the data from Mann et al. (2008).[3]

This is by far the most comprehensive publicly available database of temperatures and proxies collected to date. It contains 1209 climate proxies (with some going back as far as 8855 BC and some continuing up till 2003 AD). It also contains a database of eight global annual temperature aggregates dating 1850–2006 AD (expressed as deviations or "anomalies" from the 1961–1990 AD average[4]). Finally, there is a database of 1732 local annual temperatures dating 1850–2006 AD (also expressed as anomalies from the 1961–1990 AD average).[5] All three of these datasets have been substantially processed including smoothing and imputation of missing data [Mann et al. (2008)]. While these present interesting problems, they are not the focus of our inquiry. We *assume* that the data selection, collection, and processing performed by climate scientists meets the standards of their discipline. Without taking a position on these data quality issues, we thus take the dataset as given. We further make the assumptions of linearity and stationarity of the *relationship* between temperature and proxies, an assumption employed throughout the climate science literature [NRC (2006)] noting that "the stationarity of the relationship does not require stationarity of the series themselves" [NRC (2006)]. Even with these substantial assumptions, the paleoclimatological reconstructive endeavor is a very difficult one and we focus on the substantive *modeling* problems encountered in this setting.

Our paper structure and major results are as follows. We first discuss the strength of the proxy signal in this $p \gg n$ context (i.e., when the number of covariates or parameters, p, is much larger than the number of datapoints, n) by comparing the performance, in terms of holdout RMSE, of the proxies against several alternatives. Such an exercise is important because, when $p \gg n$, there is a sizeable risk of overfitting and in-sample performance is often a poor benchmark for out-of-sample performance. We will show that the proxy record easily does better at predicting out-of-sample global temperature than simple rapidly-mixing stationary processes generated independently of

[3]In the sequel, we provide a link to *The Annals of Applied Statistics* archive which hosts the data and code we used for this paper. The Mann et al. (2008) data can be found at http://www.meteo.psu.edu/~mann/supplements/MultiproxyMeans07/. However, we urge caution because this website is periodically updated and therefore may not match the data we used even though at one time it did. For the purposes of this paper, please follow our link to *The Annals of Applied Statistics* archive.

[4]For details, see http://www.cru.uea.ac.uk/cru/data/temperature/.

[5]The Mann et al. (2008) original begins with the HadCRUT3v local temperature data given in the previous link. Temperatures are given on a five degree longitude by five degree latitude grid. This would imply 2592 cells in the global grid. Mann et al. (2008) disqualified 860 such cells because they contained less than 10% of the annual data thus leaving 1732.

the true temperature record. On the other hand, the proxies do not fare so well when compared to predictions made by more complex processes also generated independently of any climate signal. That is, randomly generated sequences are as "predictive" of holdout temperatures as the proxies.

Next, we show that various models for predicting temperature can perform similarly in terms of cross-validated out-of-sample RMSE *but* have very different historical temperature backcasts. Some of these backcasts look like hockey sticks while others do not. Thus, cross-validation is inadequate on its own for model and backcast selection.

Finally, we construct and fit a full probability model for the relationship between the 1000-year-old proxy database and Northern Hemisphere average temperature, providing appropriate *pathwise* standard errors which account for parameter uncertainty. While our model offers support to the conclusion that the 1990s were the warmest decade of the last millennium, it does not predict temperature as well as expected even in-sample. The model does much worse on contiguous 30-year holdout blocks. Thus, we remark in conclusion that natural proxies are severely limited in their ability to predict average temperatures and temperature gradients.

All data and code used in this paper are provided in the supplementary materials [McShane and Wyner (2011)].

3. Model evaluation.

3.1. *Introduction.* A critical difficulty for paleoclimatological reconstruction is that the temperature signal in the proxy record is surprisingly weak. That is, very few, if any, of the individual natural proxies, at least those that are uncontaminated by the documentary record, are able to explain an appreciable amount of the *annual* variation in the local instrumental temperature records. Nevertheless, the proxy record is quite large, creating an additional challenge: there are many more proxies than there are years in the instrumental temperature record. In this setting, it is easy for a model to overfit the comparatively short instrumental record and therefore model evaluation is especially important. Thus, the main goals of this section are twofold. First, we endeavor to judge regression-based methods for the specific task of predicting blocks of temperatures in the instrumental period. Second, we study specifically how the determination of statistical significance varies under different specifications of the null distribution.

Because the number of proxies is much greater than the number of years for which we have temperature data, it is unavoidable that some type of dimensionality reduction is necessary even if there is no principled way to achieve this. As mentioned above, early studies [Mann, Bradley and Hughes (1998, 1999)] used principal components analysis for this purpose. Alternatively, the number of proxies can be lowered through a threshold screening

process [Mann et al. (2008)] whereby each proxy sequence is correlated with its closest local temperature series and only those proxies whose correlation exceeds a given threshold are retained for model building. This is a reasonable approach, but, for it to offer serious protection from overfitting the temperature sequence, it is necessary to detect "spurious correlations."

The problem of spurious correlation arises when one takes the correlation of two series which are themselves highly autocorrelated and is well studied in the time series and econometrics literature [Yule (1926), Granger and Newbold (1974), Phillips (1986)]. When two independent time series are nonstationary (e.g., random walk), locally nonstationary (e.g., regime switching), or strongly autocorrelated, then the distribution of the empirical correlation coefficient is surprisingly variable and is frequently large in absolute value (see Figure 4). Furthermore, standard model statistics (e.g., t-statistics) are inaccurate and can only be corrected when the underlying stochastic processes are both known and modeled (and this can only be done for special cases).

As can be seen in Figures 5 and 6, both the instrumental temperature record as well as many of the proxy sequences are not appropriately modeled by low order stationary autoregressive processes. The dependence structure in the data is clearly complex and quite evident from the graphs. More quantitatively, we observe that the sample first-order autocorrelation of the CRU Northern Hemisphere annual mean land temperature series is nearly 0.6 (with significant *partial* autocorrelations out to lag four). Among the proxy sequences, a full one-third have empirical lag one autocorrelations of at least 0.5 (see Figure 7). Thus, standard correlation coefficient test statistics are not reliable measures of significance for screening proxies against local or global temperatures series. A final more subtle and salient concern is that,

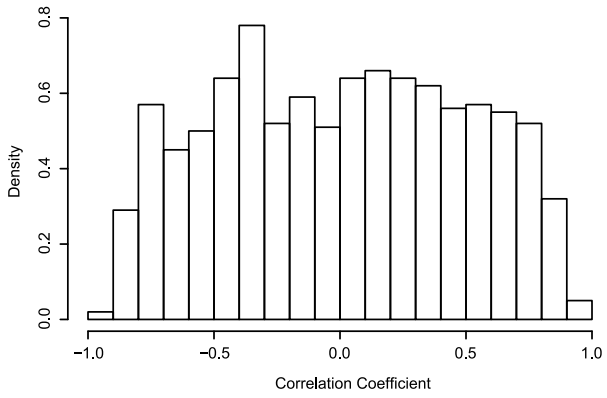

FIG. 4. *Simulated sample correlation coefficient distribution of two independent random walks. One thousand independent pairs of random walks each of length* 149 *were sampled to generate the above histogram.*

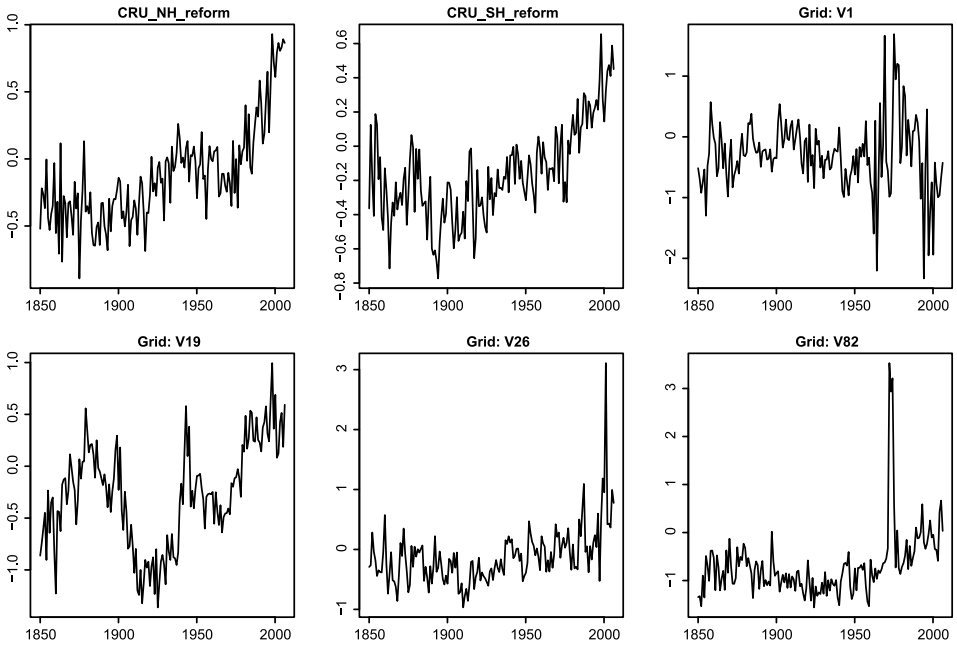

FIG. 5. *CRU Northern Hemisphere annual mean land temperature, CRU Southern Hemisphere annual mean land temperature, and four local temperatures the grids of which contain* (i) *Spitsbergen island in the Svalbard archipelago in the Artic,* (ii) *the north portion of the Omsk oblast in southwestern Siberia,* (iii) *Attu Island, the westernmost island in the Aleutian islands arcipelago, and* (iv) *Baysuat in the Aktobe Province, Kazakhstan. The x-axis gives the year and the y-axis gives the temperature anomaly from 1961–1990 AD average in degrees Celsius.*

if the screening process involves the entire instrumental temperature record, it corrupts the model validation process: no subsequence of the temperature series can be truly considered out-of-sample.

To solve the problem of spurious correlation, climate scientists have used the technique of out-of-sample validation on a reserved holdout block of data. The performance of any given reconstruction can then be benchmarked and compared to the performance of various null models. This will be our approach as well. However, we extend their validation exercises by (i) expanding the class of null models and (ii) considering interpolated holdout blocks as well as extrapolated ones.

3.2. *Preliminary evaluation.* In this subsection, we discuss our validation scheme and compare the predictive performance of the proxies against two simple models which use only temperature itself for forecasting, the in-sample mean and ARMA models. We use as our response y_t the CRU Northern Hemisphere annual mean land temperature. $X = \{x_{tj}\}$ is a cen-

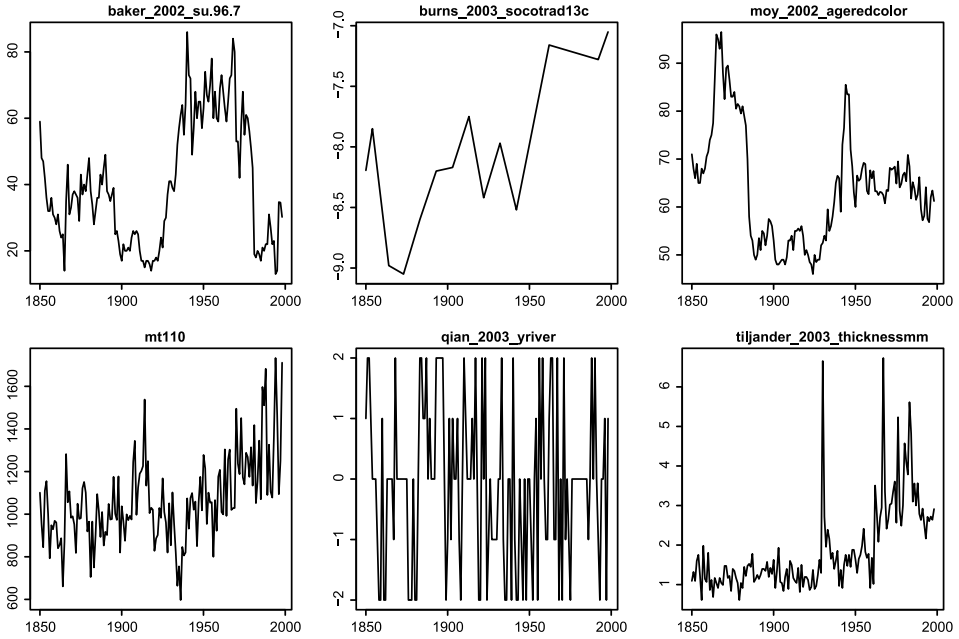

FIG. 6. *Six proxy time series plotted during the instrumental period: speleothems in Scotland, monsoons in India, lake sediment in Ecuador; tree rings in Montana, dry/wet variation on the Yellow River, and lake sediments in Finland.*

tered and scaled matrix of 1138 of the 1209 proxies, excluding the 71 Lutannt series found in Luterbacher et al. (2004).[6] We use the years 1850–1998 AD for these tests because very few proxies are available after 1998 AD.[7]

To assess the strength of the relationship between the natural proxies and temperature, we cross-validate the data. This is a standard approach, but our situation is atypical since the temperature sequence is highly autocorrelated. To mitigate this problem, we follow the approach of climate scientists in our *initial* approach and fit the instrumental temperature record using *only* proxy covariates. Nevertheless, the errors and the proxies are temporally correlated which implies that the usual method of selecting random holdout sets will not provide an effective evaluation of our model. Climate scientists have instead applied "block" validation, holding out two contiguous blocks of instrumental temperatures: a "front" block consisting of the

[6]These Lutannt "proxies" are actually reconstructions calibrated to local temperatures in Europe and thus are not true natural proxies. The proxy database may contain other nonnatural proxies though we do not believe it does. The qualitative conclusions reached in this section hold up, however, even when all 1209 proxies are used.

