

Learning from mistakes in climate research

Rasmus E. Benestad¹ · Dana Nuccitelli² · Stephan Lewandowsky^{3,4} ·
Katharine Hayhoe⁵ · Hans Olav Hygen¹ · Rob van Dorland⁶ · John Cook^{2,7,8}

Received: 21 May 2014 / Accepted: 27 July 2015 / Published online: 20 August 2015
© The Author(s) 2015. This article is published with open access at Springerlink.com

Abstract Among papers stating a position on anthropogenic global warming (AGW), 97 % endorse AGW. What is happening with the 2 % of papers that reject AGW? We examine a selection of papers rejecting AGW. An analytical tool has been developed to replicate and test the results and methods used in these studies; our replication reveals a number of methodological flaws, and a pattern of common mistakes emerges that is not visible when looking at single isolated cases. Thus, real-life scientific disputes in some cases can be resolved, and we can learn from mistakes. A common denominator seems to be missing contextual information or ignoring information that does not fit the conclusions, be it other relevant work or related geophysical data. In many cases, shortcomings are due to insufficient model evaluation, leading to results that are not universally valid but rather are an artifact of

a particular experimental setup. Other typical weaknesses include false dichotomies, inappropriate statistical methods, or basing conclusions on misconceived or incomplete physics. We also argue that science is never settled and that both mainstream and contrarian papers must be subject to sustained scrutiny. The merit of replication is highlighted and we discuss how the quality of the scientific literature may benefit from replication.

1 Introduction

There is a strong degree of agreement in climate sciences on the question regarding anthropogenic climate change. Anderegg et al. (2010) suggested that 97–98 % of the actively publishing climate researchers support the main conclusions by the Intergovernmental Panel on Climate Change (IPCC) (IPCC 2007). Cook et al. (2013) reviewed nearly 12,000 climate abstracts and received 1200 self-ratings from the authors of climate science publications. Based on both the abstracts and the self-ratings, they found a 97 % consensus in the relevant peer-reviewed climate science literature on humans causing global warming. This consensus was also noted by Oreskes (2004), yet a notable proportion of Americans doubt the anthropogenic cause behind the recent climate change (Leiserowitz et al. 2013; Doran and Zimmerman 2009). There is a lack of public awareness about the level of scientific agreement underpinning the view on anthropogenic global warming. Doran and Zimmerman (2009) reported that 52 % of Americans think that most climate scientists agree that the Earth has been warming in recent years, and 47 % think that climate scientists agree that human activities are a major cause of that warming. Theissen (2011) argued that many US undergraduate students are confused by a number of myths concerning climate change, propagated by blogs and media, and a similar

Electronic supplementary material The online version of this article (doi:10.1007/s00704-015-1597-5) contains supplementary material, which is available to authorized users.

✉ Rasmus E. Benestad
rasmus.benestad@met.no

¹ The Norwegian Meteorological Institute, Oslo, Norway

² Skeptical Science, Brisbane, Australia

³ University of Bristol, Bristol, UK

⁴ University of Western Australia, Crawley, Australia

⁵ Texas Tech University, Lubbock, TX, USA

⁶ The Royal Netherlands Meteorological Institute, De Bilt, The Netherlands

⁷ Global Change Institute, The University of Queensland, St. Lucia, Australia

⁸ University of Western Australia, Crawley, Australia

“consensus gap” exists in other countries, for example Australia (Leviston et al. 2012; Lewandowsky et al. 2013).

