Earth Syst. Dynam. Discuss., 4, 451–505, 2013
www.earth-syst-dynam-discuss.net/4/451/2013/
doi:10.5194/esdd-4-451-2013

Earth System Dynamics Discussions Open Access

Agnotology: learning from mistakes

R. E. Benestad[1], H. O. Hygen[1], R. van Dorland[2], J. Cook[3], and D. Nuccitelli[4]

[1]The Norwegian Meteorological Institute, Norway
[2]The Royal Netherlands Meteorological Institute, The Netherlands
[3]Global Change Institute, the University of Queensland, Australia
[4]Tetra Tech, Inc., McClellan, California, USA

Received: 4 April 2013 – Accepted: 22 April 2013 – Published: 3 May 2013

Correspondence to: R. E. Benestad (rasmus.benestad@met.no)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Replication is an important part of science, and by repeating past analyses, we show that a number of papers in the scientific literature contain severe methodological flaws which can easily be identified through simple tests and demonstrations. In many cases, shortcomings are related to a lack of robustness, leading to results that are not universally valid but rather an artifact of a particular experimental set-up. Some examples presented here have ignored data that do not fit the conclusions, and in several other cases, inappropriate statistical methods have been adopted or conclusions have been based on misconceived physics. These papers may serve as educational case studies for why certain analytical approaches sometimes are unsuitable in providing reliable answers. They also highlight the merit of replication. A lack of common replication has repercussions for the quality of the scientific literature, and may be a reason why some controversial questions remain unanswered even when ignorance could be reduced. Agnotology is the study of such ignorance. A free and open-source software is provided for demonstration purposes.

1 Introduction

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 for the first time coined in Proctor and Schiebinger (2008), which provided a collection of essays addressing the question "why we do not know what we do not 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. In principle, the scientific way of thinking is the ideal means of resolving questions about causality, and science can provide valuable guidance when there are conflicting views on matters concerning physical relationships. One of the scientific virtues is debate and disagreement about different hypotheses, making it dynamic and providing a driving force for progress. In this process, controversial questions should be addressed with the scientific method and rigour, and in order to provide convincing answers, it is important that the process is transparent, the results are replicable, the hypotheses testable, and the tests objective. It is also important that critiques and debates are conveyed by the scientific literature when past findings are challenged.

An agnotological study of the climate sciences can shed light on some recent controversies which take place when groups with different scientific backgrounds and mindsets dispute each other's conclusions. Some case studies may include a number of recent papers that have suggested a strong influence on Earth's climate from solar variability, Jupiter, Saturn or the lunar orbit (Friis-Christensen and Lassen, 1991; Scafetta, 2010; Scafetta and West, 2008, 2007, 2006a, b, 2005; Svensmark, 1998; Svensmark and Friis-Christensen, 1997). These papers have also argued that greenhouse gases (GHG) such as CO_2 play a relatively small role for Earth's climate, and dispute the view presented by the mainstream climate research community (National Research Council (US), 2001; Oreskes, 2004; Solomon et al., 2007). In this respect, it is important to ask whether these differences reflect legitimate uncertainties and gaps in our knowledge. In order to get to the bottom of such issues, one needs to follow the line from the original information source, via analysis, to the interpretation of the results and the final conclusions. True and universal answers should in principle be replicated independently, especially if they have been published in the peer reviewed scientific literature. A message from Proctor and Schiebinger (2008) is that ignorance in these issues may stem from the culture neglecting replication, not sharing methods and data, or not testing the methods in different settings. The most persuasive arguments are the ones where everybody can repeat the analyses for themselves, examine the methods, and get the same results. Scientific truths should in principle be universal, which means that they should be generally valid and the methods objective.

