Nullifying the climate null hypothesis

Journal:	*WIREs Climate Change*
Manuscript ID:	WCC-457.R1
Wiley - Manuscript type:	Opinion
Date Submitted by the Author:	n/a
Complete List of Authors:	Curry, Judith
Choose 1-3 topics to categorize your article:	Knowledge and practice (ANAD) < The Social Status of Climate Change Knowledge (ANAA)

Nullifying the climate null hypothesis

Judith Curry

Abstract

This essay addresses Trenberth's statement that "*Given that global warming is 'unequivocal', and is 'very likely' due to human activities to quote the 2007 IPCC report, the null hypothesis should now be reversed, thereby placing the burden of proof on showing that there is no human influence."* We examine how the concept of a null hypothesis is being used implicitly and explicitly in the scientific and policy debate on climate change, in the context of scientific hypothesis testing, as a framework for "burden of proof" arguments and policy deliberations, and metaphorically in the context of a polemic. It is argued that the statement of a null hypothesis is not particularly useful in the broader context of the scientific inferences surrounding the topic of the attribution of climate change and also policy decisions.

1. Introduction

In a presentation made at the Annual Meeting of the American Meteorological Society (AMS), Trenberth (2011) made a dramatic statement:

> *Prior to the 2007 IPCC report, it was appropriate for the null hypothesis to be that "there is no human influence on climate" and the task was to prove that there was. . . Given that global warming is "unequivocal", and is "very likely" due to human activities to quote the 2007 IPCC report, the null hypothesis should now be reversed, thereby placing the burden of proof on showing that there is no human influence. . . As a whole the community is making too many type II errors (erroneously accept the null hypothesis when it is in fact false).*

This statement appeals to the IPCC consensus as a basis for being relieved from the burden of proof and implies that scientists are making too many errors by not reversing the null hypothesis. It reinforces the idea of an overwhelming scientific consensus. Further, it removes the need for scientists to continually make an argument that anthropogenic activities are influencing extreme weather events, which is the original context for Trenberth's (2011) statement.

Trenberth's statement motivates an examination of the concept of the null hypothesis, and how it is being used implicitly and explicitly in the scientific and policy debate on climate change. The null hypothesis is used in different contexts: scientific hypothesis testing; as a framework for "burden of proof" arguments and decision making; and metaphorically in the context of a polemic. It is argued here that the concept of a null hypothesis is not particularly useful in the broader context of the scientific inferences surrounding the topic of the attribution of climate change and related policy decisions.

2. Role of the null hypothesis in scientific argument regarding attribution

Scientists continually strive for rigorous, objective approaches for making valid inferences concerning science questions. One such approach is null hypothesis testing. The term "null hypothesis" originates from Ronald Fisher, an English geneticist and statistician. The null hypothesis is typically paired with an alternative hypothesis. Fisher's concept of the null is described in his 1935 text on "The Design of Experiments:"

> *Different types of experimentation are considered with reference to their logical structure, to show that valid conclusions may be drawn from them without using the disputed theory of inductive inferences, i.e., of arguing from observation to explanatory theory. This is possible if a null hypothesis is explicitly formulated when the experiment is designed; this hypothesis can never be proved, but may be disproved with whatever probability one will accept as demonstrating a positive result*. (Fisher 1935)

In simple laboratory experiments or in clinical trials, the null hypothesis is easily formulated in the context of "no effect." Although the null hypothesis is rarely discussed explicitly in the context of climate research, it is implied in analyses that result in a confidence interval. The null hypothesis is typically of the "no effect" type, although it need not be. The choice of the null and alternative hypotheses is determined by the question that the scientist seeks to investigate. There is no unique argument of logic to determine the appropriate null hypothesis for a particular alternative hypothesis, or to determine which hypothesis in an opposing pair should be considered as the null hypothesis.

The applications, utility and misunderstanding of the null hypothesis have been the subject of substantial discussion (e.g. Cohen 1994, Nickerson 2000, Anderson et al. 2000). It has been argued that null hypothesis testing leads to relatively little increase in understanding and diverts attention from issues of estimation of effect size and meaningful mechanistic understanding and modeling of predictive and causal relationships. A frequent criticism of null hypothesis testing is the formulation of the analysis to reject a null hypothesis that is trivially known to be false. The rejection of such 'strawmen' does not advance science, and fallacious arguments are often made that use this rejection as evidence in support of the alternative hypothesis.

