BEFORE THE UNITED STATES
ENVIRONMENTAL PROTECTION AGENCY

EPA Docket No. EPA-HQ-OAR-2009-0171

In re:

Endangerment and Cause or Contribute
Findings for Greenhouse Gases Under
Section 202(a) of the Clean Air Act

PACIFIC LEGAL FOUNDATION'S PETITION FOR RECONSIDERATION OF ENDANGERMENT AND CAUSE OR CONTRIBUTE FINDINGS FOR GREENHOUSE GASES UNDER SECTION 202 OF THE CLEAN AIR ACT

THEODORE HADZI-ANTICH
LUKE WAKE
RALPH W. KASARDA
Pacific Legal Foundation
3900 Lennane Drive, Suite 200
Sacramento, California 95834
Telephone: (916) 419-7111
Facsimile: (916) 419-7747
tha@pacificlegal.org
lw@pacificlegal.org
rwk@pacificlegal.org

Attorneys for Petitioner
Pacific Legal Foundation

I

INTRODUCTION

A. Scope of This Petition

Pursuant to Section 307(d) of the Clean Air Act, 42 U.S.C. § 7607(d), and based upon new information of central relevance not available during the public comment period, Pacific Legal Foundation (PLF) hereby petitions the United States Environmental Protection Agency (EPA) to convene a proceeding for the reconsideration of the "Endangerment and Cause or Contribute Finding for Greenhouse Gases Under Section 202(a) of the Clean Air Act" (Endangerment Finding) published by EPA on December 15, 2009, 74 Fed. Reg. 66,496 (Dec. 15, 2009). The new information raises questions as to the validity of a substantial segment of the data upon which the Endangerment Finding rests, which is of central relevance to the Endangerment Finding. Consequently, in response to this petition under 307(d), EPA must: (1) reconvene the regulatory proceeding, (2) provide the public with the opportunity to comment on the newly available information, and (3) provide such information to the Science Advisory Board (SAB) for its review and comment.

B. Summary of the Bases of this Petition

1. The Data from the University of East Anglia Climate Research Unit Has Been Put in Question by Those Who Developed the Data, and This Is of Central Relevance to the Endangerment Finding, Requiring EPA To Reconvene the Regulatory Proceeding and To Provide the Opportunity for Public Comment and SAB Review

On November 17, 2009, well after the close of the public comment period on the proposed Endangerment Finding, internal e-mails and documents from the Climate Research Unit (CRU) at the University of East Anglia were released to the public. The e-mails and documents, many of which were apparently authored by the CRU scientists working at CRU on the project (CRU scientists), discuss the manner in which CRU's global warming data (CRU Data) were developed, analyzed, and handled. Specifically, the e-mails and documents suggest that the CRU scientists questioned the reliability of their own data, the methodologies used in developing and analyzing such data, and the conclusions based thereon.

PLF's petition is premised on the fact that CRU scientists themselves have put in question the quality of their own data and analyses. Thus, the data and analyses were put in question by those in a position to know the strengths and weaknesses thereof.

The following two issues raised by this petition are of central relevance to the Endangerment Finding:

(1) whether the CRU Data are reliable and, if they are not,

(2) whether data from sources other than CRU can support the Endangerment Finding.

2. EPA must Provide SAB with the Opportunity To Review and Comment on the New Information

When EPA reopens the public comment period in connection with the reconsideration proceeding, it is imperative that EPA provide the newly discovered CRU e-mails and documents to the SAB pursuant to 42 U.S.C. § 4365. The SAB is a statutorily mandated scientific panel comprised of scientists who advise EPA regarding the scientific judgments upon which regulations and regulatory findings are based. Upon a careful investigation of the newly released e-mails and documents, SAB will be in a position to address the issue of whether the CRU Data can be relied upon. If SAB finds that the CRU Data are unreliable, SAB will also be in a position to address whether the remaining data in the record are sufficient to support the Endangerment Finding without reliance on the CRU Data.

3. Because CRU Used the Same Raw Data as the Other Organizations upon Whose Work EPA Relied, and Because the Other Organizations Apparently Reached Conclusions Similar to Those of CRU, an Inquiry Is Required as to Whether the Raw Data Are Suspect, the Methodologies of the Other Organizations Are Suspect, or Both

CRU and other organizations studying global climate change and its causes all use the same raw data, and scientists from the National Oceanic and

Atmospheric Administration (NOAA) have acknowledged that the three surface temperature data sets from NOAA, the National Aeronautics and Space Administration (NASA), and CRU are not independent but are based on a common set of underlying raw data.

If NASA and NOAA used the same raw data as CRU, and if they have reached conclusions similar to those of CRU, then a finding of data unreliability with regard to the CRU Data may indicate systemic problems with all three of the data sets upon which EPA relied in promulgating its Endangerment Finding. Because these issues are of central relevance to the Endangerment Finding, they must be addressed by EPA, the public, and SAB in a proceeding for the reconsideration of the finding.

C. This Petition Differs Substantially from Previous Petitions for Reconsideration of the Endangerment Finding

On December 23, 2009, Southeastern Legal Foundation (SLF) filed a petition for reconsideration of the Endangerment Finding, asserting that the newly released internal CRU e-mails indicate that the CRU Data was the subject of "systematic manipulation" and "collusion" by CRU scientists, arguing that reconsideration is appropriate because the public did not have an adequate opportunity to consider and comment on the data used to support the Endangerment Finding in light of the statements made by CRU scientists. Southeastern Legal Foundation, Petition for Reconsideration of

"Endangerment and Cause or Contribute Findings for Greenhouse Gases Under Section 202(a) of the Clean Air Act," (Dec. 23, 2009). PLF takes no position regarding the specific allegations of systemic manipulation and collusion in SLF's petition.

On October 5, 2009, the Competitive Enterprise Institute (CEI) filed a petition to reopen the public comment period based upon the admission by CRU "that it had destroyed the raw data for its data set of global surface temperatures." CEI Petition to Reopen This Proceeding on the Basis of New Evidence Concerning the Destruction of Data, and To Reopen the Comment Period for Public Response to This New Information, at 2 (Oct. 5, 2009). On December 2, 2009, CEI amended its petition on the basis that "several thousand e-mails and documents" from CRU show "conscious efforts by leading climate scientists to misrepresent or falsify data, to evade FOIA requests, and to inject intentional bias into the scientific peer review process." Supplement to the October 5th Petition of the Competitive Enterprise Institute to Reopen This Proceeding in Light of Newly Released Information (Dec. 2, 2009). CEI's amendment to the petition also suggests that admissions in the CRU e-mails show that "global warming largely ceased in the last 10-15 years, despite public claims that it was continuing." *Id.* PLF takes no position as to the specific allegations of fraud, willful destruction of data, and global warming trends in the CEI petition or its Supplement.

PLF's petition does not assume or even argue that the statements in the CRU e-mails or documents are true. PLF believes strongly that the issue of the veracity of the e-mails and documents should not be addressed by lawyers but should be carefully reviewed, examined, and analyzed by competent climatologists and other qualified scientists, and that, for the reasons set forth herein, such a review must be conducted in connection with a reconsideration petition, with full opportunity for public comment and SAB review.

D. Freedom of Information Act Follow-Up

PLF reserves the right to supplement this petition in light of forthcoming information to be obtained through our December 17, 2010, Freedom of Information Act, 5 U.S.C. § 552, *et seq.*, request for documents relating to EPA's peer review process and the treatment of data relied upon in the Endangerment Finding.

The following sections explain in detail the bases for this petition. Section II sets forth the legal standard governing petitions for reconsideration under the Clean Air Act (CAA). Section III explains why PLF's petition meets the standard. Section IV offers a brief conclusion as to why this petition for reconsideration should be granted.

II

STANDARD OF REVIEW

A. General Requirements

The CAA provides for judicial review of the Endangerment Finding for objections raised during the public comment period. In addition, the CAA provides that EPA's Administrator shall convene a proceeding for reconsideration of the Endangerment Finding where a person: (1) raises an objection, (2) within the time specified for judicial review, (3) of central relevance to the outcome of the Endangerment Finding proceedings, (4) on grounds arising after the comment period has expired. 42 U.S.C. § 7607(d)(7)(B).

B. Timeliness

In this case, PLF advances an objection within the sixty-day period for judicial review, which extends until February 16, 2010. *See* 74 Fed. Reg. at 66,496. As set forth in more detail below, PLF's objection rests upon newly available information pertaining to the credibility of the data upon which EPA relied in promulgating the Endangerment Finding. Since this newly released information was unavailable during the public comment period, PLF could not have commented upon it during that time. Accordingly, PLF's petition for reconsideration raises a timely objection. 42 U.S.C. § 7607(d)(7(B).

Therefore the petition should be granted if it raises questions of central relevance to the outcome of the Endangerment Finding. *Id.*

C. Central Relevance

An objection based on procedural grounds filed under 42 U.S.C. § 7607(d)(7)(B) is of central relevance if "there is a substantial likelihood that the rule would have been significantly changed if such errors had not been made." *Kennecott Corp. v. E.P.A.*, 684 F.2d 1007, 1017 (D.C. Cir. 1982). In *Kennecott*, the court held that EPA's failure to abide by procedural rules was of central relevance to the outcome of EPA's decisionmaking process. In that case, EPA had failed to place "in the docket for public comment certain data relating to the basis" of a promulgated rule. EPA asserted that there was no need to reopen the public comment period because, even if the data in question were included in the docket, EPA would have come to the same conclusion inevitably. In reviewing EPA's failure to grant the petition, the court determined that in "all events, absence of those documents, or of comparable materials" shedding light on the underlying basis of the rule, "makes impossible any meaningful comment on the merits of EPA's assertions" that EPA did not rely on the documents omitted from the record. *Id.* at 1018. "In all the circumstances, EPA's failure to include such documents constitutes reversible error, for the uncertainty that might be clarified by those documents with respect to the test, the data, and the methodology used by EPA indicates

a 'substantial likelihood' that the regulations would 'have been significantly changed.'" *Id.* at 1019-20. Under the *Kennecott* formulation, in connection with a petition for reconsideration under 42 U.S.C. § 7(d)(7)(B), there is a substantial likelihood that a regulation would be changed if information not available in the rulemaking record for public comment casts substantial uncertainty over the final regulation. In such cases EPA must grant a timely filed petition for reconsideration.