[7]Only 103 of the 1209 proxies are available in 1999 AD, 90 in 2000 AD, eight in 2001 AD, five in 2002 AD, and three in 2003 AD.

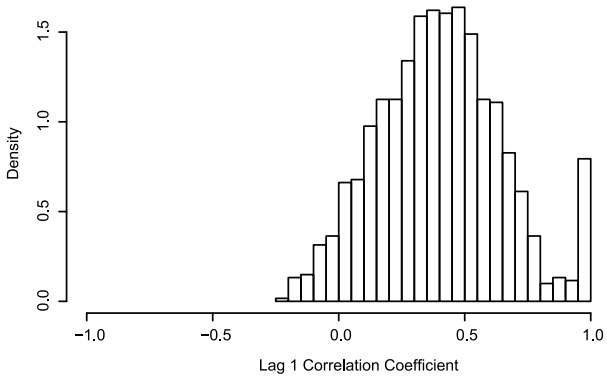

FIG. 7. *Sample lag one autocorrelation coefficient for the 1209 proxies during the instrumental period.*

first 50 years of the instrumental record and a "back" block consisting of the last 50 years.

On the one hand, this approach makes sense since our ultimate task is to extrapolate our data backward in time and only the first and last blocks can be used for this purpose specifically. On the other hand, limiting the validation exercise to these two blocks is problematic because both blocks have very dramatic and obvious features: the temperatures in the initial block are fairly constant and are the coldest in the instrumental record, whereas the temperatures in the final block are rapidly increasing and are the warmest in the instrumental record. Thus, validation conducted on these two blocks will *prima facie* favor procedures which project the local level and gradient of the temperature near the boundary of the in-sample period. However, while such procedures perform well on the front and back blocks, they are not as competitive on interior blocks. Furthermore, they cannot be used for plausible historical reconstructions! A final serious problem with validating on only the front and back blocks is that the extreme characteristics of these blocks are widely known; it can only be speculated as to what extent the collection, scaling, and processing of the proxy data as well as modeling choices have been affected by this knowledge.

Our approach is to consecutively select all possible contiguous blocks for holding out. For example, we take a given contiguous 30-year block from the 149-year instrumental temperature record (e.g., 1900–1929 AD) and hold it out. Using only the remaining 119 years (e.g, 1850–1899 AD and 1930–1998 AD), we tune and fit our model. Finally, we then use the fitted model to obtain predictions for each of the 30 years in the holdout block and then calculate the RMSE on this block.

We then repeat the procedure outlined in the previous paragraph over all 120 possible contiguous holdout blocks in order to approximate the

distribution of the holdout RMSE that is expected using this procedure.[8] We note this test only gives a sense of the ability of the proxies to predict the *instrumental temperature record* and it says little about the ability of the proxies to predict temperature several hundred or thousand years back. Climate scientists have argued, however, that this long-term extrapolation is scientifically legitimate [Mann, Bradley and Hughes (1998), NRC (2006)].

Throughout this section, we assess the strength of the proxy signal by building models for temperature using the Lasso [Tibshirani (1996)]. The Lasso is a penalized least squares method which selects

$$\hat{\beta}^{\text{Lasso}} = \arg\min_{\beta} \left\{ \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j \right)^2 + \lambda \sum_{i=1}^{p} |\beta_i| \right\}.$$

As can be seen, the intercept β_0 is not penalized. Typically (and in this paper), the matrix of predictors X is centered and scaled, and λ is chosen by cross-validation. Due to the L_1 penalty, the Lasso tends to choose sparse $\hat{\beta}^{\text{Lasso}}$, thus serving as a variable selection methodology and alleviating the $p \gg n$ problem. Furthermore, since the Lasso tends to select only a few of a set of correlated predictors, it also helps reduce the problem of spatial correlation among the proxies.

We select the Lasso tuning parameter λ by performing ten repetitions of five-fold cross-validation on the 119 in-sample years and choosing the value $\lambda = \hat{\lambda}$ which provides the best RMSE. We then fit the Lasso to the full 119-year in-sample dataset using $\lambda = \hat{\lambda}$ to obtain $\hat{\beta}^{\text{Lasso}}$. Finally, we can use $\hat{\beta}^{\text{Lasso}}$ to obtain predictions for each of the 30 years in the holdout block and then calculate the RMSE on this block.

We chose the Lasso because it is a reasonable procedure that has proven powerful, fast, and popular, and it performs comparably well in a $p \gg n$ context. Thus, we believe it should provide predictions which are as good or better than other methods that we have tried (evidence for this is presented in Figure 12). Furthermore, we are as much interested in how the proxies fare as predictors when varying the holdout block and null distribution (see Sections 3.3 and 3.4) as we are in performance. In fact, all analyses in this

[8] We performed two variations of this procedure. In the first variation, we continued to hold out 30 years; however, we calculated the RMSE for only the middle 20 years of the 30-year holdout block, leaving out the first five and last five years of each block in order to reduce the correlation between holdout blocks. In the second variation, we repeated this procedure using 60-year holdout blocks. In both cases, all qualitative conclusions remained the same. Considering smaller holdout blocks such as 15 years could be an interesting extension. However, over such short intervals, the global temperature series itself provides substantial signal even without the use of proxies. Furthermore, given the small size of the dataset and lack of independence between 15-, 30-, and 60-year holdout blocks, this might raise concerns about overfitting and over-interpreting the data.

section have been repeated using modeling procedures other than the Lasso and qualitatively all results remain more or less the same.

As an initial test, we compare the holdout RMSE using the proxies to two simple models which only make use of temperature data, the in-sample mean and ARMA models. First, the proxy model and the in-sample mean seem to perform fairly similarly, with the proxy-based model beating the sample mean on only 57% of holdout blocks. A possible reason the sample mean performs comparably well is that the instrumental temperature record has a great deal of annual variation which is apparently uncaptured by the proxy record. In such settings, a biased low-variance predictor (such as the in-sample mean) can often have a lower out-of-sample RMSE than a less biased but more variable predictor. Finally, we observe that the performance on different validation blocks are not independent, an issue which we return to in Section 3.4.

We also compared the holdout RMSE of the proxies to another more sophisticated model which, like the in-sample mean, only makes use of temperature data and makes no reference to proxy data. For each holdout block, we fit various ARMA(p, q) models; we let p and q range from zero to five and chose the values which give the best AIC. We then use this model to forecast the temperature on the holdout block. This model beats the proxy model 86% of the time.

In Figure 8, we focus on one particular holdout block, the last 30 years of the series.[9] The in-sample mean and the ARMA model completely miss the rising trend of the last 30 years; in fact, both models are essentially useless for backcasting and forecasting since their long-term prediction is equal to the in-sample mean. On the other hand, the record of 1138 proxies does appear to capture the rising trend in temperatures (in the sequel, we will assess the statistical significance of this). Furthermore, the differences in temperature and the differences in the proxy forecast are significantly correlated ($p = 0.021$), with the same sign in 21 out of the 29 years ($p = 0.026$).

3.3. *Validation against pseudo-proxies.* Because both the in-sample mean and the ARMA model always forecast the mean in the long-term, they are not particularly useful models for the scientific endeavor of temperature reconstruction. Furthermore, the fact that the Lasso-selected linear combination of the proxies beats the in-sample mean on 57% of holdout blocks and the ARMA model on 14% of holdout blocks is difficult to interpret without solid benchmarks of performance.

[9] In this and all subsequent figures, smooths are created by using the `loess` function in R with the span set to 0.33.

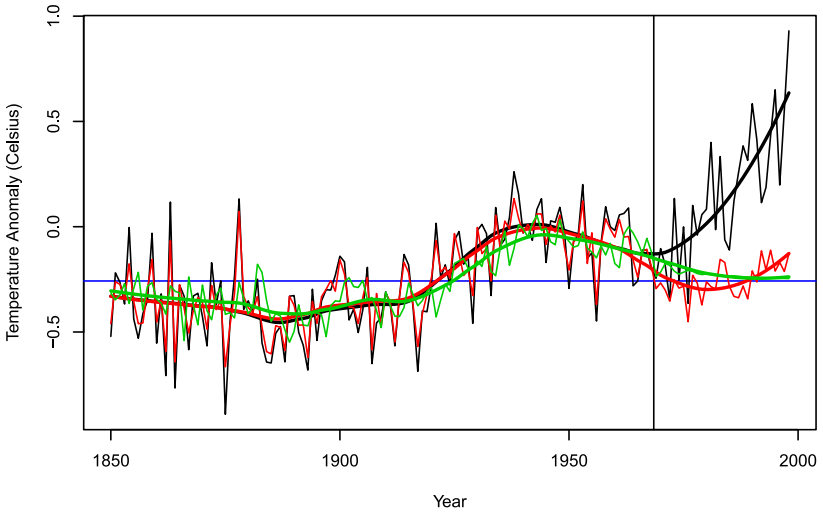

FIG. 8. *CRU Northern Hemisphere annual mean land temperature is given by the thin black line and a smoothed version is given by the thick black line. The forecast produced by applying the Lasso to the proxies is given by the thin red line and a smoothed version is given by the thick red line. The in-sample mean is given by the horizontal blue line. The forecast produced by ARMA modeling is given by the thin green line and a smoothed version is given by the thick green line. The Lasso and ARMA models and the mean are fit on 1850–1968 AD and forecast on 1969–1998 AD.*

One way to provide benchmarks is to repeat the Lasso procedure outlined above using 1138 "pseudo-proxies" in lieu of the 1138 real proxies. That is, replace the natural proxies of temperature by an alternate set of time series. Any function of the proxies, with their resultant temperature reconstruction, can be validated by comparing the ability of the proxies to predict out-of-sample instrumental temperatures to the ability of the pseudo-proxies.

The use of pseudo-proxies is quite common in the climate science literature where pseudo-proxies are often built by adding an AR1 time series ("red noise") to natural proxies, local temperatures, or simulated temperatures generated from General Circulation Models [Mann and Rutherford (2002), Wahl and Amman (2006)]. These pseudo-proxies determine whether a given reconstruction is "skillful" (i.e., statistically significant). Skill is demonstrated with respect to a class of pseudo-proxies when the true proxies outperform the pseudo-proxies with high probability (probabilities are approximated by simulation). In our study, we use an even *weaker* benchmark than those in the climate science literature: our pseudo-proxies are random numbers generated *completely independently* of the temperature series.

The simplest class of pseudo-proxies we consider are Gaussian White Noise. That is, we apply the Lasso procedure outlined above to a 149×1138 matrix of standard normal random variables. Formally, let $\varepsilon_t \overset{\text{i.i.d.}}{\sim} N(0,1), t =$

$1, 2, \ldots$. Then, our White Noise pseudo-proxies are defined as $X_t \equiv \varepsilon_t$ and we generate 1138 such series, each of length 149.

We also consider three classes of AR1 or "red noise" pseudo-proxies since they are common in the climate literature [Mann, Bradley and Hughes (1998), von Storch et al. (2004), Mann et al. (2008)]. Again, if $\varepsilon_t \overset{\text{i.i.d.}}{\sim} N(0,1)$, then an AR1 pseudo-proxy is defined as $X_t \equiv \phi X_{t-1} + \varepsilon_t$. Two of the classes are AR1 with the ϕ coefficient set in turn to 0.25 and 0.4 [these are the average sample proxy autocorrelations reported in Mann, Bradley and Hughes (1998) and Mann et al. (2008), resp.]. The third class is more complicated. First, we fit an AR1 model to each of the 1138 proxies and calculate the sample AR1 coefficients $\hat{\phi}_1, \ldots, \hat{\phi}_{1138}$. Then, we generate an AR1 series setting $\phi = \hat{\phi}_i$ for each of these 1138 estimated coefficients. We term this the empirical AR1 process. This approach is similar to that of McIntyre and McKitrick (2005a, 2005c) who use the full empirical autocorrelation function to generate trend-less pseudo-proxies.

We also consider Brownian motion pseudo-proxies formed by taking the cumulative sums of $N(0,1)$ random variables. That is, if $\varepsilon_t \overset{\text{i.i.d.}}{\sim} N(0,1)$, then a Brownian motion pseudo-proxy is defined as $X_t \equiv \sum_{j=1}^{t} \varepsilon_j = X_{t-1} + \varepsilon_t$.

White Noise and Brownian motion can be thought of as special cases of AR1 pseudo-proxies. In fact, they are the *extrema* of AR1 processes: White Noise is AR1 with the ϕ coefficient set to zero and Brownian motion is AR1 with the ϕ coefficient set to one.