All the examples discussed in this paper have been cited in the public discourse to dispute the causes of climate change, where their conclusions have tended to be

taken at face value without replication or verification. Many of the examples here have been used to back up claims in the public discourse on climate in the media (Rahmstorf, 2012). A high proportion of Americans doubt the anthropogenic cause behind the recent climate change and seem to be unaware about the level of scientific agreement underpinning the view about anthropogenic global warming. Doran and Zimmerman (2009) reported that 52 % of Americans think most climate scientists agree that the Earth has been warming in recent years, and 47 % think climate scientists agree that there is a scientific consensus about human activities being a major cause of that warming. Anderegg et al. (2010), on the other hand, presented a survey that suggested that 97–98 % of the climate researchers most actively publishing in the field support the main conclusions by the IPCC. Cook et al. (2013) reviewed nearly 12 000 climate abstracts and received 1200 self-ratings from the authors of climate science publications. Using both methodologies, they found a 97 % consensus in the peer-reviewed climate science literature that humans are causing global warming. There appears to be a gap in the understanding of the climate between experts and the lay public, and a common denominator between all the examples reported here and in the supporting material is that they all represent a contribution towards the agnotology associated with the climate change issue.

In the US, the "Nongovernmental International Panel on Climate Change" (NIPCC) report (Idso and Singer, 2009), the "Science & Environmental Policy Project" (SEPP) and the Heartland Institute have played an active role in the public discourse, promoting the ideas from some of these cases. In Norway, there have been campaigns led by an organisation called "klimarealistene", who dismiss the conclusion drawn by the Intergovernmental Panel on Climate Change (IPCC), to feed the conclusions from some of these cases into schools through leaflets sent to the headmaster (Newt and Wiik, 2012). A popular web site targeting schools (www.forskning.no) has brought together the ideas from Scafetta, Humlum, Solheim, and Stordahl (dated 30 April 2012, 17 April 2012, 20 March 2012, and 29 October 2011). Humlum et al. have also written extensively for popular science and engineering magazines ("Fra Fysikkens Verden", 1 November, 21–26, 2011; "Teknisk ukeblad", numbers 1011, 1611, and 2711), in which they have promoted the purports from a number of the agnotological cases presented here.

There are also some accounts in Proctor and Sciebinger (2008) suggesting deliberate attempts to manufacture doubts and controversies about well-established scientific conclusions, such as about climate change. Some of these efforts have targeted the scientific literature, being viewed as an authority on technical and scientific questions. When detailed replication and scrutiny shed light on some of these controversies, it is evident that peer review publication by itself does not assure validity. Indeed, some of these cases beg the question whether peer-reviewing has been sufficiently valued. 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 (2011) resigned from the editorship of Remote Sensing over the publication of Spencer and Braswell (2010). These papers have not been retracted, however, correction or errata are expected to be published when severe flaws are discovered to avoid that others unfamiliar with the papers later on base their work on incorrect information. A continuous replication of published results and dissemination through scientific fora can nevertheless contribute towards a convergence towards the most convincing explanations. However, some journals do not allow comments, and single comments may not pick up patterns of related papers, issues and authors.

There are few papers in the literature providing a comprehensive views of a several papers, rather than responding to single papers (such as "comments") and hence a pattern of similarities between these may go unnoticed. There may also be some misgivings against direct criticism of others' work in the fear of being inflammatory or "unfair" to the original authors if their work is criticised elsewhere, even if all published scientific results in principle should be up for scrutiny.

Here we show how knowledge may progress and how ignorance may be reduced for a number of controversial analyses. We attempt to provide a comprehensive review by

examining the methods used in an effort to replicate the results of a range of different studies, highlighting the value of replication. As a result, we attempt to make a contribution to agnotology. The emphasis must be on specific and detailed scientific/technical aspects in order to understand why different efforts lead to different results. For example depending on how the analytical set-up is designed, the application of statistics, in addition to physics considerations. Hence the need to examine specific examples and go into the details in order to understand why certain conclusions are drawn. A mere focus on general points rather than detailed replication may not lead to enhanced understanding for why we do not know what we do not know (Proctor and Schiebinger, 2008). We draw on a list of publications which is chosen because they represent good examples of the methodological flaws which we seek to clarify. The choice is also based on clear-cut examples for which there are easy demonstrations and obvious logical shortcomings. All these cases have also contributed to the public confusion around anthropogenic climate change. Often the methods can be tested (Pebesma et al., 2012), and some of these claims have already been revealed as flawed analysis (Benestad and Schmidt, 2009).