Thus, after *Kennecott*, EPA cannot excuse its failure to include documents in a rulemaking record merely by asserting that the documents would not have made a difference in its final decision. Indeed, *Kennecott* dismissed such an assertion as "disingenuous." *Id.* at 1018. Consequently, failure to allow for public comment on documents which cast substantial uncertainty over a final CAA regulation cannot be excused by EPA's self-serving assertion that its decision would have remained the same had public comment been allowed.

D. Role of SAB

Importantly, where substantial uncertainty surrounds EPA's scientific determinations, warranting the reopening of the public comment period, SAB review is required by 42 U.S.C. § 4365(c)(1) ("The Administrator . . . shall make available to the Board such proposed criteria document, standard, limitation, or regulation, together with relevant scientific and technical

information in the possession of the Environmental Protection Agency on which the proposed action is based."). By statute, SAB review is necessary in order to ensure public confidence in EPA's decisions and, indeed, as acknowledged in EPA's regulations, SAB was created as an independent agency explicitly to give an assurance of credibility to EPA's scientific determinations. 40 C.F.R. § 1.25(c) (1999) (SAB's mission is to provide "expert and independent advice to the [EPA] on the scientific and technical issues facing the Agency" and to assist EPA "in identifying emerging environmental problems."); *id.*, *see also* Joe G. Conley, *Conflict of Interest and the EPA's Science Advisory Board*, 86 Tex. L. Rev. 165, 168 (2007) ("Congress established the EPA Science Advisory Board in 1978 to provide independent scientific and technical advice to EPA."). Accordingly, 42 U.S.C. § 4365 requires EPA to provide SAB with all relevant technical information and data in connection with any rulemaking proceeding for SAB's independent review whenever the public comment period is open. *Id.*

III

ARGUMENT

A. This Petition Is Timely Filed Because Release of the CRU Data Occurred after the Public Comment Period Closed, and Consequently, the Issues Could Not Have Been Raised During the Public Comment Period

Under Section 307(d)(7)(B) of the CAA, a petitioner rasing an objection to the rule must demonstrate that the petition was filed within the time specified for judicial review and that it was either: (1) impracticable to raise the objection within the prescribed comment period, or (2) the grounds for such objection arose after the end of the comment period. 42 U.S.C. § 7607(d)(7)(B). Here, it was impossible to raise objections to the Endangerment Finding based upon the CRU e-mails and documents during the public comment period, because they were not released until mid-November, 2009, well after the time for public comment had closed.

1. Procedural History of the Endangerment Finding

In response to the Supreme Court's decision in *Massachusetts v. EPA*, 549 U.S. 497 (2007), EPA began examining the scientific and technical basis for an endangerment finding concerning greenhouse gas emissions. 74 Fed. Reg. at 66,500 (Dec. 15, 2009). On April 24, 2009, EPA responded to the Court's decision by proposing a finding that greenhouse gases emitted by new motor vehicles contribute to air pollution that may endanger public

health or welfare. 74 Fed. Reg. 18,886 (Apr. 24, 2009). The proposed finding was subject to a sixty-day public comment period, which ended on June 23, 2009. 74 Fed. Reg. at 66,500.

After the public comment period ended, but before EPA finalized its endangerment finding, two important developments occurred. The first development was the admission by CRU that it either lost or destroyed all of the original data that would allow a third party to reconstruct and verify CRU's global temperature record.[1]

As a result of that disclosure by CRU, a petition asking EPA to reopen the public comment period was made on October 5, 2009. EPA's Response to Public Comments, Volume 11, at 2. The EPA did not reopen the public comment period.

The second development occurred just over a month later. On November 17, 2009, an electronic file containing over a thousand e-mails and documents from CRU, began to appear on the internet.[2]

---

[1] Andrew Orlowski, *Global Warming Ate My Data*, The Register, Aug. 13, 2009, *available at* http://www.theregister.co.uk/2009/08/13/cru_missing/ (last visited on Jan. 25, 2010).

[2] Jonathan Leake, *The Great Climate Change Science Scandal*, The Sunday Times, Nov. 29, 2009, *available at* http://www.timesonline.co.uk/tol/news/environment/article6936289.ece (last visited on Jan. 25, 2010); Andrew C. Revkin, *Hacked E-Mail Is New Fodder for Climate Dispute*, N.Y. Times, Nov. 20, 2009, *available at* http://www.nytimes.com/2009/11/21/science/earth/21climate.html?_r=3&hp (last visited on Jan. 25, 2010).

As a result of the released information, on December 2, 2009, a supplement to the October, 2009, petition to reopen was submitted to EPA. The supplement to the petition alleged that the recent disclosure of released e-mail messages and documents from CRU undermined the Intergovernmental Panel on Climate Change (IPCC) science and assessment process upon which the EPA's Technical Support Document and the Findings primarily relied. *Id.* The EPA did not reopen the public comment period.

The EPA's proposed Endangerment Finding was finalized and signed by the Administrator on December 7, 2009, just five days after the submission of the December 2, 2009, supplement to the October, 2009, petition. The Endangerment Finding was published in the Federal Register on December 15, 2009. 74 Fed. Reg. 66,496.[3]

In rejecting the petition to reopen the public comment period, EPA responded that "it appears that the scientific issues raised in the e-mails were also raised in public comments." EPA's Response to Public Comments, Volume 11, at 2. Consequently, EPA asserted that it had "addressed many of the issues raised in the hacked e-mails, among the hundreds of issues raised by commenters." *Id.*

---

[3] *See* Endangerment and Cause or Contribute Findings for Greenhouse Gases under the Clean Air Act, *available at* http://www.epa.gov/climatechange/endangerment.html (last visited Jan. 25, 2010) (Administrator signed findings on Dec. 7, 2009).

2. EPA's Insufficient Response to the CRU Disclosures

As indicated, the material from CRU, consisting of over 1,000 e-mails and over 3,000 documents, were released on or about November 17, 2009. Despite EPA's assertion that the e-mails appear to present the same scientific issues contained in the public comments, the fact remains that these materials were not available during the Endangerment Finding's public comment period, which closed on June 23, 2009. Additionally, EPA's equivocal statement about what the e-mails may contain demonstrates that in the five-day period between the supplement to the petition and the signing of the Endangerment Finding, the EPA did not, and could not have, thoroughly examined the thousands of e-mails and documents released from CRU.

Contrary to EPA's assertion, the CRU e-mails and documents do not address issues raised during the comment period. Rather, as set forth in more detail below, they constitute a large volume of documentation authored in large part by the CRU scientists themselves, putting in question the reliability of the CRU Data. This voluminous documentation was wholly absent from the record during the public comment period. Indeed, as detailed above, the documentation could not have been included in the record because it was not made available to EPA until well after the public comment period had closed. Respectfully, this is precisely the type of new information contemplated by

Section 307(d)(7)(B) of the CAA as the basis for requiring EPA to reconvene a regulatory proceeding. 42 U.S.C. § 7607(d)(7)(B).

For the foregoing reasons, this petition was timely filed and is based on new information that could not have been addressed during the public comment period.

B. The Information in the E-mails from CRU Are of Central Relevance to the Outcome of the Endangerment Finding

1. The Nature of the Released CRU E-mails and Documents

The release of e-mails and documents from CRU occurred on or about November 17, 2009, when an apparent whistleblower leaked one large computer file containing thousands of e-mails and documents to various internet sites. *See* Terry Hurlbut, *Who leaked the Hadley CRU files and why*, Essex County Conservative Examiner, Nov. 21, 2009 ("The timeline begins on November 17, when the user named 'FOIA' left this comment at The Air Vent site: We feel that climate science is, in the current situation, too important to be kept under wraps. We hereby release a random selection of correspondence, code, and documents. Hopefully it will give some insight into the science and the people behind it."). These e-mails and documents were

possibly key to a Freedom of Information request submitted to CRU, because the title of the released file is FOI2009.zip.[4]

Because CRU scientists themselves have raised questions in the e-mails and documents regarding the quality of their own data and analyses, the data and analyses were put in question by those in a position to know the strengths and weaknesses thereof. Further, because the CRU Data were considered by EPA in connection with its Endangerment Finding, it is important to determine in a duly reconvened proceeding for reconsideration the extent to which the CRU Data was compromised and whether the CRU Data may properly be relied upon as a basis for the Endangerment Finding.

If it is determined that the CRU Data are unreliable, it must further be determined whether the Endangerment Finding can stand without reliance on such data. Of course, such a question cannot be resolved by a mere conclusory assertion by EPA that it would have come to the same conclusion without the CRU Data. Although EPA has made such an assertion in the record (*see* EPA's Response to Public Comments, Volume 2, Response 2-39, Dec. 15, 2010), the United States Court of Appeals for the District of Columbia Circuit has held that a similar assertion made by EPA in connection with an earlier

---

[4] Terry Hurlbut, *Hadley CRU Hacked With Release of Hundreds of Docs and E-mails*, Essex County Conservative Examiner, Nov. 19, 2009, *available at* http://www.examiner.com/x-28973-Essex-County- Conservative-Examiner ~y2009m11d19-Hadley-CRU-hacked-with-release-of-hundreds-of-docs-and-e-mails (last visited on Jan. 26, 2010).

CAA rulemaking was "disingenuous"and insufficient grounds to deny a petition for reconsideration, as set forth in more detail above. *See Kennecott Corp. v. E.P.A.*, 684 F.2d at 1017.

2. The Contents of the Released CRU E-mails and Documents

The contents of FOI2009.zip hold one top-level directory, ./FOIA.[5] That directory is divided into two main directories, ./mail and ./documents. The ./mail directory contains 1073 text files ordered by date. The text files are named in order with increasing but not sequential numbers. Each text file holds the body of an e-mail. The ./documents directory contains documents in MS Word and Adobe Acrobat format, computer code, and other data in several sub-directories.

The e-mails suggest at least four areas of concern that SAB should be given the opportunity to address during a rigorous public examination of the files. First, the e-mails appear to call into question the credibility of the peer review process. The authors of the e-mails seem to have cooperated covertly to ensure that only papers favorable to their position on global warming were published, and that editors and journals publishing contrary papers were punished. They also appear to show attempts to "discipline" scientists and

---

[5] The contents of FOIA2009.zip are now available on the internet at various websites. *See, e.g.*, http://www.eastangliae-mails.com; http://www.climate-gate.org; http://www.assassinationscience.com/climategate/1/FOIA/mail/ (last visited Jan. 26, 2010).

journalists who published skeptical information. *See* e-mails 1047388489, 1256765544, 1255352257, 1051190249, 1210367056, 1249503274, 1054756929, 1106322460, 1132094873, and 1139521913.