Before discussing the results of these simulations, it is worth emphasizing why this exercise is necessary. That is, why can't one evaluate the model using standard regression diagnostics (e.g., F-statistics, t-statistics, etc.)? One cannot because of two problems mentioned above: (i) the $p \gg n$ problem and (ii) the fact that proxy and temperature autocorrelation causes spurious correlation and therefore invalid model statistics (e.g., t-statistics). The first problem has to be dealt with via dimensionality reduction; the second can only be solved when the underlying processes are known (and then only in special cases).

Given that we do not know the true underlying dynamics, the nonparametric, prediction-based approach used here is valuable. We provide a variety of benchmark pseudo-proxy series and obtain holdout RMSE distributions. Since these pseudo-proxies are generated independently of the temperature series, we know they cannot be truly predictive of it. Hence, the real proxies—if they contain linear signal on temperatures—should outperform our pseudo-proxies, at least with high probability.

For any given class of pseudo-proxy, we can estimate the probability that a randomly generated pseudo-proxy sequence outperforms the true proxy record for predicting temperatures in a given holdout block. A major focus of our investigation is the sensitivity of this outperformance "p-value" to

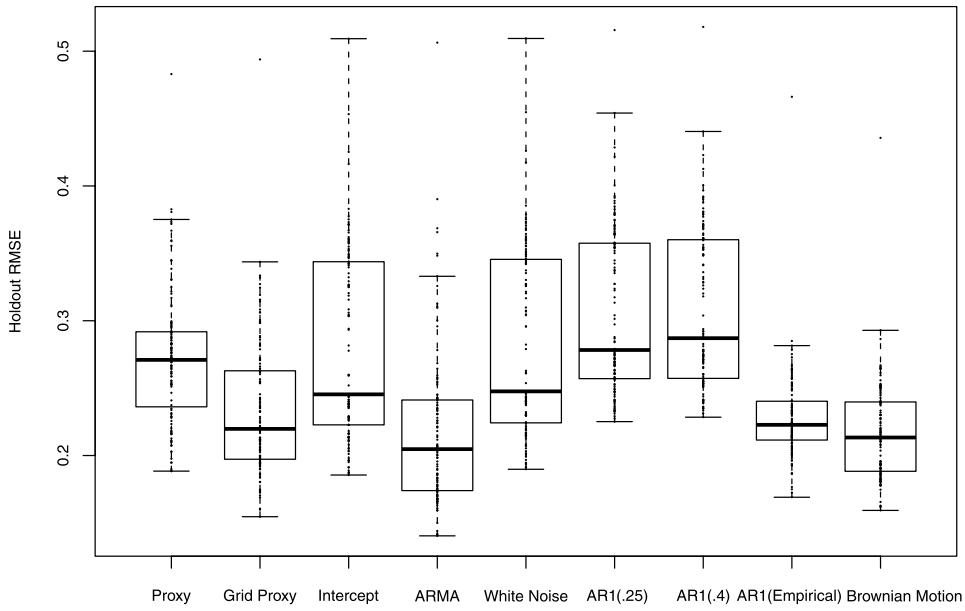

FIG. 9. *Cross-validated RMSE on* 30*-year holdout blocks for various models fit to proxies and pseudo-proxies. The procedures used to generate the Proxy, Intercept, and ARMA boxplots are discussed in Section 3.2. The procedures used to generate the White Noise, AR1, and Brownian motion boxplots are discussed in Section 3.3. The procedure used to generate the Grid Proxy boxplot is discussed in Section 3.6.*

various factors. We proceed in two directions. We first consider the level and variability in holdout RMSE for our various classes of pseudo-proxies marginally over all 120 holdout blocks. Second, since these 120 holdout blocks are highly correlated with one another, we study how the holdout RMSE varies from block to block in Section 3.4.

We present our results in Figure 9, with the RMSE boxplot for the proxies given first. As can be seen, the proxies have a slightly worse median RMSE than the intercept-only model (i.e., the in-sample mean) but the distribution is narrower. On the other hand, the ARMA model is superior to both. When the Lasso is used on White Noise pseudo-proxies, the performance is similar to the intercept-only model because the Lasso is choosing a very parsimonious model.

The proxies seem to outperform the AR1(0.25) and AR1(0.4) models, with both a better median and a lower variance. While this is encouraging, it is also raises a concern: AR1(0.25) and AR1(0.4) are the models frequently used as "null benchmarks" in the climate science literature and they seem to perform worse than both the intercept-only and White Noise benchmarks. This suggests that climate scientists are using a particularly weak null benchmark to test their models. That the null models may be too

weak and the associated standard errors in papers such as Mann, Bradley and Hughes (1998) are not wide enough has already been pointed out in the climate literature [von Storch et al. (2004)]. While there was some controversy surrounding the result of this paper [Wahl, Ritson and Amman (2006)], its conclusions have been corroborated [von Storch and Zorita (2005), von Storch et al. (2006), Lee, Zwiers and Tsao (2008), Christiansen, Schmith and Thejll (2009)].

Finally, the empirical AR1 process and Brownian motion both substantially outperform the proxies. They each have a lower average holdout RMSE and lower variability than that achieved by the proxies. This is extremely important since these two classes of time series are generated *completely independently* of the temperature data. They have *no* long term predictive ability, and they cannot be used to reconstruct historical temperatures. Yet, they significantly outperform the proxies at 30-year holdout prediction!

In other words, our model performs better when using highly autocorrelated noise rather than proxies to "predict" temperature. The real proxies are less predictive than our "fake" data. While the Lasso-generated reconstructions using the proxies are highly statistically significant compared to simple null models, they do not achieve statistical significance against sophisticated null models.

We are not the first to observe this effect. It was shown, in McIntyre and McKitrick (2005a, 2005c), that random sequences with complex local dependence structures can predict temperatures. Their approach has been roundly dismissed in the climate science literature:

> To generate "random" noise series, MM05c apply the full autoregressive structure of the real world proxy series. In this way, they in fact train their stochastic engine with significant (if not dominant) low frequency *climate signal* rather than purely nonclimatic noise and its persistence. [Emphasis in original]
>
> Ammann and Wahl (2007)

Broadly, there are two components to any climate signal. The first component is the local time dependence made manifest by the strong autocorrelation structure observed in the temperature series itself. It is easily observed that short term future temperatures can be predicted by estimates of the local mean and its first derivatives [Green, Armstrong and Soon (2009)]. Hence, a procedure that fits sequences with complex local dependencies to the instrumental temperature record will recover the ability of the temperature record to self-predict in the short run.

The second component—long-term changes in the temperature series—can, on the other hand, only be predicted by meaningful covariates. The autocorrelation structure of the temperature series does not allow for self-prediction in the long run. Thus, pseudo-proxies like ours, which inherit their ability at short-term prediction by borrowing the dependence structure of

the instrumental temperature series, have no more power to reconstruct temperature than the instrumental record itself (which is entirely sensible since these pseudo-proxies are generated independently of the temperature series).

Ammann and Wahl (2007) claim that significance thresholds set by Monte Carlo simulations that use pseudo-proxies containing "short term climate signal" (i.e., complex time dependence structures) are invalid:

> Such thresholds thus enhance the danger of committing Type II errors (inappropriate failure to reject a null hypothesis of no climatic information for a reconstruction).

We agree that these thresholds decrease power. Still, these thresholds are the correct way to preserve the significance level. The proxy record has to be evaluated in terms of its innate ability to reconstruct historical temperatures (i.e., as opposed to its ability to "mimic" the local time dependence structure of the temperature series). Ammann and Wahl (2007) wrongly attribute reconstructive skill to the proxy record which is in fact attributable to the temperature record itself. Thus, climate scientists are overoptimistic: the 149-year instrumental record has significant local time dependence and therefore far fewer independent degrees of freedom.

3.4. *Interpolation versus extrapolation.* In our analysis, we expanded our set of holdout blocks to include all contiguous 30-year blocks. The benefits of this are twofold. First, this expansion increases our sample size from two (the front and back blocks) to 120 (because there are 118 possible interior blocks). Second, by expanding the set of holdout blocks, we mitigate the potential effects of data snooping since salient characteristics of the first and last blocks are widely known. On the other hand, this expansion imposes difficulties. The RMSEs of overlapping blocks are highly dependent. Furthermore, since temperatures are autocorrelated, the RMSEs of neighboring nonoverlapping blocks are also dependent. Thus, there is little new information in each block.[10] We explore this graphically by plotting the RMSE of each holdout block against the first year of the block in Figure 10.

We begin our discussion by comparing RMSE of the Lasso model fitted to the proxies to RMSE of the in-sample mean and the RMSE of the ARMA model in upper left panel of Figure 10. As can be seen, the ARMA model either dominates or is competitive on every holdout block. The proxies, on the other hand, can match the performance of the ARMA model only on

[10] As noted in a previous footnote, we considered a variation of our procedure where we maintained 30-year holdout blocks but only calculated the RMSE on the middle 20 years of the block, thus reducing the dependence between overlapping and nearby blocks. All qualitative conclusions remained the same.

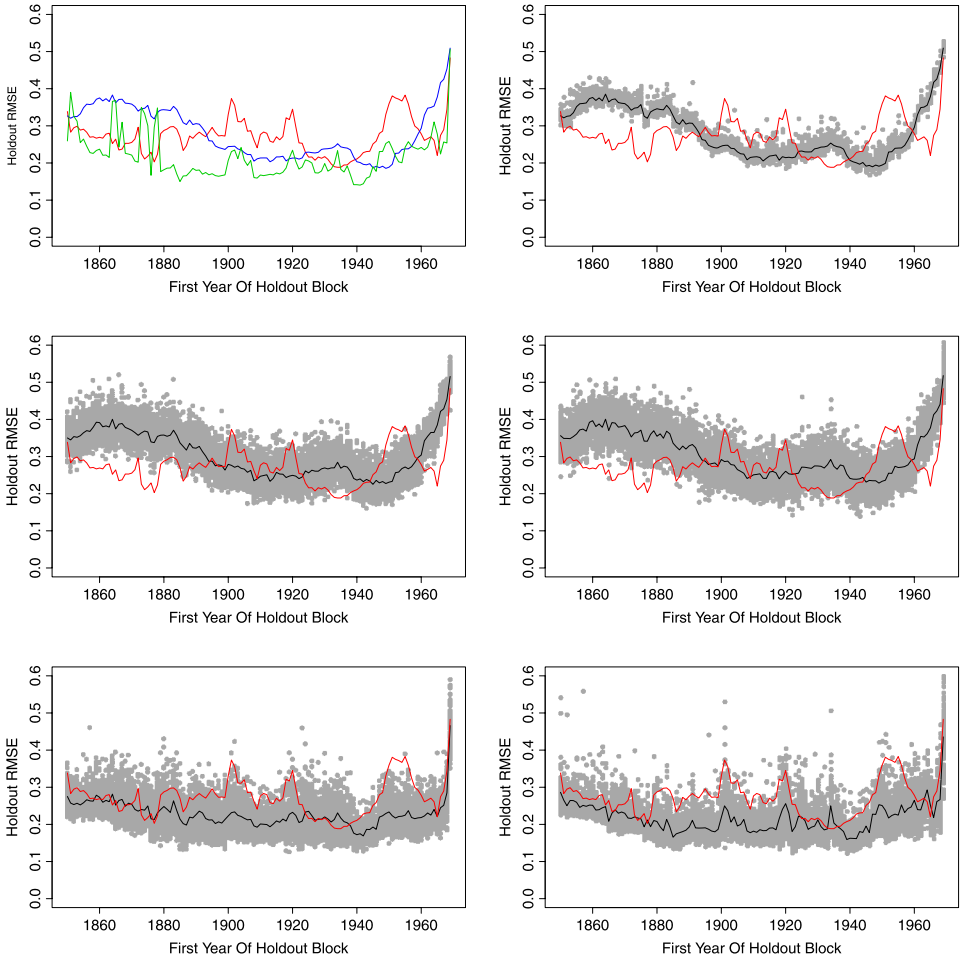

FIG. 10. *Holdout RMSE by first year of holdout block. In all panels, the Lasso-selected linear combination of the proxies is given in red. In the upper-left panel, the in-sample mean is given in blue and the ARMA model in green. In the upper-right panel, the average for the White Noise pseudo-proxy is given in black. In the middle-left panel, the average for the* AR1(0.25) *pseudo-proxy is given in black. In the middle-right panel, the average for the* AR1(0.4) *pseudo-proxy is given in black. In the lower-left panel, the average for the Empirical AR*1 *pseudo-proxy is given in black. In the lower-right panel, the average for the Brownian motion pseudo-proxy is given in black. Confidence intervals for the pseudo-proxies are given in gray and are formed by taking* 100 *samples of the pseudo-proxy matrix for each holdout block.*

the first 20 or so holdout blocks, but on other blocks, they perform quite a bit worse.