Hypotheses about complex problems such as climate change are either implicitly or explicitly built upon a collection of sub-hypotheses that are related to each other in the context of a syllogistic paradigm or other type of logical analysis. The challenges of testing a complex hypothesis involving a causal chain was addressed in Curry et al. (2006) in the context of testing the hypothesis that greenhouse warming is causing an increase in global hurricane intensity. The central hypothesis was broken down into three sub-hypotheses that were each necessary for the central hypothesis to be true, and further comprised a causal chain. A null hypothesis was formulated for each of the sub-hypotheses, and evidence was presented for both the null and sub-hypotheses. The conclusion was that the evidence did not support the rejection of any of the sub-hypotheses, and hence did not support rejection of the central hypothesis. Neither was the evidence sufficient for rejecting the nulls of the sub-hypotheses, given uncertainties in the data and models. Further, it was acknowledged that the specific causal chain and sub-hypotheses proposed for the central hypothesis might be incomplete or otherwise incorrect. While the analysis was inconclusive, it nevertheless clarified the evidence and arguments for and against the central hypothesis that greenhouse warming is causing an increase in global hurricane intensity.

Consider Trenberth's (2011; AMS paper) version of the current climate null hypothesis statement: *"there is no human influence on climate."* This null hypothesis is regarded as trivially false by nearly everyone: no one denies the role of land use changes, pollution aerosol and anthropogenically-produced greenhouse gases in modifying the climate. The key scientific question is the importance of human influences relative to natural modes of climate variability.

Ghil (2001) describes the issues surrounding the climate change null hypothesis in the following way, discriminating between local and global change and separating out the two related issues of detection and attribution:

> *More precisely, we ask whether the impact of human activities on the climate is observable and identifiable in the instrumental records of the last century-and-a-half and in recent paleoclimate records? The answer to this question depends on the null hypothesis against which such an impact is tested. The current approach that is generally pursued assumes essentially that past climate variability is indistinguishable from a stochastic red-noise process, whose only regularities are those of periodic external forcing. Given such a null hypothesis, the official consensus of IPCC (1995) tilts towards a global warming effect of recent trace-gas emissions, which exceeds the cooling effect of anthropogenic aerosol emissions. . . The presence of internally arising regularities in the climate system with periods of years and decades suggests the need for a different null hypothesis. Essentially, one needs to show that the behaviour of the climatic signal is distinct from that generated by natural climate variability in the past, when human effects were negligible, at least on the global scale. . . Can we identify with measurable certainty deviations of the current record from predictions based on past natural variability? If so, such deviations have to be attributed to new causes. The "suspects" clearly include human effects, and attribution to*

them will become thereby both easier and more reliable.

Key issues under debate in climate change detection and attribution arguments are the credibility of historical data sets and paleo-reconstructions, interpretation of natural modes of climate variability (forced and unforced), and sensitivity of the climate system to nonlinear interactions and feedbacks within the climate system in response to external forcing. Regardless of our capabilities to detect global anthropogenic climate change with certainty, it is generally accepted that humans do modify their environment in a manner that can impact regional and global climate. The detection and attribution of individual extreme weather events, or collections of extreme weather events, is complicated further by their inadequate simulation by climate models, including their dependence on weather regimes and internal multi-decadal variability.

While Trenberth (2011, AMS paper) does not articulate a specific hypothesis regarding anthropogenic climate change, for the sake of argument we adopt the IPCC AR4 attribution statement as the relevant hypothesis, H1:

> H1: "*Most of the observed increase in global average temperatures since the mid-20th century is very likely [>90%] due to the observed increase in anthropogenic greenhouse gas concentrations.*"

Is Trenberth's statement of the current null hypothesis "*there is no human influence on climate"* useful as a null hypothesis in the context of H1? It is not, since the statement is generally accepted as trivially false and its falsification lends no support to H1. Attempting to formulate a null hypothesis for the IPCC's attribution statement reveals problems with H1 as a useful hypothesis. The utility of the IPCC's attribution statement is aptly summarized by the following quote[1] from a document discussing climate change and national security:

> "*For the past 20 years, scientists have been content to ask simply whether most of the observed warming was caused by human activities. But is the percentage closer to 51 percent or to 99 percent? This question has not generated a great deal of discussion within the scientific community, perhaps because it is not critical to further progress in understanding the climate system. In the policy arena, however, this question is asked often and largely goes unanswered.*"

I regard this question as the critical consideration for understanding the climate system. The high likelihood [>90%] of the imprecise 'most' seems rather meaningless. From the InterAcademy Council Review of the IPCC[2]: "*In the Committee's view, assigning probabilities to imprecise statements is not an appropriate way to characterize uncertainty."* Given the absence of traceability regarding the reasoning that led to formulating the attribution statement in this way (as well as the justification for the likelihood level), it is difficult to infer the underlying reasoning and rationale for the form and content of the IPCC AR4 attribution statement.