This gap of perception can be traced in part to a small number of contrarian papers that have appeared in the scientific literature and are often cited in the public discourse disputing the causes of climate change (Rahmstorf 2012). The message from these has been picked up by the media, a number of organizations, and blogs and has been turned into videos. For instance, a claim that the atmospheric greenhouse effect is “saturated” by a Canadian organization called “Friends of Science” is supported by one contrarian paper (Miskolczi 2010). A handful of papers (Shaviv 2002; Svensmark 1998; Friis-Christensen and Lassen 1991; Marsh and Svensmark 2000) have provided a basis for videos with titles such as “The Global Warming Swindle” and “The cloud mystery.” These have targeted the lay public, who have been left with the impression that greenhouse gases (GHGs) play a minor role in global warming and that the recent warming has been caused by changes in the sun. In the USA, the “Nongovernmental International Panel on Climate Change” (NIPCC) report (Idso et al. 2009; NIPCC 2013), the “Science & Environmental Policy Project” (SEPP), and the Heartland Institute have played an active role in the public discourse, providing a platform for the public dissemination of papers at variance with the notion of anthropogenic climate change. In Norway, there have been campaigns led by an organization called “Klimarealistene” which dismisses the conclusions drawn by the IPCC. This Norwegian organization has fed the conclusions from contrarian papers into schools through leaflets sent to the headmasters (Newt and Wiik 2012), following an example set by the Heartland Institute. They have also used a popular website (www.forskning.no) to promote such controversial papers targeting schools and the general public.

Misrepresentation of the climate sciences is a concern, and Somerville and Hassol (2011) have called for the badly needed voice of rational scientists in modern society. There have been attempts in the scientific literature to correct some misconceptions, such as a myth regarding an alleged recent “slow-down” in global warming, a so-called hiatus. Easterling and Wehner (2009) showed that natural variations give rise to reduced or even negative temperature trends over brief periods; however, this is due to stochastic fluctuations about an underlying warming trend (Foster and Rahmstorf 2011). Balmaseda et al. (2013) suggested that changes in the winds have resulted in a recent heat accumulation in the deep sea that has masked the surface warming and that the ocean heat content shows a steady increase. Examples of setting the record straight include both scientific papers (Legras et al. 2010; Masuda et al. 2006) and blogs such as Climate Dialogue (Vasileiadou 2013), ScepticalScience.com, and RealClimate.org (Rapley 2012).

The current situation for the climate sciences has been described as “a struggle about the truth of the state of climate”

(Romm 2010), and a number of books even claim that climate science myths have been introduced to society in a distorted way, causing more confusion than enlightenment (Oreskes and Conway 2008; Gelbspan 1997; Hoggan et al. 2009; Mooney 2006). Unjustified claims and harsh debates are not new (Sherwood 2011); history shows that they have been part of the scientific discourse for a long time. There are few papers in the literature that provide comprehensive analyses of several contrarian papers, and hence, a pattern of similarities between these may go unnoticed. Writing collections of replications of past papers is not the norm, but it is difficult to get published in journals with a set of expected formats or because of high likelihood that one reviewer does not like the implications or conclusions. Some journals do not even allow comments.

An interesting question is whether mistakes are random events or if a number of papers share common flaws of logic or methodology. We expect that scientific papers in general form networks by citing one another, and it is interesting to ask whether conclusions drawn in flawed papers are independent of each other or if errors propagate through further citation. To address this question, we need to identify errors through replicating previous work, following the line from the original information source, via analysis, to the interpretation of the results and the final conclusions, testing methods and assumptions. The objective of this paper is to present an approach to documenting and learning from mistakes. Errors and mistakes are often considered to be an essential ingredient of the learning process (Bedford 2010; Bedford and Cook 2013), creating potential learning material. The supporting material (SM) contains a number of case studies with examples of scrutiny and replication, providing an in-depth analysis of each paper (Benestad 2014a). Accompanying *open-source* software (also part of the SM) includes the source code for all of the analyses (Benestad 2014b, c). An important point is that this software too is open to scrutiny by other experts and, in the case of replication, represents the “hard facts” on which the SM and this paper are based.

2 Results

We review and summarize differences and common features of 38 contrarian papers that dispute anthropogenic global warming. We first grouped the papers into five categories describing how their conclusions depend the analytical setup, statistics, mathematics, physics, and representation of previous results. Most papers fell into the “analytical setup” category, and a common logical failure found in these papers was either starting with false assumptions or executing an erroneous analysis. Starting with false assumptions was common in the attribution

studies reporting that astronomical forcings influence Earth's climate. Examples of an erroneous analysis found in these papers included improper hypothesis testing and incorrect statistics.