More interesting is the examination of the performance of the pseudo-proxies, as shown in the remaining five panels of Figure 10. In these graphs, we compare the RMSE of the proxies on each holdout block to the RMSE

of the pseudo-proxies. We also provide confidence intervals for the pseudo-proxies at each block by simulating 100 draws of the pseudo-proxy matrix and repeating our fitting procedure to each draw. As can be seen in the upper-right panel, the proxies show statistically significant improvement over White Noise for many of the early holdout blocks as well as many of the later ones. However, there are blocks, particularly in the middle, where they perform significantly worse.

When the AR1(0.25) and AR1(0.4) pseudo-proxies preferred by climate scientists are used, the average RMSE on each is comparable to that given by White Noise but the variation is considerably higher as shown by the middle two panels of Figure 10. Hence, the proxies perform statistically significantly better on very few holdout blocks, particularly those near the beginning of the series and those near the end. This is a curious fact because the "front" holdout block and the "back" holdout block are the only two which climate scientists use to validate their models. Insofar as this front and back performance is anomalous, they may be overconfident in their results.

Finally, we consider the AR1 Empirical and Brownian motion pseudo-proxies in the lower two panels of Figure 10. For almost all holdout blocks, these pseudo-proxies have an average RMSE that is as low or lower than that of the proxies. Further, for no block is the performance of true proxies statistically significantly better than that of either of these pseudo-proxies. Hence, we cannot reject the null hypothesis that the true proxies "predict" equivalently to highly correlated and/or nonstationary sequences of random noise that are independent of temperature.

A little reflection is in order. By cross-validating on interior blocks, we are able to greatly expand the validation test set. However, reconstructing interior blocks is an interpolation of the training sequence and paleoclimatological reconstruction requires extrapolation as opposed to interpolation. Pseudo-proxy reconstructions can only extrapolate a climate trend accurately for a very short period and then only insofar as the local dependence structure in the pseudo-proxies matches the local dependence structure in the temperature series. That is, forecasts from randomly generated series can extrapolate successfully only by chance and for very short periods.

On the other hand, Brownian motions and other pseudo-proxies with strong local dependencies are quite suited to interpolation since their in-sample forecasts are fitted to approximately match the the training sequence datapoints that are adjacent to the initial and final points of a test block. Nevertheless, true proxies also have strong local dependence structure since they are temperature surrogates and therefore should similarly match these datapoints of the training sequence. Furthermore, unlike pseudo-proxies, true proxies are *not* independent of temperature (in fact, the scientific presumption is that they are *predictive* of it). Therefore, proxy interpolations on interior holdout blocks should be expected to outperform pseudo-proxy forecasts notwithstanding the above.

TABLE 1
Percent of pseudo-proxies selected by the Lasso

Pseudo-proxy	**Percent selected**
White Noise	37.8%
AR1(0.25)	43.5%
AR1(0.4)	47.9%
Empirical AR1	53.0%
Brownian Motion	27.9%

3.5. *Variable selection: True proxies versus pseudo-proxies.* While the use of noise variables such as the pseudo-proxies is not unknown in statistics, such variables have typically been used to augment a matrix of covariates rather than to replace it. For example, Wu, Boos and Stefanski (2007) augment a matrix of covariates with noise variables in order to tune variable selection methodologies. Though that is not our focus, we make use of a similar approach in order to assess the the degree of signal in the proxies.

We first augment the in-sample matrix of proxies with a matrix of pseudo-proxies of the same size (i.e., replacing the 119×1138 proxy matrix with a matrix of size 119×2276 which consists of the original proxies plus pseudo-proxies). Then, we repeat the Lasso cross-validation described in Section 3.2, calculate the percent of variables selected by the Lasso which are pseudo-proxies, and average over all 120 possible blocks. If the signal in the proxies dominates that in the pseudo-proxies, then this percent should be relatively close to zero.

Table 1 shows this is far from the case. In general, the pseudo-proxies are selected about as often as the true proxies. That is, the Lasso does not find that the true proxies have substantially more signal than the pseudo-proxies.

3.6. *Proxies and local temperatures.* We performed an additional test which accounts for the fact that proxies are local in nature (e.g., tree rings in Montana) and therefore might be better predictors of local temperatures than global temperatures. Climate scientists generally accept the notion of "teleconnection" (i.e., that proxies local to one place can be predictive of climate in another possibly distant place). Hence, we do not use a distance restriction in this test. Rather, we perform the following procedure.

Again, let y_t be the CRU Northern Hemisphere annual mean land temperature where t indexes each year from 1850–1998 AD, and let $X = \{x_{tj}\}$ be the centered and scaled matrix of 1138 proxies from 1850–1998 AD where t indexes the year and j indexes each proxy. Further, let $Z = \{z_{tj}\}$ to be the matrix of the 1732 centered and scaled local annual temperatures from

1850–1998 AD where again t indexes the year and j indexes each local temperature.

As before, we take a 30-year contiguous block and reserve it as a holdout sample. Our procedure has two steps:

1. Using the 119 in-sample years, we perform ten repetitions of five-fold cross-validation as described in Section 3.2. In this case, however, instead of using the proxies X to predict y, we use the local temperatures Z. As before, this procedure gives us an optimal value for the tuning parameter $\hat{\lambda}$ which we can use on all 119 observations of y and Z to obtain $\hat{\beta}^{\text{Lasso}}$.
2. Now, for each j such that $\hat{\beta}_j^{\text{Lasso}} \neq 0$, we create a Lasso model for $z_{\cdot j}$. That is, we perform ten repetitions of five-fold cross-validation as in Section 3.2 but using X to predict $z_{\cdot j}$. Again, this procedure gives us an optimal value for the tuning parameter $\hat{\lambda}_j$ which we can use on all 119 observations of $z_{\cdot j}$ and X to obtain $\hat{\beta}^{\text{Lasso},(j)}$.

Similarly, we can predict on the holdout block using a two-stage procedure. For each j such that $\hat{\beta}_j^{\text{Lasso}} \neq 0$, we apply $\hat{\beta}^{\text{Lasso},(j)}$ to X to obtain $\hat{z}_{\cdot j}$ in the 30 holdout years. Then, we apply $\hat{\beta}^{\text{Lasso}}$ to the collection of $\hat{z}_{\cdot j}$ in order to obtain $\hat{y}_t$ in the 30 holdout years. Finally, we calculate the RMSE on the holdout block and repeat this procedure over all 120 possible holdout blocks.

As in Section 3.2, this procedure uses the Lasso to mitigate the $p \gg n$ problem. Furthermore, since the Lasso is unlikely to select correlated predictors, it also attenuates the problem of spatial correlation among the local temperatures and proxies. But, this procedure has the advantage of relating proxies to local temperatures, a feature which could be advantageous if these relationships are more conspicuous and enduring than those between proxies and the CRU global average temperature. The same is also potentially true *mutatis mutandis* of the relationship between the local temperatures and CRU.

The results of this test are given by the second boxplot in Figure 9. As can be seen, this method seems to perform somewhat better than the pure global method. However, it does not beat the empirical AR1 process or Brownian motion. That is, random series that are independent of global temperature are as effective or more effective than the proxies at predicting global annual temperatures in the instrumental period. Again, the proxies are not statistically significant when compared to sophisticated null models.

3.7. *Discussion of model evaluation.* We can think of four possible explanatory factors for what we have observed. First, it is possible that the proxies are in fact too weakly connected to global annual temperature to offer a substantially predictive (as well as reconstructive) model over the

majority of the instrumental period. This is not to suggest that proxies are unable to detect large variations in global temperature (such as those that distinguish our current climate from an ice age). Rather, we suggest it is possible that natural proxies cannot reliably detect the small and largely unpredictable changes in annual temperature that have been observed over the majority of the instrumental period. In contrast, we have previously shown that the proxy record has some ability to predict the final 30-year block, where temperatures have increased most significantly, better than chance would suggest.

A second explanation is that the Lasso might be a poor procedure to apply to these data. This seems implausible both because the Lasso has been used successfully in a variety of $p \gg n$ contexts and because we repeated the analyses in this section using modeling strategies other than the Lasso and obtained the same general results. On the other hand, climate scientists have basically used three different statistical approaches: (i) scaling and averaging (so-called "Composite Plus Scale" or CPS) [NRC (2006)], (ii) principal component regression [NRC (2006)], and (iii) "Errors in Variables" (EIV) regression [Schneider (2001), Mann et al. (2007)]. The EIV approach is considered the most reliable and powerful. The approach treats forecasting (or reconstruction) from a missing data perspective using the Expectation–Maximization algorithm to "fill-in" blocks of missing values. The EM core utilizes an EIV generalized linear regression which addresses the $p \gg n$ problem using regularization in the form of a ridge regression-like total sum of squares constraint (this is called "RegEM" in the climate literature [Mann et al. (2007)]). All of these approaches are intrinsically linear, like Lasso regression, although the iterative RegEM can produce nonlinear functions of the covariates. Fundamentally, there are only theoretical performance guarantees for i.i.d. observations, while our data is clearly correlated across time. The EM algorithm in particular lacks a substantive literature on accuracy and performance without specific assumptions on the nature of missing data. Thus, it not obvious why the Lasso regression should be substantively worse than these methods. Nevertheless, in subsequent sections we will study a variety of different and improved model variations to confirm this.

A third explanation is that our class of competitive predictors (i.e., the pseudo-proxies) may very well provide unjustifiably difficult benchmarks as claimed by Ammann and Wahl (2007) and discussed in Section 3.3. Climate scientists have calibrated their performance using either (i) weak AR1 processes of the kind demonstrated above as pseudo-proxies or (ii) by adding weak AR1 processes to local temperatures, other proxies, or the output from global climate simulation models. In fact, we have shown that the the proxy record outperforms the former. On the other hand, weak AR1 processes underperform even white noise! Furthermore, it is hard to argue that a

procedure is truly skillful if it cannot consistently outperform noise, no matter how artfully structured. In fact, Figure 6 reveals that the proxy series contain very complicated and highly autocorrelated time series structures which indicates that our complex pseudo-proxy competitors are not entirely unreasonable.

Finally, perhaps the proxy signal can be enhanced by smoothing various time series before modeling. Smoothing seems to be a standard approach for the analysis of climate series and is accompanied by a large body of literature [Mann (2004, 2008)]. Still, from a statistical perspective, smoothing time series raises additional questions and problems. At the most basic level, one has to figure out which series should be smoothed: temperatures, proxies, or both. Or, perhaps, only the forecasts should be smoothed in order to reduce the forecast variance. A further problem with smoothing procedures is that there are many methods and associated tuning parameters and there are no clear data-independent and hypothesis-independent methods of selecting among the various options. The instrumental temperature record is also very well known so there is no way to do this in a "blind" fashion. Furthermore, smoothing data exacerbates all of the statistical significance issues already present due to autocorrelation: two smoothed series will exhibit artificially high correlations and both standard errors and p-values require corrections (which are again only known under certain restrictive conditions).

4. Testing other predictive methods.

4.1. *Cross-validated RMSE.* In this section, we pursue alternative procedures, including regression approaches more directly similar to techniques used by climate scientists. We shall see, working with a similar dataset, that various fitting methods can have both (i) very similar contiguous 30-year cross-validated instrumental period RMSE distributions *and* (ii) very different historical backcasts.

Again, we use as our response the CRU Northern Hemisphere annual mean land temperature from 1850–1998 AD and augment it with the 1732 local temperature series when required. However, since we are ultimately interested in large-scale reconstructions, we limit ourselves in this section to only those 93 proxies for which we have data going back over 1000 years.[11] Hence, our in-sample dataset consists of the CRU global aggregate, the 1732 local temperatures, and the 93 proxies from 1850–1998 AD and we apply

[11]There are technically 95 proxies dating back this far but three of them (tiljander_2003_darksum, tiljander_2003_lightsum, and tiljander_2003_thicknessmm) are highly correlated with one another. Hence, we omit the latter two. Again, qualitatively, results hold up whether one uses the reduced set of 93 or the full set of 95 proxies. However, using the full set can cause numerical instability issues.

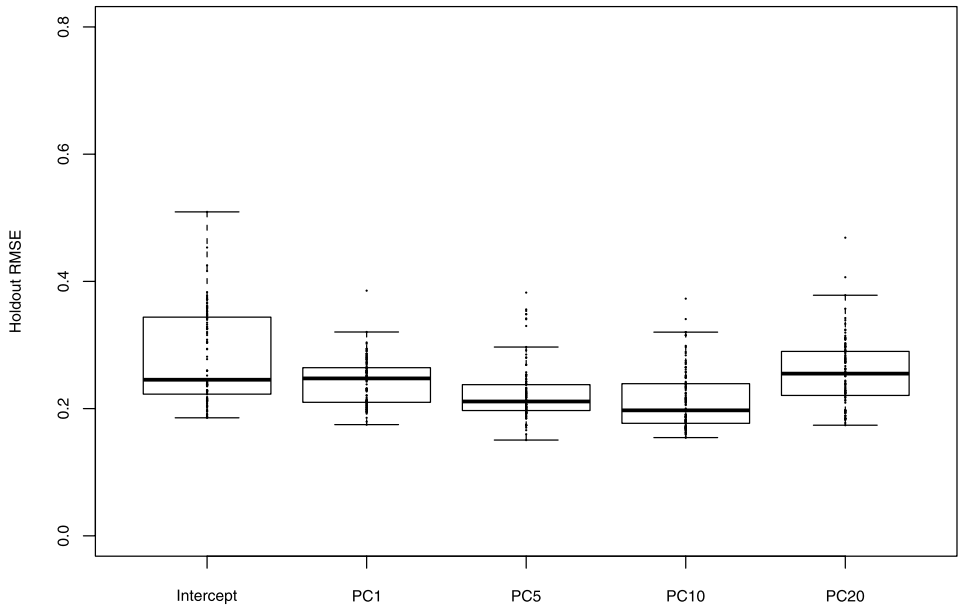

FIG. 11. *Cross-validated RMSE on* 30*-year holdout blocks for various model specifications: intercept only and regression on the first one, five, ten, and* 20 *principal components of the proxies.*

the cross-validation procedure discussed in Section 3.2 to it. We can then examine backcasts on the 998–1849 AD period for which only the proxies are available. We expect that our prediction accuracy during the instrumental period will decay somewhat since our set of proxies is so much smaller. However, the problem of millennial reconstructions is much more interesting both statistically and scientifically. It is well known and generally agreed that the several hundred years before the industrial revolution were a comparatively cool "Little Ice Age" [Matthes (1939), Lamb (1990)]. What happened in the early Medieval period is much more controversial and uncertain [Ladurie (1971), IPCC (2001)].