Second, the e-mails may suggest that the authors manipulated and "massaged" the data to strengthen the case in favor of unprecedented global warming, and to suppress their own data if it called global warming into question. *See* e-mails 0938018124, 0843161829, 0939154709, 0942777075, and 1059664704.

Third, the e-mails may suggest that the authors attempted to prevent data from being made available to other researchers by hindering data archiving requests under the Freedom of Information Acts of both the United States and the United Kingdom. *See* e-mails 1106338806, 1228330629, 1212063122, 1210367056, and 1107454306. Significantly, *The Wall Street Journal* reported recently that the United Kingdom's Information Commissioner's Office determined that the CRU violated the British Freedom of Information Act by "failing to comply with requests for raw data about global warming." *See* Guy Chazan, *U.K. Says University Broke Law on Turning Over Data*, Wall St. J., Jan. 29, 2010, at A8. The article also stated that Sir Muir Russell, a former British civil servant, is leading an examination of the e-mails to "check for evidence that data were handled in ways that are 'at odds with acceptable scientific practice and may therefore call into question

any of the research outcomes.'" *Id.* Revelation of this information indicates that EPA was incorrect in asserting that "the science on which the Administrator has based her determinations regarding the endangerment of both public health and welfare . . . have been fully open and transparent." EPA's Response to Public Comments, 11-2.

Fourth, the "HARRY READ ME" file may suggest that the data selected, used, and analyzed by CRU not only had substantial gaps but was also potentially irreparably compromised. *See* Appendix A, 10-21.

Appendix A of this petition provides specific examples of the content of the CRU e-mails and other newly identified information. Because of the bulk of the documentation, the sampling is not intended to be inclusive, but we have reproduced what we believe are representative examples of the content of the statements in the four specific categories of information set forth above, namely: (1) credibility of the peer review process, (2) adjustment of data to produce a desired outcome, (3) obstruction of freedom of information requests, and (4) the "HARRY READ ME" file. As can be ascertained by a review of Appendix A, the e-mails and documents authored by the CRU scientists question the reliability of the CRU Data and analyses.

3. The CRU E-mails and Documents Raise Questions of Central Relevance to the Outcome of the Endangerment Finding

a. There Is a Substantial Likelihood That EPA May Change Its Decision Because Fundamental Questions Have Been Raised by the New Information Regarding the Reliability of the Data upon Which EPA Based Its Endangerment Finding

In response to a petition under CAA Section 307(d)(7)(B), EPA must convene a proceeding for reconsideration of a rulemaking when the petitioner raises an objection of central relevance to the outcome of the decision. The credibility of the data upon which EPA relied in promulgating its Endangerment Finding is of central relevance to the outcome of the finding because there is a substantial likelihood that the finding would have been significantly changed had the revelations regarding the CRU Data been made prior to the close of the comment period on the proposed Endangerment Finding. *See Kennecott Corp. v. E.P.A.*, 684 F.2d at 1017; 42 U.S.C. § 7607(d)(7)(B). Had such revelations been included in the record, they would have provided the public and SAB with substantial additional opportunities to review the manner in which the data were developed, handled, and analyzed.

Because the leaked CRU e-mails were written by scientists working on the CRU Data, and because they address specifically what the scientists did (and did not do) with their own data, they merit special attention, since they

may be interpreted as admissions by the CRU scientists that the CRU data is not reliable. *See* Appendix A, 1-21.

This is all the more true given the fact that the e-mails suggest that the CRU Data may have been falsified. The scientists referred to "false codes" and running programs to "allow bad databases to pass unnoticed, and good data bases to become bad," *see* Appendix A, 15-16, leading to the possible inference that the CRU scientists may have falsified or manipulated the data. The fact that the British Information Commissioner's Office has found that CRU violated British law by failing to make disclosures under British freedom of information laws lends substantial credence to the proposition that CRU engaged in at least some wrongdoing in connection with the CRU Data. *See* Guy Chazan, *U.K. Says University Broke Law on Turning Over Data*, Wall St. J., Jan. 29, 2010, at A8. *See also* David Derbyshire, *Scientists broke the law by hiding climate change data: But legal loophole means they won't be prosecuted*, Daily Mail, Jan. 30, 2010.[6]

Further, the CRU scientists working on the project made the statements in written communications which they likely believed would remain private and confidential. Thus, the statements were not only made by the very scientists who developed and analyzed the CRU Data, but they were made

---

[6] *Available at* http://www.dailymail.co.uk/news/article-1246661/New-scandal-Climate-Gate-scientists-accused-hiding-data-global-warming-sceptics.html (last visited Jan. 30, 2010).

under circumstances which tend to assure their credibility, because the scientists would have made their communications privately, expecting them to remain so.

Moreover, the statements may be considered to be statements against self-interest, because they suggest that the scientists may have falsified CRU Data, which would not only reflect poorly upon their professional reputations, but could also reflect poorly upon CRU and the University of East Anglia. An analogy to the law of evidence is appropriate. Statements made against one's own interest are deemed inherently credible and therefore are not subject to exclusion from evidence under the hearsay rule. *See* Kenneth S. Broun, *Determining What is Against Interest; confrontation problems*, 2 McCormick on Evid. § 319 (6th ed.) ("[T]he theory underlying the hearsay exception for declarations against interest is that people do not make statements that are harmful to their interests without substantial reason to believe that the statements are true."). Given the circumstances in which the statements were made, there is a significant possibility that the authors of the e-mails believed that what they wrote was true.

Unmistakably, the quality of the underlying data and the methodologies relied upon in formulating the scientific conclusions forming the bases of the Endangerment Finding necessarily are of central relevance to the finding. Moreover, under *Kennecott*, an objection is of central relevance to a decision

under the CAA when "there is a substantial likelihood that the rule would have been significantly changed" if the errors pointed out in a petition for reconsideration had been addressed. *Kennecott* makes clear that the substantial likelihood test does not require a showing of near certainty. In *Kennecott*, the court found that there was a substantial likelihood that EPA would significantly change its decision because EPA failed to include certain data in the rulemaking record. This failure created uncertainty that could only be clarified by including the data in the record and reopening the public comment period. As such, *Kennecott* demonstrates that the substantial likelihood test is met when EPA's failure to abide by a required procedure raises a substantial uncertainty as to whether EPA's decision may have been made in error.

Of course, the substantial likelihood test requires more than unfounded speculation that EPA may change its decision, *see, e.g.*, *Husqvarna AB v. E.P.A.*, 254 F.3d 195, 203 (D.C. Cir. 2001) (Petitioner failed to show that alleged procedural error was serious enough to create a substantial uncertainty or likelihood that EPA would change its position). In the instant case, however, the allegations regarding the unreliability of the CRU Data were made by the very scientists who developed that data. *See* Appendix A. *See also* Guy Chazan, *U.K. University in Climate Controversy Broke Law, Officials Say*, Wall St. J., Jan. 29, 2010, at A8. The statements of such scientists directly reflect on the data and raise substantial questions as to its

quality. Axiomatically, the quality of the data upon which a regulatory decision is based is of central relevance to the regulatory decision. Consequently, in response to this petition, EPA must reconvene the regulatory proceeding so that EPA, the public, and the SAB may have the opportunity to address the issue of whether the newly revealed e-mails and documents undercut the reliability of the data upon which the Endangerment Finding rests.

It is true that "Congress was concerned that EPA's rulemaking not be casually overturned for procedural reasons," *Sierra Club v. Costle*, 657 F.2d 298, 391 (D.C. Cir. 1981), but the substantial likelihood test has never been understood to require a showing of near certainty. Rather, reconsideration is required where a failure to reconsider would raise substantial uncertainty regarding what the outcome would have been had reconsideration occurred. *Kennicott*, 684 F.2d at 1019-20 (Substantial likelihood is demonstrated where uncertainties may change resolution of the decision.).

Certainly where EPA fails to abide by a procedure which would have little or no effect on its final decision, its procedural error is harmless, and the substantial likelihood test is not met. In the instant case, however, the ground upon which EPA's decision rests, namely, the data underlying the Endangerment Finding, has been put into question by those who developed it. This poses substantial uncertainty regarding whether the Endangerment

Finding would have been made had the newly discovered information been fully and openly addressed by the public, SAB, and EPA. Under these circumstances, reconvening the regulatory proceeding may well result in a change to the Endangerment Finding if the statements made by the scientists about the unreliability of their own data are true. Thus, if the data supporting the Endangerment Finding are unreliable, the finding itself cannot stand, and failure to reconvene the regulatory proceeding in light of such new information would constitute reversible error under *Kennecott*.

b. EPA Cannot Rest Its Endangerment Finding on NOAA and NASA Datasets Without Further Review and Public Comment

EPA has asserted that the Endangerment Finding may be independently supported by data gathered and analyzed by NASA and NOAA. *See* Endangerment and Cause or Contribute Finding for Greenhouse Gases Under Section 202(a) of the Clean Air Act: EPA's Response to Public Comments, Vol. 11, 11-2 (rebutting a commenters's assertion that the CRU Data is unique, and explaining that, in addition to the CRU Data, EPA relied upon two similar data sets, namely, those of NASA and NOAA). But EPA's assertion is not supported by the facts.

According to Dr. Roger A. Pielke, Sr., a University of Colorado climatologist, and Dr. Richard Anthes, President of the University Corporation for Atmospheric Research, CRU and other organizations studying global climate change and its causes all use the same underlying raw data. *See* Roger A. Pielke Sr., et al., *Unresolved issues with the assessment of multidecadal global land surface temperature trends*, J. Geophysical Research, Vol. 112 at 13[7] (2007) ("The raw surface temperature data from which all the different global surface temperature trend analyses are derived are essentially the same. The best estimate that has been reported is that 90-95% of the raw data in each of the analyses is the same. That the analyses produce similar trends should therefore come as no surprise.") (citing a leaked 2003 E-Mail from Phil Jones). *See also* Dr. Richard Anthes, President of the University Corporation for Atmospheric Research, Testimony Before Congress (Mar. 2009).

NOAA scientists have also acknowledged that the three surface temperature data sets from NOAA, NASA, and CRU are not independent. In the publication *Temperature Trends in the Lower Atmosphere: Steps for Understanding and Reconciling Differences*, written for the U.S. Climate

---

[7] *Available at* http://pielkeclimatesci.files.wordpress.com/2009/10/r-321.pdf (last visited Jan. 28, 2010).