A common feature across all categories was a neglect of contextual information, such as relevant literature or other evidence at variance with their conclusions. Several papers also ignored relevant physical interdependencies and consistencies. There was also a typical pattern of insufficient model evaluation, where papers failed to compare models against independent values not used for model development (out-of-sample tests). Insufficient model evaluation is related to over-fitting, where a model involves enough tunable parameters to provide a good fit regardless of the model skill. Another term for over-fitting is "curve fitting," and several such cases involved wavelets, multiple regression, or long-term persistence null models for trend testing. More stringent evaluation would suggest that the results yielded by several papers on our list would fail in a more general context. Such evaluation should also include tests for self-consistency or applying the methods to synthetic data for which we already know the answer.

False dichotomy was also a common theme, for example, when it is claimed that the sun is the cause of global warming, leaving no room for GHGs even though in reality the two forcings may coexist. In some cases, preprocessing of the data emphasized certain features, leading to logical fallacies. Other issues involved ignoring tests with negative outcomes ("cherry picking") or assuming untested presumed dependencies; in these cases, proper evaluation may reduce the risk of such shortcomings. Misrepresentation of statistics leads to incorrect conclusions, and "contamination" by external factors caused the data to represent aspects other than those under investigation. The failure to account for the actual degrees of freedom also resulted in incorrect estimation of the confidence interval.

One common factor of contrarian papers included speculations about cycles, and the papers reviewed here reported a wide range of periodicities. Spectral methods tend to find cycles, whether they are real or not, and it is no surprise that a number of periodicities appear when carrying out such analyses. Several papers presented implausible or incomplete physics, and some studies claimed celestial influences but suffered from a lack of clear physical reasoning: in particular, papers claiming to report climate dependence to the solar cycle length (SCL). Conclusions with weak physics basis must still be regarded as speculative.

3 Discussion and conclusions

Here, we focus on a small sample of papers that have made a discernable mark on the public discourse about climate change; they were selectively picked for close-up study rather

than randomly sampled for a statistical representation. Perhaps the most common problem with the cases examined here was missing contextual information ("the prosecutor's fallacy" (Wheelan 2013)), and there are several plausible explanations for why relevant information may be neglected. The most obvious explanation is that the authors were unaware of such facts. It takes experts to make proper assessments, as it requires scientific skills, an appreciation of both context and theory, and hands-on experience with computer coding and data analysis. There is tacit knowledge, such as the limitation of spectral methods and over-fitting, which may not be appreciated by newcomers to the fields of climate science and climatology. In some cases, the neglect of relevant information may be linked to defending one's position, as seen in one of the cases with a clear misrepresentation of another study (Benestad 2013). We also note that several of these papers involved the same authors and that the different cases were not independent even if they involved different shortcomings. Some of the cases also implied interpretations that were incompatible with some of the other cases, such as pronounced externally induced geophysical cycles and a dominant role of long-term persistence (LTP); slow stochastic fluctuations associated with LTP make the detection of meaningful cycles from solar forcing difficult if they shape the dominant character in the geophysical record.

There were also a group of papers (Gerlich and Tschuschner 2009; Lu 2013; Scafetta 2013) that were published in journals whose target topics were remote from climate research. Editors for these journals may not know of suitable reviewers and may assign reviewers who are not peers within the same scientific field and who do not have the background knowledge needed to carry out a proper review. The peer review process in itself is not perfect and does not guarantee veracity (Bohannon 2013). It is well known that there have been some glitches in the peer review: a paper by Soon and Baliunas (2003) caused the resignation of several editors from the journal *Climate Research* (Kinne 2003), and Wagner (Wagner et al. 2011) resigned from the editorship of *Remote Sensing* over the publication of a paper by Spencer and Braswell (2010). Copernicus Publications decided on 17 January 2014 to cease the publication of *PRP* to distance itself from malpractice in the review process that may explain some unusual papers (Benestad 2013; Scafetta 2013). The common denominators identified here and in the SM can provide some guidelines for future peer reviews.