We now examine how well the proxies predict under alternative model specifications. In the first set of studies, we examine RMSE distributions using an intercept-only model and ordinary least squares regression on the first one, five, ten, and 20 principal components calculated from the full 1001×93 proxy matrix. Our results are shown in Figure 11. As can be seen, all of these methods perform comparably, with five and ten principal component models perhaps performing slightly better than the others.

In a second set of validations, we consider various variable selection methodologies and apply them to both the raw proxies and the principal components of the proxies. The methods considered are the Lasso and stepwise regression designed to optimize AIC and BIC, respectively. We plot our results

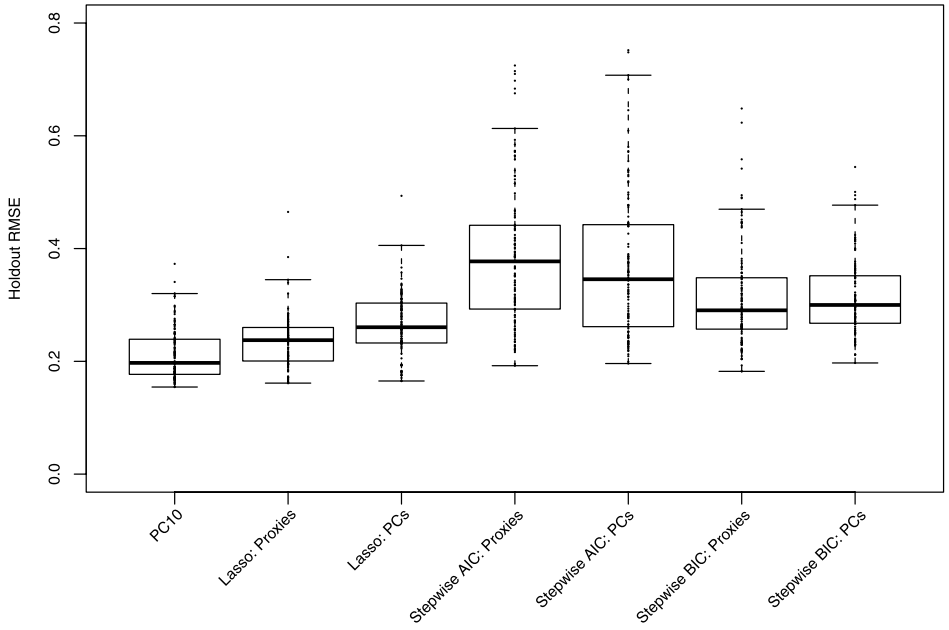

FIG. 12. *Cross-validated RMSE on* 30*-year holdout blocks for various model specifications: regression on the first ten principal components of the proxies, the Lasso applied to the proxies and the principal components of the proxies, stepwise regression to maximize AIC applied to the proxies and the principal components of the proxies, and stepwise regression to maximize BIC applied to the proxies and the principal components of the proxies.*

in Figure 12 and include the boxplot of the ten principal component model from Figure 11 for easy reference. As can be seen, the stepwise models perform fairly similarly with one another. The Lasso performs slightly better and predicts about as well as the ten principal component model.

As a final consideration, we employ a method similar to that used in the original Mann, Bradley and Hughes (1998) paper. This method takes account of the fact that local proxies might be better predictors of local temperatures than they are of global aggregate temperatures. For this method, we again use the first p principal components of the proxy matrix but we also use the first g principal components of the 149×1732 local temperature matrix. We regress the CRU global aggregate on the g principal components of local temperature matrix, and then we regress each of the g local temperature principal components on the p proxy principal components. We can then use the historical proxy principal components to backcast the local temperature principal components thereby enabling us to backcast the global average temperature.

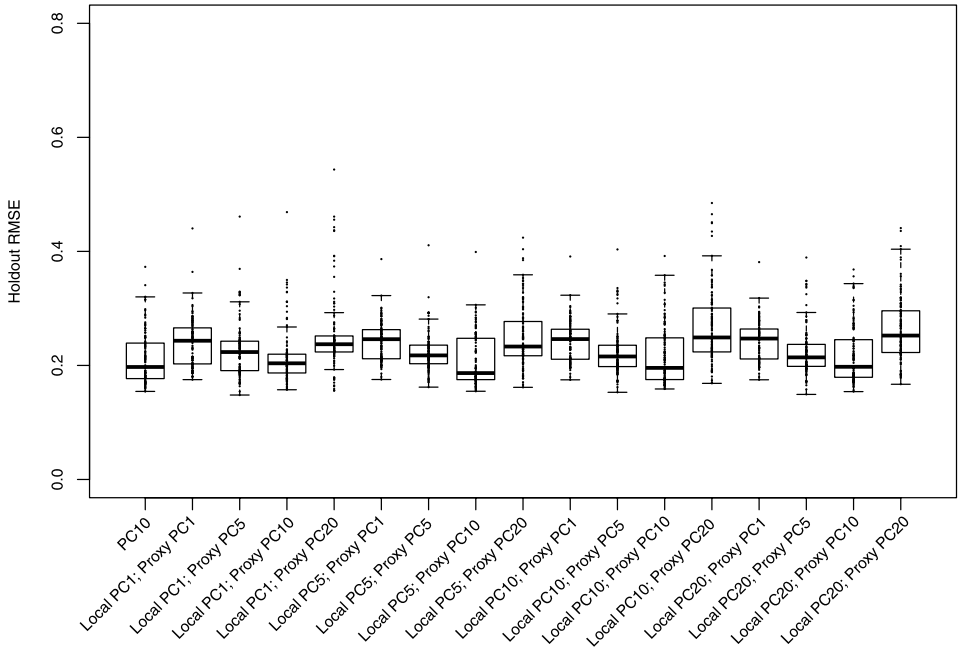

FIG. 13. *Cross-validated RMSE on* 30*-year holdout blocks for various model specifications: regression on the first ten principal components of the proxies and various two-stage models where global temperature is regressed on principal components of local temperatures which are then regressed on principal components of proxies.*

We plot our results in Figure 13 and again include the boxplot of ten principal components from Figure 11 for easy reference. As before, there is simply not that much variation in holdout RMSE across the various model specifications. No method is a clear winner.

4.2. *Temperature reconstructions.* Each model discussed in Section 4.1 can form a historical backcast. This backcast is simply the model's estimate $\hat{y}_k(\mathbf{x}_t)$ of the Northern Hemisphere average temperature in a year t calculated by inputing the proxy covariates $\mathbf{x}_t$ in the same year. The model index is k which varies over all 27 models from Section 4.1 (i.e., those featured in Figures 11–13). We plot these backcasts in Figure 14 in gray and show the CRU average in black. As can be seen, while these models all perform similarly in terms of cross-validated RMSE, they have wildly different implications about climate history.

According to some of them (e.g., the ten proxy principal component model given in green or the two-stage model featuring five local temperature principal components and five proxy principal components given in blue), the recent run-up in temperatures is not that abnormal, and similarly high temperatures would have been seen over the last millennium. Interestingly, the

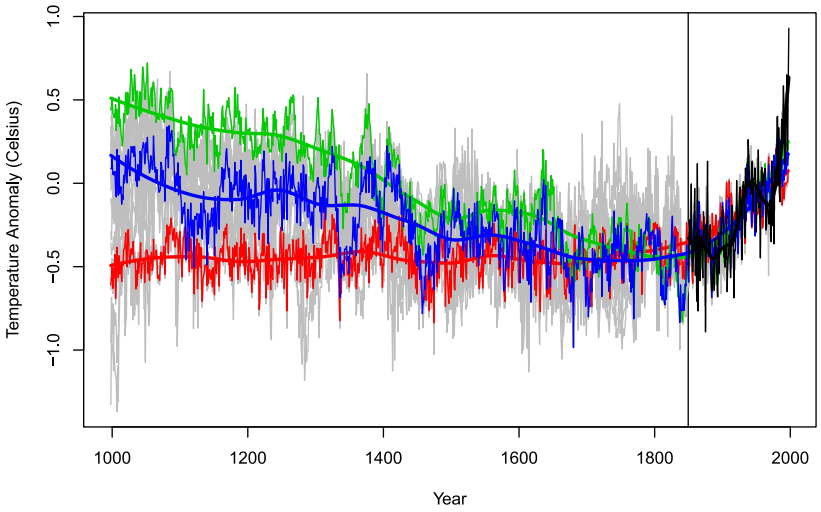

FIG. 14. *Backcasts to* 1000 *AD from the various models considered in this section are plotted in gray. CRU Northern Hemisphere annual mean land temperature is given by the thin black line with a smoothed version given by the thick black line. Three forecasts are featured: regression on one proxy principal component (red), regression on ten proxy principal components (green), and the two-stage model featuring five local temperature principal components and five proxy principal components (blue).*

blue backcast seems to feature both a Medieval Warm Period and a Little Ice Age whereas the green one shows only increasing temperatures going back in time.

However, other backcasts (e.g., the single proxy principal component regression featured in red) are in fact hockey sticks which correspond quite well to backcasts such as those in Mann, Bradley and Hughes (1999). If they are correct, modern temperatures are indeed comparatively quite alarming since such temperatures are much warmer than what the backcasts indicate was observed over the past millennium.

Figure 14 reveals an important concern: models that perform similarly at predicting the instrumental temperature series (as revealed by Figures 11–13) tell very different stories about the past. Thus, insofar as one judges models by cross-validated predictive ability, one seems to have no reason to prefer the red backcast in Figure 14 to the green even though the former suggests that recent temperatures are much warmer than those observed over the past 1000 years while the latter suggests they are not.

A final point to note is that the backcasts plotted in Figure 14 are the raw backcasts themselves with no accounting for backcast standard errors. In the next section, we take on the problem of specifying a full probability model which will allow us to provide accurate, pathwise standard errors.

5. Bayesian reconstruction and validation.

5.1. *Model specification.* In the previous section, we showed that a variety of different models perform fairly similarly in terms of cross-validated RMSE while producing very different temperature reconstructions. In this section, we focus and expand on the model which uses the first ten principal components of the proxy record to predict Northern Hemisphere CRU. We chose this forecast for several reasons. First, it performed relatively well compared to all of the others (see Figures 11–13). Second, PC regression has a relatively long history in the science of paleoclimatological reconstructions [Mann, Bradley and Hughes (1998, 1999), NRC (2006)]. Finally, when using OLS regression, principal components up to and including the tenth were statistically significant. While the t-statistics and their associated p-values themselves are uninterpretable due to the complex time series and error structures, these traditional benchmarks can serve as guideposts.

However, there is at least one serious problem with this model as it stands: the residuals demonstrate significant autocorrelation not captured by the autocorrelation in the proxies. Accordingly, we fit a variety of autoregressive models to CRU time series. With an AR2 model, the residuals showed very little autocorrelation.

So that we account for both parameter uncertainty as well as residual uncertainty, we estimate our model using Bayesian procedures. Our likelihood is given by

$$y_t = \beta_0 + \sum_{i=1}^{10} \beta_i x_{t,i} + \beta_{11} y_{t+1} + \beta_{12} y_{t+2} + \varepsilon_t,$$

$$\varepsilon_t \sim N(0, \sigma^2).$$

In our equation, y_t represents the CRU Northern Hemisphere annual land temperature in year t and $x_{t,i}$ is the value of principal component i in year t. We note that the subscripts on the right-hand side of the regression equation employ pluses rather than the usual minuses because we are interested in backcasts rather than forecasts. In addition to this, we use the very weakly informative priors

$$\vec{\beta} \sim N(\vec{0}, 1000 \cdot I),$$

$$\sigma \sim \text{Unif}(0, 100),$$

where $\vec{\beta}$ is the 13 dimensional vector $(\beta_0, \beta_1, \ldots, \beta_{12})^T$, $\vec{0}$ is a vector of 13 zeros, and I is the 13 dimensional identity matrix. This prior is sufficiently noninformative that the posterior mean of $\vec{\beta}$ is, within rounding error, equal to the maximum likelihood estimate. Furthermore, the prior on σ is effectively noninformative as y_t is always between ± 1 and therefore no posterior draw comes anywhere near the boundary of 100.