Change Science Program,[8] NOAA scientists describe the data forming the bases of the findings:

> The global surface air temperature data sets used in this report are to a large extent based on data readily exchanged internationally, e.g., through CLIMAT reports and the WMO publication Monthly Climatic Data for the World. Commercial and other considerations prevent a fuller exchange, though the United States may be better represented than many other areas. In this report, we present three global surface climate records, created from available data by NASA Goddard Institute for Space Studies, NOAA National Climatic Data Center, and the cooperative project of the U.K. Hadley Centre and the Climate Research Unit of the University of East Anglia (HadCRUT2v). These are identified as TS-NASA, TS-NOAA and TS-HadCRUT2v respectively.

Id. at 47. The report further describes that the three main groups creating analyses of global surface temperature differ in "the choice of available data that are utilized as well as the manner in which these data are synthesized." *Id.* at 50. But the report clarifies that the data sets are not independent:

> Since the three chosen data sets utilize many of the same raw observations, there is a degree of interdependence. Nevertheless, there are some differences among them as to which observing sites are utilized.

Id. at 51. Thus, EPA's contention that its scientific conclusions on global warming are supported by two data sets (NOAA and NASA) that are

---

[8] *Temperature Trends in the Lower Atmosphere: Steps for Understanding and Reconciling Differences*, A Report by the Climate Change Science Program and the Subcommittee on Global Change Research, Washington, DC. (Thomas R. Karl, Susan J. Hassol, Christopher D. Miller, and William L. Murray, eds., 2006), *available at* http://www.climatescience.gov/Library/sap/sap1-1/finalreport/sap1-1-final-chap3.pdf (last visited Jan. 28, 2010).

independent of the CRU Data (*see* EPA's Response to Public Comments, Vol. 11, 11-2) simply is incorrect.

If NASA and NOAA used substantially the same raw data used by CRU, and if they reached similar conclusions to those reached by CRU, then a finding of data unreliability with regard to the CRU Data may indicate systemic problems with all three of the data sets upon which EPA relied in promulgating its Endangerment Finding. Curiously, EPA admits that the NOAA and NASA data sets show similar trends and were developed at least in part through similar methodologies to those employed by CRU. By making this admission, EPA unintentionally leads us to the inexorable conclusion that, if the CRU Data are methodologically problematic, the NOAA and NASA data may share such methodological problems. *See* EPA's Response to Public Comments, Vol. 2, 2-28 ("The NOAA global surface temperature dataset employs the same methodology for addressing urbanization as is used in the HadCRUT [CRU Data] . . . ."). Thus, any finding that the CRU Data and its conclusions are unreliable should lead to a review of the NOAA and NASA data sets as well.

In short, having used substantially the same raw data, and having employed similar methodologies, the three data sets may suffer from similar errors, and a finding that the CRU Data are unreliable may indicate methodological problems that apply to the NOAA and NASA data as well,

warranting further review of all of the data upon which the Endangerment Finding is based. As such, the questions raised by the leaked CRU e-mails and documents raise significant further issues as to all three data sets relied upon in the Endangerment Finding, suggesting that EPA's peer review, including specifically SAB's review, were insufficient during the promulgation of the Endangerment Finding because such issues had not yet been brought to light during the public comment period.

c. Past EPA Peer Review and Past SAB Review Were Insufficient in Light of New Information

EPA asserts, in the Endangerment Finding's Technical Support Document (TSD), that the scientific conclusions upon which EPA rests its Endangerment Finding have been vetted through an internal and external peer review process. *See* Technical Support Document for Endangerment and Cause or Contribute Finding for Greenhouse Gases Under Section 202(a) of the Clean Air Act, 4 (Dec. 7, 2009). Given the fact that the new information draws into question the reliability of the CRU Data and raises questions as to the other data sets, these peer review processes must be reconvened.

Only further peer review in light of the new information will adequately address the issues raised by the newly released e-mails and documents. The substantial uncertainties raised by the new information indicate that further review by SAB, as well as further public comment, may

lead EPA to question the reliability of the data underlying the Endangerment Finding and, as a result, potentially to withdraw the finding.

Where questions are raised as to the validity of EPA's scientific assessments, it is essential that EPA allow for independent review from the SAB, because Congress specifically created the SAB to ensure that EPA's scientific decisions would be subject to such independent review, so as to give an assurance of reliability. 40 C.F.R. § 1.25(c) (1999). Any SAB review that may have occurred before or during the comment period on the proposed rulemaking remains incomplete at this point because SAB has not been given an opportunity to review the Endangerment Finding in light of the pertinent new information. Therefore SAB has not been able to perform its statutory function as an independent scientific review panel.

d. EPA's Responses Have Been Insufficient To Assure the Reliability of the Data upon Which the Endangerment Finding Was Based

To date, EPA has dismissed public concerns over the leaked CRU e-mails and documents by its single, abbreviated response to a comment made by the Competitive Enterprise Institute (CEI). *See* EPA Response to Public Comments on the Endangerment Finding, Volume 11, Response 11-2). The comment was made after the close of the public comment period, and was allowed only as a supplement to CEI's previously filed petition for reconsideration. *Id.* CEI's original petition for reconsideration was based

upon concerns arising out of a news release that the CRU Data had been lost. *Id.* Only CEI was given the opportunity to comment upon the newly leaked CRU e-mails and documents, which shed additional light upon the subject of its petition for reconsideration. Accordingly, neither the general public nor SAB have had the opportunity to comment on the new information contained in the leaked CRU e-mails and documents.

EPA responded to CEI by asserting that, in its view, the leaked e-mails and documents prove nothing, and that, even if the CRU Data has been compromised, EPA may rely upon the other two data sets. Yet, for the reasons set forth above, the questions pertaining to the credibility of the CRU Data also implicate the other two data sets upon which EPA relied in making its Endangerment Finding.

EPA's response to another comment, Comment 2-39, is along the same lines. EPA stated that it did not need to rely on the CRU Data to support the Endangerment Finding because the finding could rest entirely upon the NOAA and NASA data sets. As indicated, however, because NOAA and NASA used substantially the same underlying raw data, and because they used methodologies similar to that of CRU, the new information raises questions about those data sets as well. Consequently, as set forth in more detail above, EPA's response is insufficient. *Kennecott*, 684 F.2d at 1019-20.

4. Because the CRU E-mails and Documents Have Stirred Public Controversy over the Endangerment Finding, EPA Should Reconvene the Regulatory Proceeding To Restore Public Confidence

The leaked CRU e-mails and documents have led to public concern that the CRU Data was falsified in order to manipulate CRU's scientific findings to substantiate predetermined conclusions. *See, e.g.*, Southeastern Legal Foundation, Petition for Reconsideration of "Endangerment and Cause or Contribute Finding for Greenhouse Gases Under Section 202(a) of the Clean Air Act" (Dec. 23, 2009) (endorsed by nine Congressmen). While PLF does not take a position as to the truth of the assertions in the Southeastern Legal Foundation's petition, PLF notes that the leaked e-mails and documents sparked substantial public controversy over the reliability of the CRU Data and EPA's Endangerment Finding. *See, e.g., EPA Issues Greenhouse Gas Warnings Despite Concerns Over Leaked E-Mails*, Fox News, Dec. 7, 2009[9] (attributing the Associated Press as a contributor to the story).

The controversy, suspicions, and doubt cast by the leaked e-mails and documents have been dubbed "climategate" in the popular media. *Id.* Coverage of the story has publicized the controversy and amplified public concern as to the reliability of EPA's conclusions on global warming and its Endangerment Finding. *See, e.g.*, Gerald Warner, *US Congress Investigates*

---

[9] *Available at* http://www.foxnews.com/politics/2009/12/07/republicans-slam-epa-decision-declare-public-health-danger/ (last visited Jan. 28, 2010).

Climategate e-mails: this could be the beginning of the end for AGW,[10] Telegraph.co.uk, Nov. 26, 2009 (discussing public reaction to news of the released CRU e-mails). For these reasons, not only is EPA legally obligated to reconvene the regulatory proceeding in light of this petition, but EPA should do so also to allay public fears that the Endangerment Finding is based upon unreliable data.

It is appropriate to note that some of the scientists who wrote the leaked CRU e-mails have publically denied assertions that the CRU Data was falsified or manipulated; however, they concede making the disclosed statements. *See* Guy Chazan, *U.K. University in Climate Controversy Broke Law, Officials Say*, Wall St. J., Jan. 28, 2010, at A8[11] ("Mr. Jones has said he stands by his data and said the e-mails had been taken "completely out of context."). The CRU scientists have stated that those levying assertions of data falsification have selected e-mails to give a incorrect sense of what was being discussed. *Id.* Nevertheless, it is undeniable that the e-mails and documents have stirred substantial public controversy over the legitimacy of EPA's Endangerment

---

[10] *Available at* http://blogs.telegraph.co.uk/news/geraldwarner/100017954/us-congress-investigates-climategate-e-mails-this-could-be-the-beginning-of-the-end-for-agw/ (last visited Jan. 28, 2010).

[11] *Available at* http://online.wsj.com/article/SB10001424052748704194504575031022338013284.html?mod=googlenews_wsj (last visited Jan. 28, 2010)

Finding. Reconsideration is necessary in order to address the public concerns, and to restore public confidence.

Ultimately, EPA's Endangerment Finding may or may not be supportable based on the underlying science. As indicated, that is a question for scientists to resolve, not lawyers. To ensure the integrity of EPA's decision making process, however, the public controversy generated by the leaked e-mails and documents must be addressed by EPA in a reconvened administrative proceeding that allows for a thorough ventilation of the issues raised, especially the opportunity for SAB review. Only after the newly raised issues have been properly addressed can the legitimacy of the Endangerment Finding be resolved under 42 U.S.C. § 7607(d).

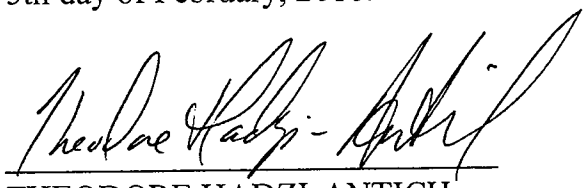
IV

CONCLUSION

PLF hereby petitions EPA under CAA 307 (d) to convene a proceeding for the reconsideration of the Endangerment Finding so that the public and SAB are afforded the opportunity to review and comment on the record regarding the new information calling into question the reliability of data upon which EPA based its Endangerment Finding. The questions raised in light of the new information pose substantial uncertainty regarding whether EPA would have made the Endangerment Finding had there been a full ventilation of the CRU e-mails and documents during the public comment period on the

proposed rulemaking. As a consequence, PLF's objection is of central relevance to the finding and, therefore, EPA must grant PLF's petition for reconsideration. *Kennecott*, 684 F.2d at 1013.