It is inferred here that the imprecision of 'most' arises in part from regarding the issue of the human influence on climate as a binary issue in the context of null hypothesis testing, whereby humans influence climate at a level greater than a specified fraction, or they do not. In fact, an infinite number of different hypotheses can be formulated, with different fractions of the warming attributed to anthropogenic greenhouse gas concentrations. The relevant issue then becomes the likelihood to assign to the various fractions. If there is significant likelihood in the vicinity of 50-

[1] *Lost in Translation: Closing the Gap Between Climate Science and National Security Policy*, published by the Center for a New American Security
http://www.cnas.org/files/documents/publications/Lost%20in%20Translation_Code406_Web_0.pdf

[2] http://reviewipcc.interacademycouncil.net/

50, then the binary demarcation at 50% implied by H1 ('most') is a poor choice.

3. Decision making and burden of proof

Whereas it is a scientist's choice regarding which hypothesis to test and whether to incorporate a null hypothesis into the argument, scientific assessment for policy makers is most usefully formulated in a context that is relevant to decision making needs. Specifically with regards to climate change, the UNFCCC[3] policy deliberations are intimately connected with the concept of "*dangerous anthropogenic interference with the climate system.*" While Trenberth (2011; AMS paper) does not specifically use the word "dangerous," he does use the phrase "climate disasters" in the context of policy making.

In the context of policy deliberations, Trenberth's argument for reversing the null hypothesis hinges on whether the IPCC's evidence and arguments meets an appropriate standard of proof. In both legal and philosophical applications, the burden of proof is always on the person asserting the claim, which is the concept underpinning the presumption of innocence in some legal systems. More generally, the burden of proof lies with those arguing against conventional understanding or wisdom.

In a burden of proof application, the conventional understanding carries the benefit of assumption, implying that there is no need to provide evidence in support of the conventional wisdom. The consequence of meeting the standard of the burden of proof is that the new claim becomes the conventional wisdom, passing the burden of proof to the opposing party. In effect, Trenberth seems to be arguing that the evidence presented by the IPCC is sufficient that dangerous anthropogenic climate change should now be regarded as the conventional wisdom with the benefit of assumption. The implication is that the burden of proof in the public debate should lie with those that are skeptical of dangerous anthropogenic climate change.

Trenberth's argument for reversing the burden of proof hinges on whether the IPCC's evidence and arguments meets an appropriate standard of proof. The commonly used levels of proof are preponderance of evidence, clear and convincing evidence, and beyond a reasonable doubt. The issues discussed in Section 2 not withstanding, the IPCC has arguably met the standard of proof for clear and convincing evidence supporting the existence of anthropogenic effects on climate, although its relative magnitude with respect to natural climate variability remains unclear. In the context of decision making, the concept of "dangerous" interference in the climate is of paramount importance: the relative roles of natural versus anthropogenic warming is meaningless in the context of policy making without providing evidence that the changed climate is dangerous to humans and ecosystems. The socially relevant issue then becomes whether anthropogenically induced climate changes are dangerous, versus too trivial to require action. The issue of what constitutes dangerous climate change on either local or global scales is complex and is not well understood owing to confounding factors associated with climate change impacts. It is my judgment that the IPCC has provided evidence that supports a reasonable suspicion of dangerous impacts but has not met the common standards of proof. Under these circumstances, the burden of proof is a tool that can be used to evaluate evidence and arguments on both sides of the climate science debate.

In public debate, the required standard of evidence can be arbitrary unless previously agreed upon. The precautionary principle introduces an interesting twist into the debate. While there are

[3] United Nations Framework Convention on Climate Change, http://unfccc.int/

many definitions and nuances to the precautionary principle, of particular relevance here is the concept of 'strong' versus 'weak' precaution (e.g. Gardiner 2006): under weak precaution, the burden of proof for justifying the need for action falls on those advocating precautionary action, whereas under strong precaution the burden of proof is on those who argue that the activity does not cause significant harm. As such, Trenberth's argument for reversing the null hypothesis and shifting the burden of proof seems less about science and more about whether strong or weak precaution should prevail in the context of climate change policy.