It is important to consider how our model accounts for the perils of temperature reconstruction discussed above. First and foremost, we deal with the problem of weak signal by building a simple model ($\mathrm{AR}\,2 + \mathrm{PC}10$) in order to avoid overfitting. Our fully Bayesian model, which accounts for parameter uncertainty, also helps attenuate some of the problems caused by weak signal. Dimensionality reduction is dealt with via principal components. PCs have two additional benefits. First, they are well-studied in the climate science literature and are used in climate scientists' reconstructions. Second, the orthogonality of principal components will diminish the pernicious effects of spatial correlation among the proxies. Finally, we address the temporal correlation of the temperature series with the AR2 component of our model.

5.2. *Comparison to other models.* An approach that is broadly similar to the above has recently appeared in the climate literature [Li, Nychka and Amman (2007)] for purposes similar to ours, namely, quantifying the uncertainty of a reconstruction. In fact, Li, Nychka and Amman (2007) is highly unusual in the climate literature in that its authors are primarily statisticians. Using a dataset of 14 proxies from Mann, Bradley and Hughes (1999), Li, Nychka and Amman (2007) confirms the findings of Mann, Bradley and Hughes (1998, 1999) but attempts to take forecast error, parameter uncertainty, and temporal correlation into account. They provide toy data and code for their model here: `http://www.image.ucar.edu/~boli/research.html`

Nevertheless, several important distinctions between their model and ours exist. First, Li, Nychka and Amman (2007) make use of a dataset over ten years old [Mann, Bradley and Hughes (1999)] which contains only 14 proxies dating back to 1000 AD and has instrumental records dating 1850–1980 AD. On the other hand, we make use of the latest multi-proxy database [Mann et al. (2008)] which contains 93 proxies dating back to 1000 AD and has instrumental records dating 1850–1998 AD. Furthermore, Li, Nychka and Amman (2007) assume an AR2 structure on the errors from the model where we assume the model is AR2 with covariates. Finally, and perhaps most importantly, Li, Nychka and Amman (2007) estimate their model via generalized least squares and therefore use (i) the parametric bootstrap in order to account for parameter estimation uncertainty and (ii) cross-validation to account overfitting the in-sample period (i.e., to inflate their estimate of the error variance σ). On the other hand, by estimating our model in a fully Bayesian fashion, we can account for these within our probability model. Thus, our procedure can be thought of as formalizing the approach of Li, Nychka and Amman (2007) and it provides practically similar results when applied to the same set of covariates (generalized least squares also produced

practically indistinguishable forecasts and backcasts though obviously narrower standard errors).

At the time of this manuscript's submission, the same authors were working on a fully Bayesian model which deserves mention [subsequently published as Li, Nychka and Amman (2010)]. In this paper, they integrate data from three types of proxies measured at different timescales (tree rings, boreholes, and pollen) as well as data from climate forcings (solar irradiance, volcanism, and greenhouse gases) which are considered to be external drivers of climate. Furthermore, they account for autocorrelated error in both the proxies and forcings as well as autocorrelation in the deviations of temperature from the model. While the methodology and use of forcing data are certainly innovative, the focus of Li, Nychka and Amman (2010) is not on reconstruction *per se*; rather, they are interested in validating their modeling approach taking as "truth" the output of a high-resolution state-of-the-art climate simulation [Amman et al. (2007)]. Consequently, all data used in the paper is synthetic and they concentrate on methodological issues, "defer[ring] any reconstructions based on actual observations and their geophysical interpretation to a subsequent paper" [Li, Nychka and Amman (2010)].

Finally, Tingley and Huybers (2010a, 2010b) have developed a hierarchical Bayesian model to reconstruct the full temperature field. They fit the model to experimental datasets formed by "corrupting a number of the [temperature] time series to mimic proxy observations" [Tingley and Huybers (2010a)]. Using these datasets, they conduct what is in essence a frequentist evaluation of their Bayesian model [Tingley and Huybers (2010a)] and then compare its performance to that of the well-known RegEM algorithm [Tingley and Huybers (2010b)]. Like Li, Nychka and Amman (2010), however, they do not use their model to produce temperature reconstructions from actual proxy observations.

5.3. *Model reconstruction.* We create a full temperature backcast by first initializing our model with the CRU temperatures for 1999 AD and 2000 AD. We then perform a "one-step-behind" backcast, plugging these values along with the ten principal component values for 1998 AD into the equation $y_t = \beta_0 + \sum_{i=1}^{10} \beta_i x_{t,i} + \beta_{11} y_{t+1} + \beta_{12} y_{t+2}$ to get a backcasted value for 1998 AD (using the posterior mean of $\vec{\beta}$ as a plug-in estimator). Similarly, we use the CRU temperature for 1999 AD, this backcasted value for 1998 AD, and the ten principal component values for 1997 AD to get a backcasted value for 1997 AD. Finally, we then iterate this process one year at a time, using the two most recent backcasted values as well as the current principal component values, to get a backcast for each of the last 1000 years.

We plot the in-sample portion of this backcast (1850–1998 AD) in Figure 15. Not surprisingly, the model tracks CRU reasonably well because it is

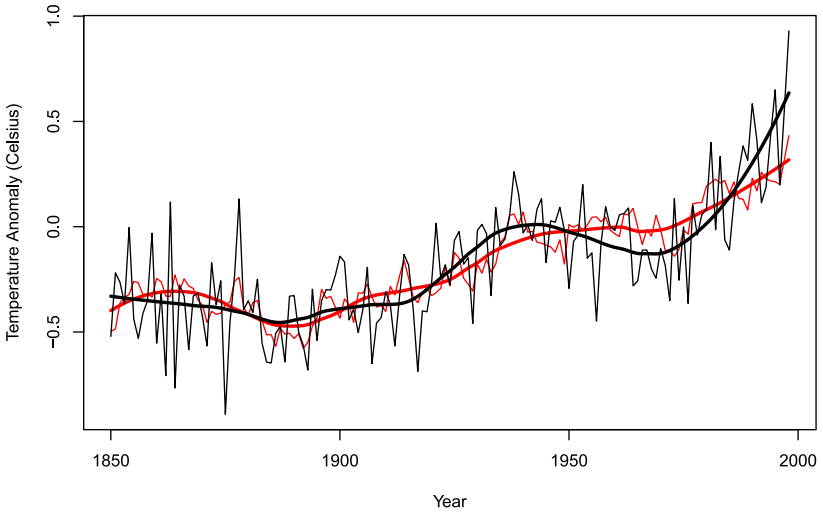

FIG. 15. *In-sample Backcast from Bayesian Model of Section 5. CRU Northern Hemisphere annual mean land temperature is given by the thin black line and a smoothed version is given by the thick black line. The backcast is given by the thin red line and a smoothed version is given by the thick red line. The model is fit on 1850–1998 AD.*

in-sample. However, despite the fact that the backcast is both in-sample and initialized with the high true temperatures from 1999 AD and 2000 AD, it still cannot capture either the high level of or the sharp run-up in temperatures of the 1990s: it is substantially biased low. That the model cannot capture run-up even in-sample does not portend well for its ability to capture similar levels and run-ups if they exist out-of-sample.

A benefit of our fully Bayesian model is that it allows us to assess the error due to both (i) residual variance (i.e., ε_t) and (ii) parameter uncertainty. Furthermore, we can do this in a fully pathwise fashion. To assess the error due to residual variance, we use the one-step-behind backcasting procedure outlined above with two exceptions. First, at each step, we draw an error from a $N(0, \sigma^2)$ distribution and add it to our backcast. These errors then propagate through the full path of backcast. Second, we perform the backcast allowing σ to vary over our samples from the posterior distribution.

To assess the error due to the uncertainty in $\vec{\beta}$, we perform the original one-step-behind backcast [i.e., without drawing an error from the $N(0, \sigma^2)$ distribution]. However, rather than using the posterior mean of $\vec{\beta}$, we perform the backcast for each of our samples from the posterior distribution of $\vec{\beta}$.

Finally, to get a sense of the full uncertainty in our backcast, we can combine both of the methods outlined above. That is, for each draw from the posterior of $\vec{\beta}$ and σ, we perform the one-step-behind backcast drawing

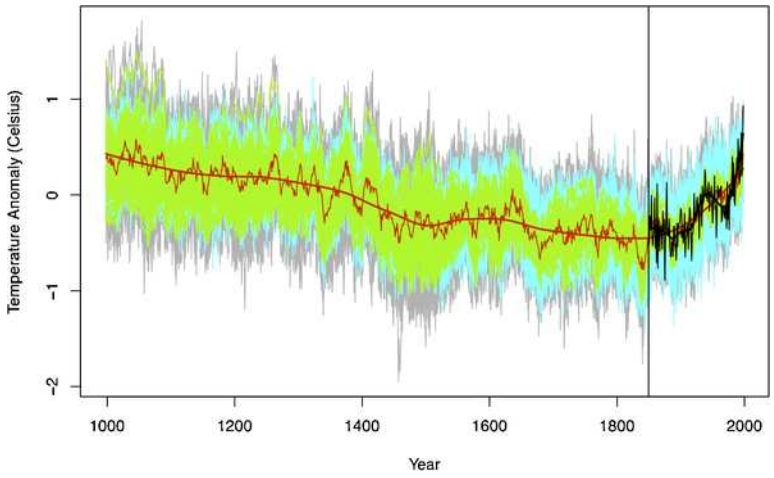

FIG. 16. *Backcast from Bayesian Model of Section 5. CRU Northern Hemisphere annual mean land temperature is given by the thin black line and a smoothed version is given by the thick black line. The forecast is given by the thin red line and a smoothed version is given by the thick red line. The model is fit on 1850–1998 AD and backcasts 998–1849 AD. The cyan region indicates uncertainty due to ε_t, the green region indicates uncertainty due to $\vec{\beta}$, and the gray region indicates total uncertainty.*

errors from the $N(0, \sigma^2)$ distribution. This gives one curve for each posterior draw, each representing a draw of the full temperature series conditional on the data and the model. Taken together, they form an approximation to the full posterior distribution of the temperature series.

We decompose the uncertainty of our model's backcast by plotting the curves drawn using each of the methods outlined in the previous three paragraphs in Figure 16. As can be seen, in the modern instrumental period the residual variance (in cyan) dominates the uncertainty in the backcast. However, the variance due to $\vec{\beta}$ uncertainty (in green) propagates through time and becomes the dominant portion of the overall error for earlier periods. The primary conclusion is that failure to account for parameter uncertainty results in overly confident model predictions.

As far as we can tell, no effort at paleoclimatological global temperature reconstruction of the past 1000 years has used a fully Bayesian probability model to incorporate parameter uncertainty into the backcast estimates [in fact, the aforementioned Li, Nychka and Amman (2007) paper is the only paper we know of that even begins to account for uncertainty in some of the parameters; see Haslett et al. (2006) for a Bayesian model used for reconstructing the local prehistoric climate in Glendalough, Ireland]. The widely used approach in the climate literature is to estimate uncertainty using residuals (usually from a holdout period). Climate scientist generally report less accurate reconstructions in more distant time periods, but this

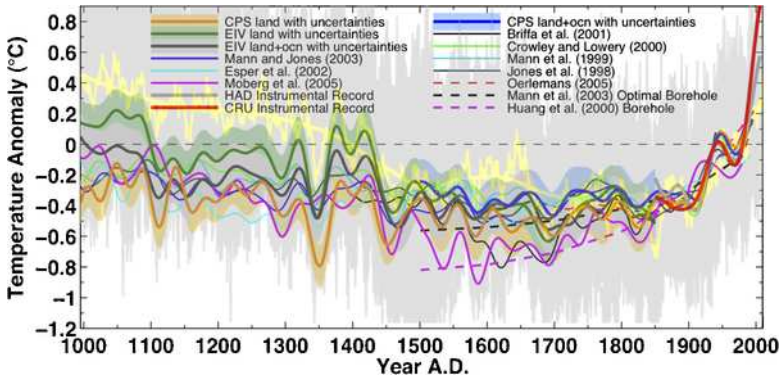

FIG. 17. *This figure modifies Figure* 3 *from Mann et al. (2008). We take that figure and superimpose the backcast from Bayesian model of Section 5. The backcast is given by the thin yellow line, the smoothed backcast by a thick yellow line, and the backcast error in gray.*

is due to the fact that there are fewer proxies that extend further back into time and therefore larger validation residuals.

5.4. *Comparison to other reconstructions and posterior calculations.* What is most interesting is comparing our backcast to those from Mann et al. (2008) as done in Figure 17. We see that our model gives a backcast which is very similar to those in the literature, particularly from 1300 AD to the present. In fact, our backcast very closely traces the Mann et al. (2008) EIV land backcast, considered by climate scientists to be among the most skilled. Though our model provides slightly warmer backcasts for the years 1000–1300 AD, we note it falls within or just outside the uncertainty bands of the Mann et al. (2008) EIV land backcast even in that period. Hence, our backcast matches their backcasts reasonably well.