Respectfully submitted this 5th day of February, 2010.



THEODORE HADZI-ANTICH
LUKE WAKE
RALPH W. KASARDA

Attorneys for Petitioner
Pacific Legal Foundation

APPENDIX A

EXCERPTS FROM THE NEWLY DISCOVERED CRU INFORMATION SUPPORTING PLF'S PETITION FOR RECONSIDERATION[1]

1) Credibility of the Peer Review Process

The e-mail string initiated on March 11, 2003 (1047388489.txt), by Dr. Phil Jones concerns a peer-reviewed paper authored by Willie Soon and Sallie Baliunas, of the Harvard-Smithsonian Center for Astrophysics, which had been published in the journal, Climate Research.[2] Dr. Jones writes:

> Tim Osborn has just come across this. Best to ignore probably, so don't let it spoil your day. I've not looked at it yet. It results from this journal having a number of editors. The responsible one for this is a well-known skeptic in NZ. He has let a few papers through by Michaels and Gray in the past. I've had words with Hans von Storch about this, but got nowhere.

In a subsequent e-mail from this series, Dr. Jones writes:

> Writing this I am becoming more convinced we should do something ... I will be emailing the journal to tell them I'm having nothing more to do with it until they rid themselves of this troublesome editor. A CRU person is on the editorial board, but papers get dealt with by the editor assigned by Hans von Storch.

In the e-mail string begun on April 24, 2003 (1051190249.txt), Tim Carter, a research professor at the Finnish Environment Institute, states:

> On the CR issue ... I wonder if a review of the refereeing policy is in order. The only way I can think of would be for all papers to go through two Editors rather than one, the former to have overall responsibility, the latter to provide a second opinion on a paper and reviewers' comments prior to publication. A General Editor would be needed to adjudicate in the event of disagreement. Of course, this could then slow down the review process enormously. However, without an editorial board to vote

---

[1] The excerpts are provided without editorial comment but are described by the Petitioner Pacific Legal Foundation in order to put them into context. They have been divided by the Petitioner Pacific Legal Foundation into four categories as follows: (1) credibility of the peer review process; (2) the adjustment of data to produce a desired outcome; (3) obstruction of freedom of information requests; and (4) the HARRY READ ME File.

[2] Willie Soon & Sallie Baliunas, *Proxy Climatic and Environmental Changes of the Past 1000 Years*, Climate Res., Vol. 23:89-110 (2003), *available at* http://www.int-res.com/articles/cr2003/23/c023p089.pdf (last visited Jan. 27, 2010).

someone off, how can suspect Editors be removed except by the Publisher (in this case, Inter-Research).

Tom Wigley, a former CRU director, replies:

> Re Climate Research, I do not know the best way to handle the specifics of the editoring. Hans von Storch is partly to blame—he encourages the publication of crap science "in order to stimulate debate". One approach is to go direct to the publishers and point out the fact that their journal is perceived as being a medium for disseminating misinformation under the guise of refereed work. I use the word "perceived" here, since whether it is true or not is not what the publishers care about—it is how the journal is seen by the community that counts.

Mr. Wigley states further:

> I think we could get a large group of highly credentialed scientists to sign such a letter—50+ people. Note that I am copying this view only to Mike Hulme and Phil Jones. Mike's idea to get editorial board members to resign will probably not work—must get rid of von Storch too, otherwise holes will eventually fill up with people like Legates, Balling, Lindzen, Michaels, Singer, etc. I have heard that the publishers are not happy with von Storch, so the above approach might remove that hurdle too.

The e-mail series initiated on January 20, 2005 (1106322460.txt), by Steve Mackwell, Editor in Chief of Geophysical Research Letters, to Mike Mann, concerns Dr. Mann's complaint that he was not able to review a manuscript, critical of his own work, prior to its publication:

> Dear Prof. Mann
>
> In your recent email to Chris Reason, you laid out your concerns that I presume were the reason for your phone call to me last week. I have reviewed the manuscript by McIntyre, as well as the reviews. The editor in this case was Prof. James Saiers. He did note initially that the manuscript did challenge published work, and so felt the need for an extensive and thorough review. For that reason, he requested reviews from knowledgeable scientists. All three reviews recommended publication.
>
> While I do agree that this manuscript does challenge (somewhat aggressively) some of your past work, I do not feel that it takes a particularly harsh tone. On the other hand, I can understand your reaction. As this manuscript was not written as a Comment, but rather as a full-up scientific manuscript, you would not in general be asked to look it over. And I am satisfied by the credentials of the reviewers. Thus, I do not feel that we have sufficient reason to interfere in the timely publication of this work.

As part of the e-mail chain, Dr. Mann forwards this response to a number of his colleagues, including those at CRU:

> Dear All,
>
> Just a heads up. Apparently, the contrarians now have an "in" with GRL. This guy Saiers has a prior connection w the University of Virginia Dep of Environmental Sciences that causes me some unease.
>
> I think we now know how the various Douglass papers w Michaels and Singer, the Soon paper, and now this one have gotten published in GRL.

Tom Wigley responded in his e-mail:

> This is truly awful. GRL has gone downhill rapidly in recent years. I think the decline began before Saiers. I have had some unhelpful dealings with him recently with regard to a paper Sarah and I have on glaciers—it was well received by the referees, and so is in the publication pipeline. However, I got the impression that Saiers was trying to keep it from being published.
>
> Proving bad behavior here is very difficult. If you think that Saiers is in the greenhouse skeptics camp, then, if we can find documentary evidence of this, we could go through official AGU channels to get him ousted. Even this would be difficult.

In the e-mail of August 5, 2009 (1249503274.txt), Grant Foster requests the names of five experts who can review his paper, as required by the Journal of Geophysical Research. That journal's requirement is:

> Please list the names of 5 experts who are knowledgeable in your area and could give an unbiased review of your work. Please do not list colleagues who are close associates, collaborators, or family members.

In his response, Dr. Jones suggests individuals who would "know the sorts of things to say":

> [a]gree with Kevin that Tom Karl has too much to do. Tom Wigley is semi-retired, and, like Mike Wallace, may not be responsive to requests from JGR.
>
> We have Ben Santer in common! Dave Thompson is a good suggestion. I'd go for one of Tom Peterson or Dave Easterling.
>
> To get a spread, I'd go with US, [o]ne Australian, and one in Europe. So Neville Nicholls and David Parker.

All of them know the sorts of things to say—about our [C]omment and the awful original, without any prompting.

In his e-mail of October 28, 2009 (1256765544.txt), Dr. Jones of CRU has taken exception with an e-mail written by Professor Sonja Boehmer-Cristiansen, a geologist in the Department of Geology at the University of Hull.

Subject: Dr Sonja BOEHMER-CHRISTIANSEN

Dear Professor Haughton,

The email below was brought to my attention ... It was sent by the person named in the header of this email. I regard this email as very malicious. Dr Boehmer-Christiansen states that it is beyond her expertise to assess the claims made. If this is the case then she shouldn't be sending malicious emails like this. The two Canadians she refers to have never developed a tree-ring chronology in their lives and McIntyre has stated several times on his blog site that he has no aim to write up his results for publication in the peer-review literature.

I'm sure you will be of the same opinion as me that science should be undertaken through the peer-review literature as it has been for over 300 years. The peer-review system is the safeguard science has developed to stop bad science being published.

In his response, Professor Haughton states:

Dear Phil, sorry to hear this. I don't see much of her these days, but when I do see Sonja next I'll try and have a quiet word with her about the way the affiliation to us is used, but at the moment in fairness she is entitled to use it in the way she does. Fortunately I don't get to see many of these email exchanges but I do occasionally hear about them or see them and frankly am rarely convinced by what I read. But as with all academics, I'd want to protect another academic's freedom to be contrary and critical, even if I personally believe she is probably wrong. I agree with you that it'd be better for these exchanges to be conducted through the peer review process but these forms of e-communication are now part of the public debate and it's difficult to do much about it other than to defend your position in this and other fora, or just ignore it as being, in your words, malicious.

2) The Adjustment of Data To Produce a Desired Outcome

In the September 22, 1999, e-mail series (0938018124.txt), Keith Briffa, a scientist at the CRU, raises an issue concerning the divergence of tree ring data from actual measured temperatures.

> I know there is pressure to present a nice tidy story as regards "apparent unprecedented warming in a thousand years or more in the proxy data" but in reality the situation is not quite so simple. We don't have a lot of proxies that come right up to [today] and those that do (at least a significant number of tree proxies) [have] some unexpected changes in response that do not match the recent warming. I do not think it wise that this issue be ignored in the chapter.
>
> * * *
>
> I believe that the recent warmth was probably matched about 1000 years ago.

Dr. Mann provides an e-mail to describe how he will engineer a graph mixing tree ring data from the past with modern thermometer temperature readings from the present:

> I am perfectly amenable to keeping Keith's series in the [graph], and can ask Ian Macadam to add it to the [graph] he has been preparing (nobody liked my own color [and graphing] conventions so I've given up doing this myself). The key thing is making sure the [lines] are vertically aligned in a reasonable way. I had been using the entire 20th century, but in the case of Keith's, we need to align the first half of the 20th century [with] the corresponding values of the other lines, due to the late 20th century decline.

Dr. Mann next discusses the 20th century decline in temperatures indicated by the recent tree ring data:

> So if Chris and Tom (?) are ok with this, I would be happy to add Keith's [line to the graph]. That having been said, it does raise a conundrum: We demonstrate ... that the major discrepancies between Phil's and our [line] can be explained in terms of [statistical excuses]. But that explanation certainly can't rectify why Keith's [data], which has similar [properties] to Phil's [data], differs in large part in exactly the opposite direction that Phil's does from ours. This is the problem we all picked up on [—] everyone in the room at IPCC was in agreement that this was a problem and a potential distraction/detraction from the reasonably consensus viewpoint we'd like to show [with] the Jones and Mann [results].