The implications of the weak and strong versions of the precautionary principle can also be understood in the context of Type I and Type II errors. A Type I error rejects a true null hypothesis (false positive), whereby a Type II error fails to reject a false null hypothesis (false negative.) In a more colloquial sense, a Type I error is associated with failing to believe the truth, whereas a Type II error is associated with continuing to believe what is a falsehood for lack of a better alternative. Emphasizing the importance of Type II errors is the basis of the precautionary principle. Prudent decision makers are concerned about both Type I and Type II errors. Reversing the null hypothesis and changing the burden of proof will invert the Type II errors, which may then become Type I errors or disappear as new errors are created.

In deciding who should bear the burden of proof in a public debate, we should ask "Where lies wisdom?" In the case of testing the efficacy of a new drug, the assumption of "no effect" is regarded as more wisely assumed in the absence of strong evidence to the contrary, with the burden of proof on those demonstrating the efficacy of a new drug. Even in the context of this relatively straightforward example, the level of proof remains a debatable issue for terminally ill individuals with nothing to lose and a chance of being helped by the experimental drug.

In the case of dangerous anthropogenic climate change, the judgment involves political, economic and ethical issues, as well as scientific assessments. If the null hypothesis is that the anthropogenic impact on climate change is dangerous, and the null hypothesis cannot be falsified, the implication is that the appropriate policy response is a serious effort to mitigate the adverse impacts. If the null hypothesis is that human impacts on climate change are not dangerous, failure to falsify the null hypothesis would justify a decision to refrain from serious mitigation and adaptation efforts. The problem with both of these null hypotheses is that neither is falsifiable in any scientifically objective way.

Application of the burden of proof as a tool in public debate requires a clear understanding of whom it is that the party with the burden of proof has to convince. If the jury is the IPCC, then H1 reflects the consensus understanding and conventional wisdom. If the jury is the broader world of policy makers, the choice is more problematic. Several European governments have implemented policies that are consistent with the IPCC consensus, and hence the burden of proof on changing policy in these countries lies with those that are skeptical of climate change. In the U.S., where federal policies on carbon mitigation have not been implemented, the burden of proof remains with the IPCC consensus.

Specifically with regards to Trenberth's focus on the attribution of extreme weather events, the policy relevant issue seems primarily related to generating sufficient evidence of "dangerous anthropogenic interference with the climate system" so as to trigger global precautionary policies. In the context of local adaptation measures, where the typical concern is recent loss rather than future risk, the recent occurrence of an extreme event in a particular locale commonly motivates consideration of local adaptation measures, independent of whether the event was caused or exacerbated in some way by anthropogenic global warming.

4. Polemics and the metaphorical null

I have argued that the debate about reversing the null hypothesis does not make much sense in the context of the scientific debate surrounding the attribution of climate change. The debate regarding which side bears the burden of proof hinges primarily on societal tolerance for risk, i.e. whether the weak or strong version of the precautionary principle should prevail. The debate about reversing the null hypothesis is arguably most significant in the context of the political debate, where the concept is more metaphorical than scientific.

The original paper of Trenberth's recommending reversal of the null hypothesis (Trenberth 2011; AMS paper) provides the following context for his recommendation:

> *"The climate change deniers have very successfully caused major diversions from the much needed debate about what to do about climate change and how to implement it. It is important that climate scientists learn how to counter the distracting strategies of deniers."*
>
> *"The media have been complicit in the disinformation campaign of the deniers."*
>
> *"My own observation is that while some politicians are indeed well informed and understand their role, most are not. The corrupting influence of funding from all sources of vested interests prevents many of them from doing the right thing on behalf of the country and civilization as a whole. It is clear that climate science has become politicized, and scientists are slow to recognize this. Politicians hide behind the apparent uncertainties and have failed to act. Hence while politicians are often also part of the problem, implementation of policies necessarily goes through them."*
>
> *"Unfortunately, society is not ready to face up to these challenges and the needed changes in the way we create order and govern ourselves."*

My personal reaction to Trenberth's (2011; AMS paper) essay is that the statement seems less about scientific analysis than about policy and winning a battle against the 'deniers' and reluctant politicians. In this sense, his essay comes across as a polemic. Trenberth is using the idea of reversing the null hypothesis as a metaphor for changing the political balance in the climate change debate. Trenberth's proposal to reverse the null hypothesis is basically an argument that the natural political response should be to take strong actions to address anthropogenic climate change, and that those who oppose such actions should bear the burden of providing evidence that the actions are unnecessary. In this context, Trenberth's proposal has little to do with scientific hypothesis testing, but is based on an opinion concerning the appropriate actions for responding to dangers of anthropogenic climate change.