The major difference between our model and those of climate scientists, however, can be seen in the *large width* of our uncertainty bands. Because they are pathwise and account for the uncertainty in the parameters (as outlined in Section 5.3), they are much larger than those provided by climate scientists. In fact, our uncertainty bands are so wide that they *envelop* all of the other backcasts in the literature. Given their ample width, it is difficult to say that recent warming is an extraordinary event compared to the last 1000 years. For example, according to our uncertainty bands, it is possible that it was as warm in the year 1200 AD as it is today. In contrast, the reconstructions produced in Mann et al. (2008) are completely pointwise.

Another advantage of our method is that it allows us to calculate posterior probabilities of various scenarios of interest by simulation of alternative sample paths. For example, 1998 is generally considered to be the warmest

year on record in the Northern Hemisphere. Using our model, we calculate that there is a 36% posterior probability that 1998 was the warmest year over the past thousand. If we consider rolling decades, 1997–2006 is the warmest on record; our model gives an 80% chance that it was the warmest in the past 1000 years. Finally, if we look at rolling 30-year blocks, the posterior probability that the last 30 years (again, the warmest on record) were the warmest over the past thousand is 38%.

Similarly, we can look at posterior probabilities of the run-up in (or derivative of) temperatures in addition to the levels. For this purpose, we defined the "derivative" as the difference between the value of the loess smooth of the temperature series (or reconstruction series) in year t and year $t-k$. For $k=10$, $k=30$, and $k=60$, we estimate a zero posterior probability that the past 1000 years contained run-ups larger than those we have experienced over the past ten, 30, and 60 years (again, the largest such run-ups on record). This suggests that the temperature derivatives encountered over recent history are unprecedented in the millennium. While this does seem alarming, we should temper our alarm somewhat by considering again Figure 15 and the fact that the proxies seem unable to capture the sharp run-up in temperature of the 1990s. That is, our posterior probabilities are based on derivatives from our model's proxy-based reconstructions and we are comparing these derivatives to derivatives of the *actual* temperature series; insofar as the proxies cannot capture sharp run-ups, our model's reconstructions will not be able to either and therefore will tend to understate the probability of such run-ups.

5.5. *Model validation.* Though our model gives forecasts and backcasts that are broadly comparable to those provided by climate scientists, our approach suggests that there is substantial uncertainty about the ability of the model to fit and predict new data. Climate scientists estimate out-of-sample uncertainty using only two holdout blocks: one at the beginning of the instrumental period and one at the end. We pursue that strategy here. First, we fit on 1880–1998 AD and attempt to backcast 1850–1879 AD. Then, we fit on 1850–1968 AD and forecast 1969–1998 AD. These blocks are arguably the most interesting and important because they are not "tied" at two endpoints. Thus, they genuinely reflect the most important modeling task: reconstruction.

Figure 18 illustrates that the model seems to perform reasonably well on the first holdout block. Our reconstruction regresses partly back toward the in-sample mean. Compared to the actual temperature series, it is biased a bit upward. On the other hand, the model is far more inaccurate on the second holdout block, the modern period. Our reconstruction, happily, does not move toward the in-sample mean and even rises substantively at first. Still, it seems there is simply not enough signal in the proxies to detect either the

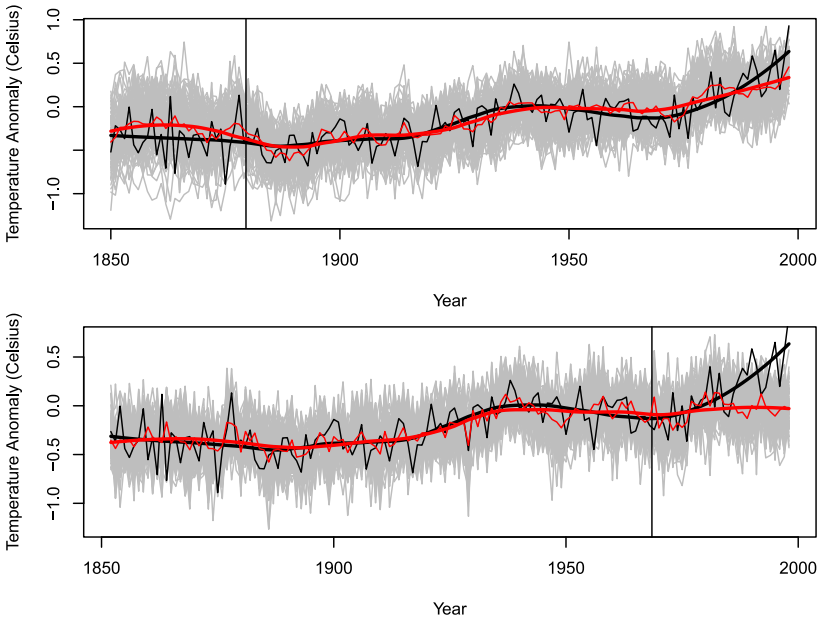

FIG. 18. *Predictions from the Bayesian model of Section 5 when the first* 30 *years of instrumental data are held out* (top) *and when the last* 30 *years of instrumental data are held out* (bottom). *CRU is given in black and the model predictions in red. The raw data and predictions are given by the thin lines and loess smooths are given by the thick lines. Uncertainty bands are indicated by the gray region.*

high levels of or the sharp run-up in temperature seen in the 1990s. This is disturbing: if a model cannot predict the occurrence of a sharp run-up in an out-of-sample block which is contiguous with the in-sample training set, then it seems *highly unlikely* that it has power to detect such levels or run-ups in the more distant past. It is even more discouraging when one recalls Figure 15: the model cannot capture the sharp run-up even *in-sample*. In sum, these results suggest that the 93 sequences that comprise the 1000-year-old proxy record simply lack power to detect a sharp increase in temperature.[12]

As mentioned earlier, scientists have collected a large body of evidence which suggests that there was a Medieval Warm Period (MWP) at least in

[12]On the other hand, perhaps our model is unable to detect the high level of and sharp run-up in recent temperatures because anthropogenic factors have, for example, caused a regime change in the relation between temperatures and proxies. While this is certainly a consistent line of reasoning, it is also fraught with peril for, once one admits the possibility of regime changes in the instrumental period, it raises the question of whether such changes exist elsewhere over the past 1000 years. Furthermore, it implies that up to half of the already short instrumental record is corrupted by anthropogenic factors, thus undermining paleoclimatology as a statistical enterprise.

TABLE 2
Percent of time various null models outperform the Bayesian model of Section 5

Pseudo-proxy	**First block p-value**	**Last block p-value**
White Noise	0.0%	0.0%
AR1(0.25)	0.1%	0.0%
AR1(0.4)	0.1%	0.0%
Empirical AR1	24.1%	20.6%
Brownian Motion	16.4%	32.2%

portions of the Northern Hemisphere. The MWP is believed to have occurred c. 800–1300 AD (it was followed by the Little Ice Age). It is widely hoped that multi-proxy models have the power to detect (i) how warm the Medieval Warm Period was, (ii) how sharply temperatures increased during it, and (iii) to compare these two features to the past decade's high temperatures and sharp run-up. Since our model cannot detect the recent temperature change, detection of dramatic changes hundreds of years ago seems out of the question.

This is not to say that the proxy record is unrelated to temperatures. We can compare our model's RMSE in these two holdout periods to various null models which we know have no signal. That is, we can perform a test similar to that of Section 3.4. On each holdout block, we generate a 149×93 matrix of pseudo-proxies from each of the six null models known to be independent of the temperature series. Then, analogously to our model, we take the first ten principal components of these pseudo-proxies, regress the in-sample temperature on the ten in-sample principal components, and compute the RMSE on the holdout block. We perform this procedure 1000 times for each holdout block and then calculate the percentage of time that the model fit to the pseudo-proxies beats our model.

Our model, with an RMSE of 0.26 on the first holdout block and an RMSE of 0.36 on the second handily outperforms the relatively unsophisticated white noise and weak AR1 process pseudo-proxies (see Table 2). Again, this is not surprising. These pseudo-proxies cannot capture the local dependence in the instrumental record, so they regress sharply to the in-sample mean. On the other hand, the Empirical AR1 processes and Brownian motion have more complex local structure so they provide respectable competition to our model. These models capture only the local dependence in the temperature record: in the long term, forecasts based off the AR1 processes will slide slowly back to the in-sample mean and forecasts based off Brownian motion will wander aimlessly. Taken together, it follows that our model is at best weakly significant relative to the Empirical AR1 process or Brownian motion on either holdout block.

In tandem, Figure 18 and Table 2 should make us very cautious about using our model to extrapolate, even with wide standard errors. The second panel of Figure 18 demonstrates that these standard errors are too narrow even for very temporally short forecasts. While we are able to replicate the significance tests in Mann, Bradley and Hughes (1998), our Table 2 shows that our model does not pass "statistical significance" thresholds against savvy null models. Ultimately, what these tests essentially show is that the 1000-year-old proxy record has little power given the limited temperature record.

6. Conclusion. Research on multi-proxy temperature reconstructions of the earth's temperature is now entering its second decade. While the literature is large, there has been very little collaboration with university-level, professional statisticians [Wegman, Scott and Said (2006), Wegman (2006)]. Our paper is an effort to apply some modern statistical methods to these problems. While our results agree with the climate scientists findings in some respects, our methods of estimating model uncertainty and accuracy are in sharp disagreement.

On the one hand, we conclude unequivocally that the evidence for a "long-handled" hockey stick (where the shaft of the hockey stick extends to the year 1000 AD) is lacking in the data. The fundamental problem is that there is a limited amount of proxy data which dates back to 1000 AD; what is available is weakly predictive of global annual temperature. Our backcasting methods, which track quite closely the methods applied most recently in Mann (2008) to the same data, are unable to catch the sharp run up in temperatures recorded in the 1990s, even in-sample. As can be seen in Figure 15, our estimate of the run up in temperature in the 1990s has a much smaller slope than the actual temperature series. Furthermore, the lower frame of Figure 18 clearly reveals that the proxy model is not at all able to track the high gradient segment. Consequently, the long flat handle of the hockey stick is best understood to be a feature of regression and less a reflection of our knowledge of the truth. Nevertheless, the temperatures of the last few decades have been relatively warm compared to many of the 1000-year temperature curves sampled from the posterior distribution of our model.

Our main contribution is our efforts to seriously grapple with the uncertainty involved in paleoclimatological reconstructions. Regression of high-dimensional time series is always a complex problem with many traps. In our case, the particular challenges include (i) a short sequence of training data, (ii) more predictors than observations, (iii) a very weak signal, and (iv) response and predictor variables which are both strongly autocorrelated. The final point is particularly troublesome: since the data is not easily modeled

by a simple autoregressive process, it follows that the number of truly independent observations (i.e., the effective sample size) may be just too small for accurate reconstruction.

Climate scientists have greatly underestimated the uncertainty of proxy-based reconstructions and hence have been overconfident in their models. We have shown that time dependence in the temperature series is sufficiently strong to permit complex sequences of random numbers to forecast out-of-sample reasonably well fairly frequently (see Figures 9 and 10). Furthermore, even proxy-based models with approximately the same amount of reconstructive skill (Figures 11–13), produce strikingly dissimilar historical backcasts (Figure 14); some of these look like hockey sticks but most do not.

Natural climate variability is not well understood and is probably quite large. It is not clear that the proxies currently used to predict temperature are even predictive of it at the scale of several decades let alone over many centuries. Nonetheless, paleoclimatoligical reconstructions constitute only one source of evidence in the AGW debate.

Our work stands entirely on the shoulders of those environmental scientists who labored untold years to assemble the vast network of natural proxies. Although we assume the reliability of their data for our purposes here, there still remains a considerable number of outstanding questions that can only be answered with a free and open inquiry and a great deal of replication.

Acknowledgments. We thank Editor Michael Stein, two anonymous referees, and Tilmann Gneiting for their helpful suggestions on our manuscript. We also thank our colleagues Larry Brown and Dean Foster for many helpful conversations.

SUPPLEMENTARY MATERIAL

Code repository for "A statistical analysis of multiple temperature proxies: Are reconstructions of surface temperatures over the last 1000 years reliable?" (DOI: 10.1214/10-AOAS398SUPP; .zip). This repository archives all data and code used for "A statistical analysis of multiple temperature proxies: Are reconstructions of surface temperatures over the last 1000 years reliable?" In particular, it contains code to make all figures and tables featured in the paper.

REFERENCES

Amman, C. M., Joos, F., Otto-Bliesner, B. L. and Tomas, R. (2007). Solar influence on climate during the past millennium: Results from transient simulations with the NCAR climate system model. *Proc. Natl. Acad. Sci. USA* **104** 3713–3718.

Ammann, C. and Wahl, E. (2007). The importance of the geophysical context in statistical evaluations of climate reconstruction procedures. *Climatic Change* **85** 71–88.