Open-source algorithms for data analyses and data sharing in the first place could in some cases prevent such situations. Science often involves trial and error, and making mistakes is sometimes unavoidable. Hence, the identification of flaws in past work should be regarded as progress. Sharing of data and computer code used in analysis lead to more robust understanding, as the computer codes provide the exact recipe that lead to conclusions. Our own methods used in this paper are disclosed in the form of computer code. We show how replications using an R-package can be applied through a review of the literature, methods and analytical setup.

The outline of this paper is as follows: a description of our methods, a new freely available and open-source software for an educational tool kit demonstrating flaws in some of past analyses. This section is followed by some examples based on replication through an open-source software. A discussion towards the end of the paper summarises our experience, and tries to outline a direction for further progress.

2 The methods

The analysis was implemented in the R environment (R Development Core Team, 2004; version 2.13.1), which is a free software that runs on most platforms (Linux, Mac, Windows), and provides open-source access to the computer code in addition to user manual pages and examples. All the results and demonstrations presented in this paper are available in the R-package "replicationDemos" (version 1.10) provided as supporting material at CRAN[1]. Moreover, replicationDemos contains several different functions replicating the papers discussed here, in addition to providing both the documentation of the methods, the necessary data, and the open-source code itself. In other words, it provides both the ingredients and the recipe for the analyses presented here. The details about the functions in replicationDemos are provided in Table 1, and when reference is made to Table 1 in the text below, it will serve as a description of which functions in the replicationDemos package that were used for a particular analysis. More details about the methods are also given in the Appendix.

3 Replication demonstrations

3.1 Case 1: ignoring data which do not agree with the conclusions.

Humlum et al. (2011a) suggested that the moon and the giant planets in the solar system play a role a role in climate change on Earth, and that their influence is more important than changes in the GHG. A replication of their analysis can provide a means for turning these controversies into an educational exercise, and hence, provide a link to agnotology.

The core of the analysis carried out by Humlum et al. (2011a) involved curve-fitting and tenuous physics, with a vague idea that the gravity of solar system objects somehow can affect the Earth's climate. The most severe problem with the paper, however,

[1]http://cran.r-project.org/web/packages/replicationDemos/

was that it had discarded a large fraction of data for the Holocene which did not fit their purports. Their reason for not showing the part of the data before 4000 BP was that they "*chose to focus on the most recent 4000 years of the GISP2 series, as the main thrust of [their] investigation is on climatic variations in the recent past and their potential for forecasting the near future*" (square brackets here denotes replacing "our" with "their"). Humlum had also been a co-author on an older article in a popular technical magazine where this absent part of the data had been presented (Bye et al., 2011), and the data stretches almost 50 000 years back in time and is downloaded in one single file[2].

Humlum et al. (2011a) examined the last 4000 yr of the GISP2 (Greenland Ice Sheet Project Two) record, and constructed a mathematical model based on a set of Fourier components and only three periods: 2804, 1186, and 556 yr (for Svalbard annual mean temperature since 1912, they found Fourier components of 68.4, 25.7, and 16.8 yr). Fourier series are often discussed in science textbooks and it is a mathematical fact that any finite series can be represented in terms of a series of sinusoids, which easily can result in mere "curve-fitting" (Fourier expansion in this case). According to Stephenson (1973, p. 255), the sum of Fourier series is not necessarily equal to the function $f(x)$ from which is derived, since the function given by the Fourier expansion is mathematically bound to extend periodic regularity (known as the Dirichlet conditions). Moreover, most functions $f(x)$, defined for a finite interval, are not periodic, although it is possible to find a Fourier series that represents this function in the given interval (Williams, 1960, p. 74). Pain (1983; p. 252) also clearly states that the Fourier series represent the function $f(x)$ only within the chosen interval, and one can fit a series of observations to arbitrary accuracy without having any predictability at all. This is a form of "over-fit" (Wilks, 1995), and therefore it is important to verify model to data outside the fitted region.