The statement regarding the null hypothesis in the context of this polemic appeals to the IPCC consensus as a basis for being relieved from the burden of proof, and implies that scientists are making too many errors by not reversing the null hypothesis. I can imagine a public relations person finding this idea of changing the null hypothesis to receive a favorable public response in focus groups whose members are already convinced of the IPCC's arguments. It reinforces the idea of an overwhelming scientific consensus. It removes the need for scientists to continually make an argument that anthropogenic activities are influencing extreme weather events. Trenberth's attempt to seize the climate battlefield terrain under the metaphor of the null

hypothesis has elements of an effective polemic.
From the perspective of climate change skeptics (or those referred to by Trenberth as 'deniers'), this polemic arguably comes across as, well, a polemic. The net result of such a strategy may be to further inflame the "climate wars" and place climate scientists in a position where they are perceived as moving away from a real discussion and debate on the scientific issues.

5. Conclusions

This discussion on the null hypothesis has highlighted the fuzziness surrounding the actual hypotheses related to dangerous climate change and their falsifiability. The central scientific disputes are the relative importance of anthropogenic forcing and natural variability and whether a warmer climate will lead to changes that are likely to be judged dangerous. These issues are ill suited to null hypothesis testing, since the primary issues are the estimation of relative effect size and meaningful mechanistic understanding and model predictions. The policy debate is centered on whether anthropogenically-induced climate changes are dangerous, versus too trivial to require action or too costly to fix.

With regards to attribution, rather than trying to reject either of these hypotheses (regardless of which is the null), there should be a debate regarding the relative significance of anthropogenic warming relative to forced and unforced natural climate variability. Changing the debate to focus on the relative significance of anthropogenic forcing relative to natural variability accepts the existence of an anthropogenic effect while retaining skepticism on the issues of real interest and importance. Further, this could turn the dispute more productively away from polemics and become a real scientific debate that identifies common ground and clarifies the key uncertainties and areas of ignorance.

Central issues surrounding the responsibility for burden of proof in the public debate surrounding dangerous anthropogenic climate change hinges on who the jury is determined to be and whether strong or weak precaution is preferred. Attempts at providing a scientific rationale for mandating that one side versus the other bears the burden of proof in a public debate is futile given the broader concerns of the jury and the conflict of values on both sides of the debate.

Trenberth's proposal to reverse the null hypothesis seems most potent in the context of a polemic designed to change the political balance in the climate change debate and further marginalize the "deniers". As a polemic, the idea of reversing the null hypothesis is a strategy for scoring political points against skeptics. However, such strategies are likely to exacerbate skepticism and inflame the political debate, which can be counterproductive.

So Trenberth bears the burden of proof on reversing the null hypothesis. It remains to be seen whether he and other proponents of his argument can make a sufficiently compelling case for such reversal to the relevant juries. Perhaps Trenberth should be careful of what he asks for – one consequence of reversing the null hypothesis is that the scientific focus (and funding) should arguably reverse to attempting to disprove dangerous anthropogenic climate change, which has been a position of many skeptics.

Acknowledgements. I would like to acknowledge the Denizens of my blog Climate Etc. (http://judithcurry.com/2011/05/16/null-hypothesis-discussion-thread/) for their insightful comments and discussions on this topic. I am grateful to Peter Webster and Paul Baer for providing comments on the manuscript. This research was funded by a grant from NOAA.

References

Anderson, DR, KP Burnham, WL Thompson, 2000: Null hypothesis testing: problems, prevalence and an alternative. *J. Wildlife Mgt*., 64, 912-923.

Cohen 1994: The earth is round ($p < .05$). *American Psychologist*, 49, 997-1003.

Curry, J.A., P.J. Webster, G. J. Holland 2006: Mixing politics and science in testing the hypothsis that greenhouse warming is causing a global increase in hurricane intensity. *Bull. Amer. Meteorol. Soc*., DOI:10.1175/BAMS-87-8-1025.

Fisher, RA 1935: *The Design of Experiments*. Oliver and Boyd, Oxford England, 251 pp.

Gardiner, S.M., (2006) A Core Precautionary Principle. *The Journal of Political Philosophy,* 14, 33-60.

Ghil M. (2001) Hilbert problems for the geosciences in the 21st century. *Nonlin. Proc. in Geophys*., 8, 211-222.

Nickerson, R.S. 2000: Null hypothesis significance testing: A review of an old and continuing controversy. *Psych. Methods*, 5, 241-301.

Trenberth, K.E. (2011) *Communicating climate science and thoughts on Climategate*. American Meteorological Society Joint Presidential Session on Communicating Climate Change, 23-27 January 2011, Seattle, WA.
http://ams.confex.com/ams/91Annual/webprogram/Paper180230.html