BBC (2008). *Earth: The Climate Wars*. British Broadcasting Company, September 14.
Bradley, R. S. (1999). *Paleoclimatology: Reconstructing Climates of the Quaternary*, 2nd ed. Academic Press, San Diego.
Christiansen, B., Schmith, T. and Thejll, P. (2009). A surrogate emsemble study of climate reconstruction methods: Stochasticity and robustness. *Journal of Climate* **22** 951–976.
Gore, A. (2006). *An Inconvenient Truth*. Lawrence Bender Productions.
Granger, C. W. J. and Newbold, P. (1974). Spurious regressions in econometrics. *Journal of Econometrics* **2** 111–120.
Green, K. C., Armstrong, J. S. and Soon, W. (2009). Validity of climate change forecasting for public policy decision-making. *International Journal of Forecasting* **25** 826–832.
Haslett, J., Whiley, M., Bhattacharya, S., Salter-Townshend, M., Wilson, S. P., Allen, J. R. M., Huntley, B. and Mitchell, F. J. G. (2006). Bayesian palaeoclimate reconstruction. *J. Roy. Statist. Soc. Ser. A* **169** 395–438. MR2236914
IPCC (2001). *Climate Change 2001: The Scientific Basis*. Cambridge Univ. Press, Cambridge.
IPCC (2007). *Climate Change 2007: The Physical Science Basis*. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change.
Johnson, K. (2009). Climate emails stoke debate. *Wall Street Journal*. November 23, page A3.
Johnson, K. and Naik, G. (2009). Lawmakers probe climate emails. *Wall Street Journal*. November 24, page A8.
Jolis, A. (2009). Revenge of the climate laymen. *Wall Street Journal*. November 18.
Ladurie, E. L. (1971). *Times of Feast, Times of Famine: A History of Climate Since the Year 1000*. Doubleday, New York.
Lamb, H. H. (1990). *Climate: Past, Present and Future*. Routledge, New York.
Lee, T., Zwiers, F. W. and Tsao, M. (2008). Evaluation of proxy-based millennial reconstruction methods. *Climate Dynamics* **31** 263–281.
Li, B., Nychka, D. W. and Amman, C. M. (2007). The 'hockey stick' and the 1990s: A statistical perspective on reconstructing hemispheric temperatures. *Tellus* **59A** 591–598.
Li, B., Nychka, D. W. and Amman, C. M. (2010). The value of multi-proxy reconstructin of past climate. *J. Amer. Statist. Assoc.* **105** 883–895.
Luterbacher, J., Dieterich, D., Xoplaki, E., Grosjean, M. and Wanner, H. (2004). European seasonal and annual temperature variability, trends, and extremes since 1500. *Science* **202** 1499–1503.
Mann, M. E. (2004). On smoothing potentially non-stationary climate time series. *Geophysical Research Letters* **31**.
Mann, M. E. (2008). Smoothing of climate time series revisited. *Geophysical Research Letters* **35**.
Mann, M. E., Bradley, R. E. and Hughes, M. K. (1998). Global-scale temperature patterns and climate forcing over the past six centuries. *Nature* **392** 779–787.
Mann, M. E., Bradley, R. E. and Hughes, M. K. (1999). Northern hemisphere temperatures during the past millennium: Inferences, uncertainties, and limitations. *Geophysical Research Letters* **26** 759–762.
Mann, M. E., Bradley, R. E. and Hughes, M. K. (2004). Corrigendum: Global-scale temperature patterns and climate forcing over the past six centuries. *Nature* **430** 105.
Mann, M. E. and Rutherford, S. (2002). Climate reconstruction using pseudoproxies. *Geophysical Research Letters* **29** 1501.

Mann, M. E., Rutherford, S., Wahl, E. and Ammann, C. (2005). Testing the fidelity of methods used in proxy-based reconstructions of past climate. *Journal of Climate* **18** 4097–4107.

Mann, M. E., Rutherford, S., Wahl, E. and Ammann, C. (2007). Robustness of proxy-based climate field reconstruction methods. *Journal of Geophysical Research* **112**.

Mann, M. E., Zhang, Z., Hughes, M. K., Bradley, R. S., Miller, S. K., Rutherford, S. and Ni, F. (2008). Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millenia. *Proc. Natl. Acad. Sci. USA* **105** 13252–13257.

Matthes, F. E. (1939). Report of the committee on glaciers. *Transactions of the American Geophysical Union* **20** 518–523.

McIntyre, S. and McKitrick, R. (2003). Corrections to the Mann et al. (1998) proxy base and northern hemispheric average temperature series. *Energy and Environment* **14** 751–771.

McIntyre, S. and McKitrick, R. (2005a). Hockey sticks, principal components, and spurious significance. *Geophysical Research Letters* **32**.

McIntyre, S. and McKitrick, R. (2005b). The M&M critique of the MBH98 and northern hemisphere climate index: Update and implications. *Energy and Environment* **16** 69–100.

McIntyre, S. and McKitrick, R. (2005c). Reply to comment by Huybers on "Hockey sticks, principal components, and spurious significance." *Geophysical Research Letters* **32** L20713.

McShane, B. B. and Wyner, A. J. (2011). Supplement to: "A statistical analysis of multiple temperature proxies: Are reconstructions of surface temperatures over the last 1000 years reliable?" DOI: 10.1214/10-AOAS398SUPP.

National Research Council (2006). Surface temperature reconstructions. National Academic Press, Washington, DC.

Phillips, P. C. B. (1986). Understanding spurious regressions in econometrics. *J. Econometrics* **33** 311–340. MR0867979

Rothstein, E. (2008). Apocalypse now, via diorama. *The New York Times*. October 17, page C27.

Rutherford, S., Mann, M. E., Osborn, T. J., Bradley, R. S., Briffa, K. R., Hughes, M. K. and Jones, P. D. (2005). Proxy-based northern hemispheric surface reconstructions: Sensitivity to method, predictor network, target season, and target domain. *Journal of Climate* **18** 2308–2329.

Schneider, T. (2001). Analysis of incomplete climate data: Estimation of mean values and covariance matrices and imputation of missing values. *Journal of Climate* **14** 853–871.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *J. Roy. Statist. Soc. Ser. B* **58** 267–288. MR1379242

Tingley, M. and Huybers, P. (2010a). A Bayesian algorithm for reconstructing climate anomalies in space and time. Part I: Development and applications to paleoclimate reconstruction problems. *Journal of Climate* **23** 2759–2781.

Tingley, M. and Huybers, P. (2010b). Bayesian algorithm for reconstructing climate anomalies in space and time. Part II: Comparison with the regularized expectation–maximization algorithm. *Journal of Climate* **23** 2782–2800.

von Storch, H. E. and Zorita, E. (2005). Comment on "Hockey sticks, principal components, and spurious significance," by S. McIntyre and R. McKitrick. *Geophysical Research Letters* **32**.

von Storch, H. E., Zorita, E., Jones, J. M., Dimitriev, Y., Gonzalez-Rouco, F. and Tett, S. (2004). Reconstructing past climate from noisy data. *Science* **306** 679–682.

von Storch, H. E., Zorita, E., Jones, J. M., Dimitriev, Y., Gonzalez-Rouco, F. and Tett, S. (2006). Response to comment on "Reconstructing past climate from noisy data." *Science* **213** 529.

Wahl, E. R. and Amman, C. M. (2006). Robustness of the Mann, Bradley, Hughes reconstruction of the Northern Hemisphere surface temperatures: Examination of criticisms based on the nature and processing of proxy climate evidence. *Climatic Change* **85** 33–69.

Wahl, E. R., Ritson, D. M. and Amman, C. M. (2006). On "Reconstruction past climate from noisy data." *Science* **312** 592b.

Wegman, E. J. (2006). Response of Dr. Edward Wegman to questions posed by the honorable Bart Stupak in connection with testimony to the subcommittee on oversight and investigations. Available at http://www.uoguelph.ca/~rmckitri/research/StupakResponse.pdf.

Wegman, E. J., Scott, D. W. and Said, Y. H. (2006). Ad Hoc committee report on the 'hockey stick' global climate reconstruction. Available at http://republicans.energycommerce.house.gov/108/home/07142006_Wegman_Report.pdf.

Wu, Y., Boos, D. D. and Stefanski, L. A. (2007). Controlling variable selection by the addition of pseudovariables. *J. Amer. Statist. Assoc.* **102** 235–243. MR2345541

Yule, G. U. (1926). Why do we sometimes get nonsense correlations between time series? A study in sampling and the nature of time series. *J. Roy. Statist. Soc.* **89** 1–64.

Kellogg School of Management
Northwestern University
Leverone Hall
2001 Sheridan Road
Evanston, Illinois 60208
USA
E-mail: b-mcshane@kellogg.northwestern.edu
URL: http://www.blakemcshane.com

Department of Statistics
The Wharton School
University of Pennsylvania
400 Jon M. Huntsman Hall
3730 Walnut Street
Philadelphia, Pennsylvania 19104
USA
E-mail: ajw@wharton.upenn.edu
URL: http://www.adiwyner.com

Exhibit C

ORAL ARGUMENT NOT YET SCHEDULED

UNITED STATES COURT OF APPEALS
FOR THE DISTRICT OF COLUMBIA CIRCUIT

COMPETITIVE ENTERPRISE INSTITUTE, FREEDOMWORKS, and THE SCIENCE AND ENVIRONMENTAL POLICY PROJECT, Petitioners, v. UNITED STATES ENVIRONMENTAL PROTECTION AGENCY, Respondent.	No. 10-1318 consolidated with: 10-1234 et al.

NONBINDING STATEMENT OF ISSUES

Petitioners hereby file this Nonbinding Statement of Issues in their challenge to the Environmental Protection Agency's denial of their petition for reconsideration of EPA's greenhouse gas decision (popularly known as its Endangerment ruling). EPA's denial is published at 75 Fed. Reg. 49,556 (Aug. 13, 2010). In addition to the issues designated by petitioners in the cases consolidated with this action, Petitioners CEI *et al.* intend to raise the following issues:

1. Whether EPA's dismissal of evidence relating to the destruction of raw climate data was arbitrary, capricious, or otherwise contrary to law;

2. Whether EPA's treatment of the "Climategate" documents and of other evidence which developed or came to light after its Endangerment decision is arbitrary, capricious, or otherwise contrary to law;

3. Whether EPA's treatment of the "absurdity" issue as not being of central relevance is arbitrary, capricious, or otherwise contrary to law.

Dated: November 17, 2010.

Respectfully submitted,

/s/ Hans Bader_____
Sam Kazman
Hans Bader
COMPETITIVE ENTERPRISE INSTITUTE
1899 L Street, NW, 12th Floor
Washington, D.C. 20036
(202) 331-2265
Counsel for Petitioners in No. 10-1318

CERTIFICATE OF SERVICE

I hereby certify that on this 17th day of November, 2010, a copy of the foregoing Non-Binding Statement of Issues was served electronically through the court's CM/ECF system on all registered counsel and by first-class mail on counsel not registered as listed below:

Christopher Gene King
New York City Law Dept. 6-143
100 Church Street
New York, NY 10007

Mark J. Bennett
Office of the Attorney General
State of Hawaii, Dept. of Agriculture
425 Queen Street
Honolulu, HI 96813

Kelvin Allen Brooks
Office of the Attorney General
State of New Hampshire
33 Capitol Street
Concord, NH 03301-6397

Joseph P. Mikitish
Office of the Attorney General
State of Arizona
1275 W. Washington St.
Phoenix, AZ 85007-2926

Kimberly P. Massicotte
Office of the Attorney General
State of Connecticut
55 Elm Street
Hartford, CT 06106

/s/ Hans Bader

Exhibit D

Michael E. Mann thanks David. Expenses are covered on this, but there is a great need of additional resources for the Climate Science Legal Defense Fund to defend against the various other attacks by groups like ATI, etc. on me and other scientists. So folks should consider contributing to them if they can:
https://salsa.democracyinaction.org/o/823/p/salsa/donation/common/public/?donate_page_KEY=7935

Public Employees for Environmental Responsibility: Publications - Home
salsa.democracyinaction.org
Public Employees for Environmental Responsibility (PEER) is a national non-profi...
See More

August 27 at 6:51am · 4

Bob Fishell This paragraph is pretty ugly:

"But after months of legal wrangling to keep Virginia's attorney general, Ken Cuccinelli, from digging into his e-mails from his days at the University of Virginia, does Dr. Mann, now with Penn State, really want to open... See More
August 27 at 6:53am · 2

Michael E. Mann I agree Bob, that statement is quite misguided on several levels. A response is forthcoming.
August 27 at 7:06am · 4

Chip Mefford Absolutely fascinating. it's very difficult to sit back and attempt to respond to all attacks, I certainly understanding going on the offensive. However, tort law is at it's very best ambiguous, and there is no such thing as truth in a court of law, so it seems. Only law, not justice, not truth, none of that silliness.
I -for one- have absolutely no faith at all in the courts in matters such as these. None.
August 27 at 7:36am · 1

Michael E. Mann thanks Kevan. I've let the folks there know about that...
August 27 at 8:19am · 1

Kelly Anspaugh It is, of course, those who would teach our children that anthropogenic global warming is all a "liberal hoax" who are comparable to Jerry Sandusky. They are the ones engaged in intellectual molestation.
August 27 at 9:09am · 4