The merit of replication, by reexamining old publications in order to assess their veracity, is obvious. Science is never settled, and both the scientific consensus and alternative hypotheses should be subject to ongoing questioning, especially in the presence of new evidence and insights. True and universal answers should, in principle, be replicated independently, especially if they have been published in the peer-reviewed

scientific literature. Open-source code and data provide the exact recipe that lead to conclusions, but a lack of openness and transparency may represent one obstacle to resolving scientific disputes and progress, such as a refusal to share the code to test diverging conclusions (Le Page 2009). Indeed, open-source methods are not the current norm for published articles in the scientific journals, and data are often inaccessible due to commercial interests and political reasons.

Bedford (2010) argued that “agnotology” (the study of how and why we do not know things) presents a potentially useful tool to explore topics where knowledge is or has been contested by different people. The term “agnotology” was coined by Proctor and Schiebinger (2008), who provided a collection of essays addressing the question “why we don’t know what we don’t know?” Their message was that ignorance is a result of both cultural and political struggles as well as an absence of knowledge. The counterpart to agnotology is epistemology, for which science is an important basis. The scientific way of thinking is an ideal means for resolving questions about causality, providing valuable guidance when there are conflicting views on matters concerning physical relationships.

It is widely recognized that climate sciences have profound implications for society (IPCC 2007), but one concern is that modern research is veering away from the scientific ideal of replication and transparency (Cartlidge 2013a; The Economist 2013). Unresolved disputes may contribute to confusion if one view is based on faulty analysis or logic (Theissen 2011; Rahmstorf 2012) and are especially unfortunate if the society has to make difficult choices depending on non-transparent knowledge, information, and data. High-profile papers, results influencing decision making, and controversial propositions should be replicated, and openness is needed to avoid non-epistemic consensus and “group think.” It is important that scrutiny and debate are sustained efforts and address both the scientific consensus and alternative views supported by scientific evidence. It is also important that critiques and debates are conveyed by the scientific literature when past findings are challenged. The assessments made by the IPCC could highlight the merit of replication, confirmation, and falsification; however, critics argue that it has failed to correct myths about climate research (Pearce (2010). The message from the IPCC assessment reports would be more robust if it also made available the source code and data from which its key figures and conclusions are derived. The demonstrations provided here in the SM may serve as an example (Pebesma 2012), and there are already existing examples where there is free access to climate data (Lawrimore et al. 2013; Cartlidge 2013b).

4 The method

The 38 papers selected for this study have all contributed to the gap in perception on anthropogenic climate change

between the general public and climate scientists. The sample was drawn based on expert opinion according to the criterion of being contrarian papers with high public visibility and with results that are not in agreement with the mainstream view. The sample was highly selective and meant for replication and the identification of errors rather being a representative statistical sample reflecting the volume of scientific literature. One objection to the selection of the cases here may be that they introduce an “asymmetry” through imperfect sampling; however, the purpose was not to draw general conclusions about the entire body of scientific literature but to learn from mistakes. There may also be flawed papers agreeing with the mainstream view, but they have little effect on the gap of perception between the public perception and the scientific consensus.

The analysis was implemented in the R environment (R Development Core Team 2004). This choice was motivated by the fact that R is free and runs on all common computer platforms, and has accessible online manuals and documentation. R is a scripting language with intuitive logic that provides the opportunity to create R packages (Pebesma 2012) with open-source computer code, user manual pages, necessary data, and examples. All the results and demonstrations presented in this paper are available in the R package “replicationDemos” (version 1.12) provided as supporting material (Benestad 2014b, c). In other words, it includes both the ingredients and the recipe for the analyses discussed here.