[2]ftp://ftp.ncdc.noaa.gov/pub/data/paleo/icecore/greenland/summit/gisp2/isotopes/gisp2_temp_accum_alley2000.txt

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

The underlying data used in (Humlum et al., 2011a) analysis violate the Dirichlet conditions, and their analysis is replicated here through replicationDemos. They claimed that they could "produce testable forecasts of future climate" by extending their statistical fit, and in fact, they did produce a testable forecast of the past climate by leaving out the period between the end of the last ice age and up to 4000 yr before present. However, they did not state why the discarded data was not used for evaluation purposes, and the problem with their model becomes apparent once their fit is extended to the part of data that they left out.

We extended their analysis back to the end of the last ice age. Figure 1 shows our replication of part of their results (Table 1). It clearly shows that the curve-fit for the selected 4000 yr does not provide a good description for the rest of the Holocene. The full red line shows their model results and the dashed red lines show two different attempts to extend their model to older data. One initial attempt was made, keeping their trend; however this obviously caused a divergence. So in the second attempt, we removed the trend to give their model a better chance of making a good hindcast. Again, the fit is no longer quite as good as presented in their paper. Clearly, their hypothesis of 3 dominant periodicities no longer works when extending the data period, and this is not surprising as this is explained in text books on Fourier methods.

In other words, the analysis made by Humlum et al. (2011a) was limited to a subset of the data, but they did not use the remaining part to evaluate their model. They ignored the data which did not agree with their conclusions. Moreover, a lack of being universally valid suggests that the chosen method was not objective. Furthermore, they failed to acknowledge well-known shortcomings associated with curve-fitting, but rather based their analysis on unjustified fit to a set of Fourier series. Finally, their results lacked a well-formulated physical basis, and they failed to discuss past relevant literature concerning the physics as well as mathematics.

The same flawed approach was applied by Humlum et al. (2011a) to the Svalbard temperature and the results in Humlum et al. (2011b). The Humlum et al. (2011a) paper also provides a nice demonstration for why similar types of curve-fitting employed in

Scafetta (2012a, b, c) and Loehle and Scafetta (2011; henceforth "L&S2011") fail to provide reliable answers.

3.2 Case 2: unclear physics and non-objective analytical design

Scafetta (2012a) too argued that celestial forcing in the form of gravitational forces from the giant gas planets explains most of the past climatic changes on Earth, and especially fluctuations of ~20 and ~60 yr. He then evaluated how well global climate models reproduce the amplitude and phase of ~20 and ~60 yr periodicity, which he attributed to the influence of gravity from celestial objects. In addition, he carried out an evaluation of trends based on an arbitrary curve fitting, using different trend models for different parts of the data, which apparently gave a good fit to the data. Although the physics was vague, Scafetta argued that resonant response could amplify the weak effect from the planets, just like L&S2011. In addition to vague physics, many of the statistics presented in the paper were miscalculated. By repeating the work done by Scafetta, we can understand why his purports diverge from the mainstream climate science. In this sense, this paper is a good agnotological example

Scafetta (2012a) can be reviewed in terms of the physics and the statistical analysis. The paper failed to acknowledge that resonance is an inherent property of a system, and will pick up any forcing with matching frequency. Our replications demonstrated that the paper presented an inappropriate analytical setup which favoured one outcome due to its design.

A weak forcing and a pronounced response would imply a positive feedback, or at least an optimal balance between forcing periodicity and damping rate (a 60 yr periodicity would suggest very weak damping, which seems unlikely, and hence the most convincing argument for resonance would involve a delayed positive feedback), and a preferred frequency would be an inherent characteristic of the earth climate system. Noisy forcings embed a range of frequencies, as well as transient functions, and can therefore feed a resonance. Through our replication toolbox, we can demonstrate such cases, where a simulation of a forced damped oscillator picks up resonant variations if given a noisy forcing, even if the forcing itself has another dominant frequency (Table 1). Furthermore, a resonant system will respond to a trend in GHG forcings (in mathematical terms, the forcing is proportional to ln|CO2|), and if such a resonance implies positive feedbacks, these should also be present in a situation of GHG forcings. Hence it is extremely hard to attribute a cause for resonant response just from analysing cycles when several forcings are present.