Dr. Mann next describes what they should have done:

> So, if we show Keith's [line] in this plot, we have to comment that "something else" is responsible for the discrepancies in this case. Perhaps Keith can help us out a bit by explaining the processing that went into the [data] and the potential factors that might lead to it being "warmer" than the Jones and Mann [results]? We would need to put in a few words in this regard. Otherwise, the skeptics [would] have [a] field day casting doubt on our ability to understand the factors that influence these

> estimates and, thus, can undermine faith in the ... estimates [from paleological data]. I don't think that doubt is scientifically justified, and I'd hate to be the one to have to give it fodder!

Approximately two months later, Dr. Jones explains how the temperature decline shown by the tree ring data incorrectly for the last 20 years was hidden. In his November 16, 1999, e-mail (0942777075.txt), Dr. Jones explains:

> I've just completed Mike's Nature trick of adding in the real temperatures to each series for the last 20 years (i.e. from 1981 onwards) and from 1961 for Keith's to hide the decline.

In the e-mail series beginning on October 11, 2009 (1255352257.txt), Narasimha Rao, a Ph.D. student at Stanford University, writes to Stanford's Stephen Schneider:

> Steve,
>
> You may be aware of this already. Paul Hudson, BBC reporter on climate change, on Friday wrote that there's been no warming since 1998, and that Pacific oscillations will force cooling for the next 20–30 years. It is not outrageously biased in presentation as are other skeptics' views.
> [links included]
>
> BBC has significant influence on public opinion outside the US. Do you think this merits an op–ed response in the BBC from a scientist?

After the e-mail is forwarded, Mike Mann responds:

> [It is] extremely disappointing to see something like this appear on BBC. [I]t's particularly odd, since climate is usually Richard Black's beat at BBC (and he does a great job). [F]rom what I can tell, this guy was formerly a weather person at the Met Office.

Kevin Trenberth, of the National Center for Atmospheric Research interjects:

> Well I have my own article on where the heck is global warming? We are asking that here in Boulder where we have broken records the past two days for the coldest days on record. ...
>
> The fact is that we can't account for the lack of warming at the moment and it is a travesty that we can't. The ... data published in the August ... 2009 supplement on 2008 shows there should be even more warming: but the data are surely wrong. Our observing system is inadequate.

In an October 14, 2009, e-mail (1255523796.txt), Dr. Trenbeth expresses to Dr. Wigley his concern regarding their lack of understanding of the climate system.

> How come you do not agree with a statement that says we are nowhere close to knowing where energy is going or whether clouds are changing to make the planet brighter? We are not close to balancing the energy budget. The fact that we cannot account for what is happening in the climate system makes any consideration of geoengineering quite hopeless, as we will never be able to tell if it is successful or not! It is a travesty!

In a subsequent e-mail, Dr. Trenberth further explains:

> Here are some of the issues as I see them:
> Saying it is natural variability is not an explanation. What are the physical processes? Where did the heat go? ...
> But the resulting evaporative cooling means the heat goes into atmosphere and should be radiated to space: so we should be able to track it with CERES [sky temperature] data. The CERES data are unfortunately wonting [sic], and so too are the cloud data. The ocean data are also lacking, although some of that may be related to the ocean current changes, and burying heat at depth, where it is not picked up. If it is sequestered at depth then it comes back to haunt us later, and so we should know about it.

3) Obstruction of Freedom of Information Requests

In the January 21, 2005, e-mail (1106338806.txt), Tom Wigley asked Phil Jones about the United Kingdom's Freedom of Information Act (FOI Act):

> I got a brochure on the FOI Act from UAE [the University of East Anglia]. Does this mean that, if someone asks for a computer program we have to give it out?? Can you check this for me (and Sarah).

Phil Jones responds:

> On the FOI Act, there is a little leaflet we have all been sent. It doesn't really clarify what we might have to do [regarding] programs or data. Like all things in Britain we will only find out when the first person or organization asks. I wouldn't tell anybody about the FOI Act in Britain. I don't think UEA really knows what's involved.

Dr. Jones the advises Wigley:

> As you're no longer an employee, I would use this argument if anything comes along.

Tom Wigley replies:

> The leaflet appeared so general, but it was prepared by UEA so they may have simplified things. From their wording, computer [programs] would be covered by the FOIA. My concern was if Sarah is/was still employed by UEA. I guess she could claim that she had only written one tenth of the [programs], and release every tenth line.

Phil Jones replies:

> As for FOIA Sarah isn't technically employed by the University of East Anglia and she will likely be paid by Manchester Metropolitan University.
>
> I wouldn't worry about the [computer programs]. If FOIA does ever get used by anyone, there is also IPR [Intellectual Property Rights] to consider as well. Data is covered by all the agreements we sign with people, so I will be hiding behind them. I'll be passing any requests onto the person at UAE who has been given a post to deal with them.

The discussion of the FOI Act continues in an exchange between Dr. Jones and Dr. Mann. In Dr. Jones' February 2, 2005, e-mail (1107454306.txt), Dr. Jones writes:

> Just sent loads of ... data to Scott. Make sure he documents everything better this time! And don't leave stuff lying around on ftp sites—you never know who is trawling them. [McIntyre and McKitrick] have been after the CRU ... data for years. If they ever hear there is a Freedom of Information Act now in the UK, I think I'll delete the file rather than send it to anyone.
> Does your similar [Act] in the US force you to respond to enquiries within 20 days?—our[s] does! The UK works on precedents, so the first request will test it. We also have a data protection act, which I will hide behind.
>
> Tom Wigley has sent me a worried email when he heard about it—thought people could ask him for his [computer programs]. He has retired officially from UEA so he can hide behind that. [delete line to bring this up into this paragraph] IPR should be relevant here, but I can see me getting into an argument with someone at UEA who'll say we must adhere to it [the Freedom of Information Act]!

Mike Mann responded as follows:

> Yes, we've learned [our] lesson about FTP [sites]. We're going to be very careful in the future what gets put there. Scott really screwed up big time when he established that directory so that Tim could access the data.
>
> Yeah, there is a [F]reedom [O]f [I]nformation [A]ct in the US, and the contrarians are going to try to use it for all it's worth. But there are also intellectual property rights issues, so it isn't clear how these sorts of things will play out ultimately in the US.

In a February 21, 2005, e-mail (1109021312.txt), Dr. Jones writes:

> The skeptics seem to be building up a head of steam here!...
> Leave it to you to delete as appropriate!...
> PS I'm getting hassled by a couple of people to release the CRU ... temperature data. Don't any of you three tell anybody that the UK has a Freedom of Information Act!

In a May 29, 2008, e-mail (1212063122.txt), Dr. Jones writes to Dr. Mann:

> Mike,
>
> Can you delete any emails you may have had with Keith re AR4 [latest Intergovernmental Panel on Climate Change report]? Keith will do likewise. He's not in at the moment—minor family crisis.
>
> Can you also email Gene and get him to do the same? I don't have his new email address.
>
> We will be getting Caspar to do likewise.

Dr. Mann responds:

> I'll contact Gene about this as soon as possible.

In a series of e-mails initiated on December 3, 2008 (1228330629.txt), Dr. Jones states the following:

> The inadvertent email I sent last month has led to a Data Protection Act request sent by a certain Canadian, saying that the email maligned his scientific credibility with his peers!

If he pays 10 pounds (which he hasn't yet) I am supposed to go through my emails and he can get anything I've written about him. About 2 months ago I deleted loads of emails, so have very little—if anything at all.

4) The HARRY READ ME.txt file

The HARRY READ ME.txt file is found in the documents section of FOIA2009.zip.[3] The 274 page file contains the comments and computer code of a CRU programmer over a span of three years (2006-2009) as he attempted to piece together a new database from missing data. Within the file are hundreds of programming source files in different computer language, such as IDL and FORTRAN, that are buried in dozens of subordinate sub-folders. A complete analysis of the programming routines within this file and their implications would reasonably require more than five days, which is the number of days between the time EPA received the supplement to the petition to reopen the public comments based upon the CRU data, and the time that the Endangerment Finding was finalized.

The number of programs in the HARRY READ ME file, which subject raw climate data to repeated "transformative and filtering routines" has been called "simply staggering." Marc Sheppard, *CRU's Source Code: Climategate Uncovered*, American Thinker (Nov. 25, 2009).[4]

Analysis of the HARRY READ ME file, is ongoing, but three examples of data issues have been identified. *Id.* First, in the subfolder labeled "osborn-tree6mannoldprog" there is a program titled "Calibrate_mxd.pro" that calibrates the tree ring temperature data against available local instrumental summer (growing season) temperatures between 1911 and 1990. *Id.* The program then merges that data into a new file. That file is then further modified by another program called "Pl_calibmxd1.pro" which creates calibration statistics for the tree ring data against the stored temperature. *Id.* This is used to "estimate" or "infill" figures where such temperature readings were not available. The file created by that program is modified once again by another program called "Pl_Decline.pro" which, as the author describes, "corrects it" by "identifying" and "artificially" removing "the decline." But the series doesn't begin its "decline adjustment" in 1960, when tree ring data begins its divergence from actual recorded temperatures. Instead, all data between 1930 and 1994 are subject to a warming "correction." *Id.*

In the second example, two programs, "briffa_Sep98_d.pro" and "briffa_Sep98_e.pro," the "correction" appears to be more pronounced. *Id.* The programmer labeled the "adjustment" as

---

[3] The HARRY READ ME.txt file is available from various internet sites, such as http://www.anenglishmanscastle.com/HARRY_READ_ME.txt (last visited Jan. 26, 2010).

[4] *Available at* http://www.americanthinker.com/2009/11/crus_source_code_climategate_r.html (last visited Jan. 28, 2010).

"Apply a VERY ARTIFICAL correction for decline!!" The programmer calls the adjustment a "fudge factor" as a descriptive remark in the code:

> yrloc=[1400,findgen(19)*5.+1904]
> valadj=[0.,0.,0.,0.,0.,-0.1,-0.25,-0.3,0.,-0.1,0.3,0.8,1.2,1.7,2.5,2.6,2.6,2.6,2.6,2.6]*
> 0.75 ; fudge factor

These 2 lines may establish a 20-element array (yrloc) comprising the year 1400 and 19 years between 1904 and 1994 in half-decade increments. Then the corresponding "fudge factor" may be applied to each interval. The CRU programmer's remarks may be an admission that he was tampering with the data by multiplying it by what he himself called a "fudge factor."

In the final example, the HARRY READ ME file contains plotting programs, such as "data4alps.pro" that print this reminder to the user prior to producing a chart:

> IMPORTANT NOTE: The data after 1960 should not be used. The tree-ring density records tend to show a decline after 1960 relative to the summer temperature in many high-latitude locations. In this data set this "decline" has been artificially removed in an ad-hoc way, and this means that data after 1960 no longer represent tree-ring density variations, but have been modified to look more like the observed temperatures.