Acknowledgments Part of the material presented here has been inspired by posts written for www.RealClimate.org, but the paper has evolved over time based on input from comments in the review and discussion process, where several of the authors of the papers discussed here have had their say. We are also grateful for valuable comments from Oskar Landgren and Øyvind Nordli.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Proctor R N., Schiebinger L (eds) (2008). *Agnotology: the making and unmaking of ignorance*. Stanford University Press, Stanford
- Anderegg WRL, Prall JW, Harold J, Schneider SH (2010) Expert credibility in climate change. *PNAS* 107:12107–12110
- Balmaseda MA, Trenberth KE, and Källén E (2013) Distinctive climate signals in reanalysis of global ocean heat content. *Geophys Res Lett*. doi: [10.1002/grl.50382](https://doi.org/10.1002/grl.50382)
- Bedford D (2010) Agnotology as a teaching tool: learning climate science by studying misinformation. *J Geogr* 109:159–165

- Bedford D, Cook J (2013) Agnotology, scientific consensus, and the teaching and learning of climate change: a response to Legates, Soon and Briggs. *Sci & Educ* 22:2019–2030
- Benestad RE (2013) Comment on ‘Discussions on common errors in analyzing sea level accelerations, solar trends and global warming’ by Scafetta (2013). *Pattern Recogn Phys* 1:91–92
- Benestad R (2014a) Replication of a number of contrarian climate change studies. http://figshare.com/articles/Replication_of_a_number_of_contrarian_climate_change_studies/951971. Retrieved Aug 04, 2015 (GMT)
- Benestad R. (2014b) R-package for replication studies. http://figshare.com/articles/R_package_for_replication_studies/951973. Retrieved August 4, 2015 (GMT) (R-package for Windows)
- Benestad R. (2014c) R-package for replication studies. http://figshare.com/articles/R_package_for_replication_studies/951972. Retrieved August 4, 2015 (GMT) (R-package for Linux/Mac)
- Bohannon J (2013) Who’s afraid of peer review? *Science* 342:60–65
- Cartlidge E. (2013a) Opening data up to scrutiny. *Phys World* 13
- Cartlidge E (2013b) Opening data up to scrutiny. *Phys World* 13
- NIPCC (2013) Climate change reconsidered II: physical science. The Heartland Institute, Chicago
- IPCC (2007) Climate change: the physical science basis. Contribution of Working Group I to the fourth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, New York
- Romm J (2010) Climate of fear. *Nature* 464: 141. doi: 10.1038/464141a
- Cook J et al (2013) Quantifying the consensus on anthropogenic global warming in the scientific literature. *Environ Res Lett* 8:024024
- Doran PT, Zimmerman MK (2009) Examining the scientific consensus on climate change. *EOS Trans Am Geophys Union* 90:22–23
- Easterling DR, Wehner MF (2009) Is the climate warming or cooling? *Geophys Res Lett* 36:L08706
- Foster G, Rahmstorf S (2011) Global temperature evolution 1979–2010. *Environ Res Lett* 6:044022
- Friis-Christensen E, Lassen K (1991) Length of the solar cycle: an indicator of solar activity closely associated with climate. *Science* 254: 698–700
- Gelbspan R (1997) The heat is on: the high stakes battle over Earth’s threatened climate. Addison-Wesley Pub. Co., Boston
- Gerlich G, Tschuschner RD (2009) Falsification of the atmospheric CO₂ greenhouse effects within the frame of physics. *Int J Mod Phys B* 23: 275–364
- Hoggan J, Littlemore RD, and Ebrary I (2009) Climate cover-up: the crusade to deny global warming. Greystone Books, Vancouver
- Idso C, Singer F (2009) Climate change reconsidered: 2009 report of the Nongovernmental Panel on Climate Change (NIPCC). The Heartland Institute, Chicago
- Kinne O (2003) Climate research: an article unleashed worldwide storms. *Clim Res* 24:197–198
- Lawrimore J, Rennie J, Thorne P (2013) Responding to the need for better global temperature data. *EOS Trans Am Geophys Union* 94:61–62