Another program, such as "mxdgrid2ascii.pro," prints this warning:

> NOTE: recent decline in tree-ring density has been ARTIFICIALLY REMOVED to facilitate calibration. THEREFORE, post-1960 values will be much closer to observed temperatures then [sic] they should be which will incorrectly imply the reconstruction is more skilful [sic] than it actually is. See Osborn et al. (2004).

The following are excerpts of the HARRY READ ME.txt file, showing the CRU programmer's frustrations with the data and his attempt to adjust 11,000 files.

> Nearly 11,000 files! And about a dozen assorted 'read me' files addressing individual issues...
>
> So, uhhhh.. what in tarnation is going on? Just how off-beam are these datasets?!!
>
> Unbelievable – even here the conventions have not been followed. It's botch after botch after botch.
>
> 22. Right, time to stop pussyfooting around the niceties of Tim's labyrinthine software suites – let's have a go at producing CRU TS 3.0! since failing to do that will be the definitive failure of the entire project..

How handy – naming two different files with exactly the same name and relying on their location to differentiate! Aaarrgghh!!

If the latest precipitation database file contained a fatal data error ... then surely it has been altered since Tim last used it to produce the precipitation grids? But if that's the case, why is it dated so early?

So what's going on? I don't see how the 'final' precip file can have been produced from the 'final' precipitation database, even though the dates imply that. The obvious conclusion is that the precip file must have been produced before 23 Dec 2003, and then redated (to match others?) in Jan 04.

There is no way of knowing which Tim used to produce the current public files. The scripts differ internally but – you guessed it! – the descriptions at the start are identical. WHAT IS GOING ON?

So what is this mysterious variable 'nf' that isn't being set? Well strangely, it's in Mark N's 'rdbin.pro'. I say strangely because this is a generic prog that's used all over the place! Nonetheless it does have what certainly looks like a bug...

Where is the documentation to explain all this?!

Bear in mind that there is no working synthetic method for cloud, because Mark New lost the coefficients file and never found it again (despite searching on tape archives at UEA) and never recreated it.

DON'T KNOW, UNDOCUMENTED. Wherever I look, there are data files, no info about what they are other than their names. And that's useless..

So what the hell did Tim do?!! As I keep asking.

This is irritating as it means precip has only 9 fields and I can't do a generic mapping from any cru format to cru ts.

Then.. like an idiot.. I had to test the data!

It's halfway through April and I'm still working on it. This surely is the worst project I've ever attempted. Eeeek.

Oh bugger. What the HELL is going on?!

In fact, on examination the US database record is a poor copy of the main database one, it has more missing data and so forth. By 1870 they have diverged, so in this case it's probably OK.. but what about the others?

Oh GOD if I could start this project again and actually argue the case for junking the inherited program suite!!

Oh Tim what have you done, man?

Just another thing I cannot understand, and another reason why this should all have been rewritten from scratch a year ago!

am I the first person to attempt to get the CRU databases in working order?!!

Oh bum. But, but.. how? I know we do muck around with the header and start/end years, but still..

In the upside-down world of Mark and Tim, the numbers of stations contributing to each cell during the gridding operation are calculated not in the IDL gridding program – oh, no! – but in anomdtb! ..well that was, erhhh.. 'interesting'... So there is no guarantee that the station number files, which are produced *independently* by anomdtb, will reflect what actually happened!!

I am seriously worried that our flagship gridded data product is produced by Delaunay triangulation – apparently linear as well. As far as I can see, this renders the station counts totally meaningless. It also means that we cannot say exactly how the gridded data is arrived at from a statistical perspective – since we're using an off-the-shelf product that isn't documented sufficiently to say that. Why this wasn't coded up in Fortran I don't know – time pressures perhaps? Was too much effort expended on homogenisation, that there wasn't enough time to write a gridding procedure? Of course, it's too late for me to fix it too. Meh.

CRU NEEDS A DATA MANAGER.

Not only do both databases have unnecessary duplicates, introduced for external mapping purposes by the look of it, but the 'main' stations (2 and 4) have different station name & country. In fact one of the country names is illegal! Dealing with things like this cannot be automated as they're the results of non-automatic decisions.

What a bloody mess.

Now looking at the dates.. something bad has happened, hasn't it. COBAR AIRPORT AWS cannot start in 1962, it didn't open until 1993! Looking at the data – the COBAR station 1962-2004 seems to be an exact copy of the COBAR AIRPORT AWS station 1962-2004. And wouldn't you know it, the data for this station has missing data between 12/92 and 12/99 inclusive. So I reckon it's the old FORREST AERO station (WMO 9464600, .au ID 11004), with the new Australian bulletin updates tacked on (hence starting in 2000) So.. do I split off the 2000-present data to a new station with the new number, or accept that whoever joined them (Dave?) looked into it and decided it would be OK? The BOM website says they're 800m apart.

Hope that's right..

All 115 refs now matched in the TMin database. Confidence in the fidelity of the Australian station in the database drastically reduced. Likelihood of invalid merging of Australian stations high. Let's go..

getting seriously fed up with the state of the Australian data. so many new stations have been introduced, so many false references.. so many changes that aren't documented. Every time a cloud forms I'm presented with a bewildering selection of similar-sounding sites, some with references, some with WMO codes, and some with both. And if I look up the station metadata with one of the local references, chances are the WMO code will be wrong (another station will have it) and the lat/lon will be wrong too.

I am very sorry to report that the rest of the databases seem to be in nearly as poor a state as Australia was. There are hundreds if not thousands of pairs of dummy stations, one with no WMO and one with, usually overlapping and with the same station name and very similar coordinates. I know it could be old and new stations, but why such large overlaps if that's the case? Aarrggghhh! There truly is no end in sight.

I honestly have no idea what to do here. and there are countless others of equal bafflingness.

I suspected a couple of stations were being counted twice, so using 'comm' I looked for identical headers. Unfortunately there weren't any!! So I have invented two stations, hmm.

I have to admit, I still don't understand secondary parameter generation. I've read the papers, and the minuscule amount of 'Read Me' documentation, and it just doesn't make sense.

As I was examining the vap database, I noticed there was a 'wet' database. Could I not use that to assist with rd0 generation? well.. it's not documented, but then, none of the process is so I might as well bluff my way into it!

Units seem to vary: <DO YOU SEE? THERE'S THAT OH-SO FAMILIAR BLOCK OF MISSING CODES IN THE LATE 80S, THEN THE DATA PICKS UP AGAIN. BUT LOOK AT THE CORRELATIONS ON THE RIGHT, ALL GOOD AFTER THE BREAK, DECIDEDLY DODGY BEFORE IT. THESE ARE TWO DIFFERENT STATIONS, AREN'T THEY? AAAARRRGGGHHHHHHH!!!!!

Quite honestly I don't have time – but it just shows the state our data holdings have drifted into. Who added those two series together? When? Why? Untraceable, except anecdotally.

But I am beginning to wish I could just blindly merge based on WMO code.. the trouble is that then I'm continuing the approach that created these broken databases.

Here, the expected 1990-2003 period is MISSING – so the correlations aren't so hot! Yet the WMO codes and station names /locations are identical (or close). What the hell is supposed to happen here? Oh yeah – there is no 'supposed', I can make it up. So I have

You can't imagine what this has cost me – to actually allow the operator to assign false WMO codes!! But what else is there in such situations? Especially when dealing with a 'Master' database of dubious provenance (which, er, they all are and always will be).

False codes will be obtained by multiplying the legitimate code (5 digits) by 100, then adding 1 at a time until a number is found with no matches in the database. THIS IS NOT PERFECT but as there is no central repository for WMO codes – especially made-up ones – we'll have to chance duplicating one that's present in one of the other databases. In any case, anyone comparing WMO codes between databases – something I've studiously avoided doing except for tmin/tmax where I had to – will be treating the false codes with suspicion anyway. Hopefully.

This still meant an awful lot of encounters with naughty Master stations, when really I suspect nobody else gives a hoot about. So with a somewhat cynical shrug, I added the nuclear option – to match every WMO possible, and turn the rest into new stations (er, CLIMAT excepted). In other words, what CRU usually do. It will allow bad databases to pass unnoticed, and good databases to become bad, but I really don't think people care enough to fix 'em, and it's the main reason the project is nearly a year late.

this was a guess! We'll see how the results look Right, erm.. off I jolly well go!

The trouble is, we won't be able to produce reliable station count files this way. Or can we use the same strategy,producing station counts from the wet database route, and filling in 'gaps' with the precip station counts? Err.

...It looks as though the calculation I'm using for percentage anomalies is, not to put too fine a point on it, cobblers.

So, good news – but only in the sense that I've found the error. Bad news in that it's a further confirmation that my abilities are short of what's required here.

...unusual behaviour of CRU TS 2.10 Vapour Pressure data was observed, I discovered that some of the Wet Days and Vepour Pressure datasets had been swapped!!

Ah – and I was really hoping this time that it would just WORK. But of course not – nothing works first time in this project.

Oh, GOD. What is going on? Are we data sparse and just looking at the climatology? How can a synthetic dataset derived from tmp and dtr produce the same statistics as an 'real' dataset derived from observations?

I DON'T UNDERSTAND!!!!!

Oh, sod it. It'll do. I don't think I can justify spending any longer on a dataset, the previous version of which was completely wrong (misnamed) and nobody noticed for five years.

"Bear in mind that there is no working synthetic method for cloud, because Mark New lost the coefficients file and never found it again (despite searching on tape archives at UEA) and never recreated it. This hasn't mattered too much, because the synthetic cloud grids had not been discarded for 1901-95, and after 1995 sunshine data is used instead of cloud data anyway." As for converting sun hours to cloud cover.. we only appear to have interactive, file-by-file programs. Aaaand – another head-banging shocker! The program sh2cld_tdm.for, which describes itself thusly: program sunh2cld c converts sun hours monthly time series to cloud percent (n/N)

Does NO SUCH THING!!! Instead it creates SUN percentages! This is clear from the variable names and user interactions.

So.. if I add the sunh -> sun% process from sh2cld_tdm.for into Hsp2cldp_m.for, I should end up with asun hours to cloud percent convertor. Possibly.

It also assisted greatly in understanding what was wrong – Tim was in fact calculating Cloud Percent, despite calling it Sun Percent!! Just awful.