- Le Page M (2009) Sceptical climate researcher won’t divulge key program. *New Scientist*. <http://www.newscientist.com/article/dn18307-sceptical-climate-researcher-wont-divulge-key-program.html#.Uq1qQ6rj6ak>. Accessed 8 April 2015
- Legras B, Mestre O, Bard E, Yiou P (2010) A critical look at solar-climate relationships from long temperature series. *Clim Past* 6:745–758
- Leiserowitz A. et al (2013) Public support for climate and energy policies in April 2013. Yale University, Connecticut
- Leviston Z, Walker I, Morwinski S (2012) Your opinion on climate change might not be as common as you think. *Nat Clim Change* 3:334–337
- Lewandowsky S, Gignac GE, Oberauer K (2013) The role of conspiracist ideation and worldviews in predicting rejection of science. *PLoS One* 8:e75637
- Lu Q-B (2013) Cosmic-ray-driven reaction and greenhouse effect of halogenated molecules: culprits for atmospheric ozone depletion and global climate change. *Int J Mod Phys B* 27:1350073
- Marsh ND, Svensmark H (2000) Low cloud properties influenced by cosmic rays. *Phys Rev Lett* 85:5004–5007
- Masuda K, Asuka J, Yoshimura J, Kawamiya M (2006) Critical comments on several skeptic views about global warming. *J Jpn Sci* 41: 36–41
- Miskolczi F. (2010) The stable stationary value of the Earth’s global average atmospheric Planck-weighted greenhouse-gas optical thickness. *Energ Environ* 21: 243–263
- Mooney C (2006) The Republican war on science. Basic Books, New York
- Newt M, Wiik H (2012) Når akademikere villeder om klimaendringer. *Utdanning nr. 15/2012*
- Oreskes N (2004) The scientific consensus on climate change. *Science* 306
- Oreskes N, Conway EM (2008) Challenging knowledge: how climate science became a victim of the Cold War. In: Proctor R, Schiebinger L Agnotology: the making and unmaking of ignorance. Stanford University Press, Stanford
- Pearce F (2010) The climate files: the battle for the truth about global warming. Random House, UK
- Pebesma E, Nüst D, Bivand R (2012) The R software environment in reproducible geoscientific research. *EOS* 93:163–164
- Rahmstorf S (2012) Is journalism failing on climate? *Environ Res Lett* 7: 041003
- Rapley C (2012) Climate science: time to raft up. *Nature* 488:583–585
- Scafetta N (2013) Discussion on common errors in analyzing sea level accelerations, solar trends and global warming. *Pattern Recogn Phys* 1:37–57
- Shaviv NJ (2002) Cosmic ray diffusion from the galactic spiral arms, iron meteorites, and a possible climatic connection. *Phys Rev Lett* 89: 051102
- Sherwood S (2011) Science controversies past and present. *Phys Today* 64:39–44
- Somerville RCJ, Hassol SJ (2011) Communicating the science of climate change. *Phys Today* 64:48–53
- Soon W, Baliunas S (2003) Proxy climatic and environmental changes of the past 1000 years. *Clim Res* 23:89–110
- Spencer RW, Braswell WD (2010) On the diagnosis of radiative feedback in the presence of unknown radiative forcing. *J Geophys Res* 115: D16109
- Svensmark H (1998) Influence of cosmic rays on Earth’s climate. *Phys Rev Lett* 81:5027–5030
- R Development Core Team (2004) R: a language and environment for statistical computing. Vienna, Austria. <http://www.R-project.org>. Accessed 5 July 2015
- Theissen KM (2011) What do U.S. students know about climate change? *EOS* 92:477–478
- The Economist (2013) How science goes wrong. <http://www.economist.com/news/leaders/21588069-scientific-research-has-changed-world-now-it-needs-change-itself-how-science-goes-wrong>. Accessed 8 April 2015
- Vasileiadou E (2013) Final evaluation report of climate dialogue. <http://www.climatedialogue.org/final-evaluation-report-of-climate-dialogue/>. Accessed 8 April 2015
- Wagner W (2011) Taking responsibility on publishing the controversial paper ‘On the misdiagnosis of surface temperature feedbacks from variations in Earth’s radiant energy balance’ by Spencer and Braswell, *Remote Sens.* 2011, 3(8), 1603–1613. *Remote Sens* 3: 2002–2004
- Wheeler CJ (2013) Naked statistics: stripping the dread from the data. W. W. Norton & Company, New York