... So to CLOUD. For over a year, rumours have been circulating that money had been found to pay somebody for a month to recreate Mark New's coefficients. But it never quite gelled. Now, at last, someone's producing them! Unfortunately.. it's me.

The idea is to derive the coefficients (for the regressing of cloud against DTR) using the published 2.10 data. We'll use 5-degree blocks and years 1951-2002, then produce coefficients for each 5-degree latitude band and month. Finally, we'll interpolate toget half-degree coefficients. Apparently.

So, erm.. now we need to create our synthetic cloud from DTR. Except that's the thing we CAN'T do because pro cal_cld_gts_tdm.pro needs those bloody coefficients (a.25.7190, etc) that went AWOL.

Hunting for CDDs I found a potential problem with binary DTR (used in the construction of Frost Days, Vapour Pressure, and (eventually) Cloud. It looks as though there was a mistyping when the 2.5-degree binaries were constructed:

Another problem. Apparently I should have derived TMN and TMX from DTR and TMP, as that's what v2.10 did and that's what people expect. I disagree with publishing datasets that are simple arithmetic derivations of other datasets published at the same time, when the real data could be published instead.. but no.

I then look in the 1995 anomaly files...This whole process is too convoluted and created myriad problems of this kind. I really think we should change it.

I was going to do further backtracing, but it's been revealed that the same issues were in 2.1 – meaning that I didn't add the duff data. The suggested way forward is to not use any observations after 1989, but to allow synthetics to take over. I'm not keen on this approach as it's likely (imo) to introduce visible jumps at 1990, since we're effectively introducing a change of data source just after calculating the normals. My compromise is to try it – but to also try a straight derivation from half-degree synthetics.

So actually, this was saving with a gridsize of 5 degrees! Disquietingly, this isn't born out by the file sizes, but we'll gloss over that.

Station counts should be straightforward to derive from the anomaly files (.txt), as output by anomdtb.f90. This, however, will only work for Primary parameters, since

Secondaries are driven from synthetic data as well. Further, the synthetic element in this is usually at 2.5 degrees, so a direct relationship with half-degree coverage will be hard to establish.

So, we can have a proper result, but only by including a load of garbage!

OK, got cloud working, have to generate it now.. but distracted by starting on the mythical 'Update' program.

Of course, one of the problems is that you need a latitude value to perform the conversion – so the CLIMAT bulletins lose the value if they can't be matched in the WMO list! Not much I can do about that, and let's face it those stations are going to end up as 'new' stations with no possibility of a 61-90 normal.

So the new cloud databases I've just produced should be, if not identical, very similar? Oh, dear. There is a passing similarity, though this seems to break down in Winter. I don't have time to do detailed comparisons, of course, so we'll just run with the new one.

The procedure last time – that is, when I was trying to re-produce TS 2.10, we have no idea what the procedure was for its initial production!

So after gridding we could add these.. except that after gridding we'll have incorporated the DTR_derived synthetic cloud, which is of course based on the 1961-1990 normals as it's derived from DTR!! Arrrrggghh.

So.. {sigh}.. another problem. Well we can't change the updates side, that has to use 1995-2002 normals. But maybe we'll have to adjust the station anomalies, prior to gridding? I don't see an alternative.

The question is, IS THIS ANY GOOD? Well, we currently have published cloud data to 2002. So we can make comparisons between 1996 and 2002. Oh, my. I am sure I've written plenty of comparison routines, but as to their location or name..ah... The results were less than ideal, though they could have been much worse. Essentially, North America is totally different...

The deduction so far is that the DTR-derived CLD is waaay off. The DTR looks OK, well OK in the sense that it doesn;t have prominent bands! So it's either the factors and offsets from the regression, or the way they've been applied in dtr2cld.

Well, dtr2cld is not the world's most complicated program. Wheras cloudreg is, and I immediately found a mistake! Scanning forward to 1951 was done with a loop that, for completely unfathomable reasons, didn't include months! So we read 50 grids

instead of 600!!! That may have had something to do with it. I also noticed, as I was correcting THAT, that I reopened the DTR and CLD data files when I should have been opening the bloody station files!! I can only assume that I was being interrupted continually when I was writing this thing. Running with those bits fixed improved matters somewhat, though now there's a problem in that one 5-degree band (10S to 5S) has no stations! This will be due to low station counts in that region, plus removal of duplicate values.

Had a think. Phil advised averaging the bands either side to fill the gap, but yuk! And also the band to the North (ie, 5S to equator) is noticeably lower (extreme, even). So after some investigation I found that, well, here's the email:

<MAIL QUOTE>Phil, I've looked at why we're getting low counts for valid cloud cells in certain 5-degree latitude bands.

The filtering algorithm omits any cell values where the station count is zero, for either CLD or DTR. In general, it's the CLD counts that are zero and losing us the data. However, in many cases, the cloud value in that cell on that month is not equal to the climatology. And there is plenty of DTR data. So I'm wondering how accurate the station counts are for secondary variables, given that they have to reflect observed and synthetic inputs.

So, I'm proposing to filter on only the DTR counts, on the assumption that PRE was probably available if DTR was, so synthesis of CLD was likely to have happened, just not shown in the station counts which are probably 'conservative'?<END MAIL QUOTE> I didn't get an email back but he did verbally consent. So away we go!

Running with a DTR-station-only screening gives us lots of station values, even with duplicate filtering turned back on. Niiice. It's still not exactly smooth, but it might be enough to 'fix' the synthetic cloud.

Differences with the climatology, or with the 2.10 release, are patchy and generally below 30%. Of course it would be nice if the differences with the 2.10 release were negligable, since our regression coefficients were based on 2.10 DTR and CLD.. though of course the sun hours component is an unknown there, as is the fact that 2.10 used PRE as well as DTR for the synthetics. Anyway it gets the thumbs-up. The strategy will be to just produce it for 2003-2006.06, to tie in with the rest of the 3.00 release. So I just need to.. argh. I don't have any way to create NetCDF files 1901-2006 without the .glo.abs files to work from! I'd have to specially code a version that swallowed the existing 1901-2002 then added ours. Meh.

I really thought I was cracking this project. But every time, it ends up worse than before.

I really do hate this whole project.

No time to finish and test the fortran gridder, which will doubtless sink to some depth and never be seen again, we'll carry on with this mediocre approach.

It's not going to be easy to find 14 missing stations, is it? Since the anomalies aren't exactly the same.

Should I be worried about 14 lost series? Less than 2%. Actually, I noticed something interesting.. look at the anomalies. The anomdtb ones aren't *rounded* to 1dp, they're *truncated*! So, er – wrong!

So let's say, anomalies are done. Hurrah. Onwards, plenty more to do!

NO IDEA why, so saying they affect particular 0.5-degree cells is harder than it should be. So we'll just gloss over that entirely ;0)

Just went back to check on synthetic production. Apparently – I have no memory of this at all – we're not doing observed rain days! It's all synthetic from 1990 onwards. Probably the worst story is temperature, particularly for MCDW. Over 1000 new stations! Highly unlikely. I am tempted to blame the different lat/lon scale, but for now it will have to rest.

Oh, my giddy aunt. What a crap crap system.

Also went through the parameters one by one and fixed (hopefully) their scaling factors at each stage. What a minefield!

- I was able to look at the first problem (Guatemala in Autumn 1995 has a massive spike) and find that a station in Mexico has a temperature of 78 degrees in November 1995! This gave a local anomaly of 53.23 (which would have been 'lost' amongst the rest of Mexico as Tim just did country averages) and an anomaly in Guatemala of 24.08 (which gave us the spike)...

Oh, ****. It's the bloody WMO codes again. **** these bloody non-standard, ambiguous, illogical systems. Amateur hour again.

This whole project is SUCH A MESS.

I am seriously close to giving up, again. The history of this is so complex that I can't get far enough into it before by head hurts and I have to stop. Each parameter has a tortuous history of manual and semi-automated interventions that I simply cannot just

go back to early versions and run the update prog. I could be throwing away all kinds of corrections – to lat/lons, to WMOs (yes!), and more.

You see how messy it gets when you actually examine the problem? What we really need, and I don't think it'll happen of course, is a set of metrics (by latitude band perhaps) so that we have a broad measure of the acceptable minimum value count for a given month and location. Even better, a confidence figure that allowed the actual standard deviation comparison to be made with a looseness proportional to the sample size.

OH F*** THIS. ... I'm hitting yet another problem that's based on the hopeless state of our databases. There is no uniform data integrity, it's just a catalogue of issues that continues to grow as they're found.

PACIFIC LEGAL FOUNDATION

February 5, 2010

Administrator
United States Environmental Protection Agency
Room 3000
Ariel Rios Building
1200 Pennsylvania Avenue, N.W.
Washington, D.C. 20460

VIA FEDERAL EXPRESS

Re: Petition for Reconsideration under Section 307(d) of the Clean Air Act, 42 U.S.C. § 7607(d); Endangerment Finding, 74 Fed. Reg. 66,496 (Dec. 15, 2010); EPA Docket No. EPA-HQ-OAR-2009-0171

Dear Administrator:

Enclosed is Pacific Legal Foundation's Petition for Reconsideration, filed pursuant to Clean Air Act Section 307(d), 42 U.S.C. § 7607(d), regarding the "Endangerment and Cause or Contribute Finding for Greenhouse Gases Under Section 202(a) of the Clean Air Act," 74 Fed. Reg. 66,496 (Dec. 15, 2009).

Sincerely,

THEODORE HADZI-ANTICH
Attorney

Enclosure

cc: Mr. Jeremy Martinich (via Federal Express and E-mail)
Associate General Counsel for the Air and Radiation Law Office (via Federal Express)

Headquarters: 3900 Lennane Drive, Suite 200 • Sacramento, CA 95834 • (916) 419-7111 • Fax: (916) 419-7747
Alaska: 121 West Fireweed Lane, Suite 250 • Anchorage, AK 99503 • (907) 278-1731 • Fax: (907) 276-3887
Atlantic: 1002 SE Monterey Commons Blvd., Suite 102 • Stuart, FL 34996 • (772) 781-7787 • Fax: (772) 781-7785
Hawaii: P.O. Box 3619 • Honolulu, HI 96811 • (808) 733-3373 • Fax: (808) 733-3374 • Oregon: (503) 241-8179
Washington: 10940 NE 33rd Place, Suite 210 • Bellevue, WA 98004 • (425) 576-0484 • Fax: (425) 576-9565
E-mail: plf@pacificlegal.org • Web Site: www.pacificlegal.org