Another weakness in the analysis presented in Scafetta (2012a) is the handling of trends, as a quadratic trend that conveniently fitted the data was used for the period 1850–2000, and then a linear fit with a warming rate of 0.009 °C yr^{-1} was used after 2000. The quadratic equation for 1850–2000 $p(t) = 4.9 \times 10^{-5}\ x^2 – 3.5 \times 10^{-3} x – 0.30$ (Eq. 4, where $x = t - 1850$) gave a warming rate $dp/dt = 2 \times 4.9 \times 10^{-5} x – 3.5 \times 10^{-3} = 0.011$ °C yr^{-1} for year 2000. Hence, the method used by Scafetta implicitly assumed that the rate of warming was abruptly reduced in year 2000 for the future. It also implied that the future warming rate was smaller than the range reported in Solomon et al. (2007), and much of the recent warming was mis-attributed to natural variations based on curve-fitting similar to that of Humlum et al. (2011a).

There seem to be a number of other results in Scafetta (2012a) which are difficult to reproduce, as a replication of Fig. 5b in the paper (Table 1) suggests that it displays a lower projected trend than produced by the equations cited in the paper (Fig. 2). He also limited the confidence interval to one standard deviation (which implies a 68.6 % confidence interval) in the evaluation to see whether the model results overlapped the observations (the more commonly used 95th confidence is roughly spanned by 2 times the error estimate). Other mistakes in the paper included a misapplication of the chi-squared test used to asses the global climate models (GCMs) against the observations, where Scafetta used the squared error-estimates in the denominator; conventional chi-squared tests do not square the denominator, see e.g. Wilks (1995) and Press et al. (1989).

The gravest issue with the Scafetta (2012a) analysis involved a series of tests which in effect were “rigged” to give negatives. They involved a regression analysis to estimate

amplitude and phase of 20 and 60 yr oscillations in the global mean temperatures, assuming that these were due to the gravitational influence from celestial bodies. The phase and amplitudes found for the observations then were used as a yard stick for the GCM results, and a regression analysis was used where the covariates were the same as for the observations, with exactly the same phase and amplitude specified for the 20 and 60 yr oscillations. We know a priori that the planets are not accounted for in the CMIP3 climate simulations (Meehl et al., 2007), and hence Scafetta’s strategy is not suitable to provide an objective answer. A more appropriate null hypothesis would be that the amplitudes seen for the 20 and 60 yr variations would be due to noise. Hence, it is important to allow the phase to be unconstrained in the analysis, as we have done (Fig. 3). When we repeat the analysis using a suitable setup, we do not see a falsification of the null-hypothesis, especially if we account for the fact that the analysis involves multiple tests and take the field significance into account (Wilks, 2006).

Scafetta (2012a) assumed that his method was validated if it was calibrated on one cycle of 60 yr and then was able to reproduce the next 60 yr cycle in the data that was not part of the calibration. However, this argument is not justified, as this type of approach fails for the El Niño Southern Oscillation (ENSO), for which it is well known that there were two El Niños during the 1980s, which taken together, resemble two periods of a periodic cycle (Fig. 4). While each event provides a good fit for the other cycle, if calibration is performed on half of the decade and evaluated against the other, the predictions by this regression model fails to capture the variations outside this interval. It is therefore important to capture many cycles in a time series before one can establish a periodic signal, as only two cycles will likely not be representative of the entire system.

In other words, the results in Scafetta (2012a) were incorrect due to inappropriate strategy for which the one answer was favoured, in addition to wrong statistics. The same type of shortcomings were also present in the curve-fitting presented in Scafetta (2012b). Furthermore, a clear physical basis was lacking.

4 Discussion and conclusions

More examples, listed in Table 1, are provided in the appendix. These involve issues such as ignoring tests with negative outcomes, insufficient model evaluation, untested presumed dependencies, misrepresentation of statistics, a failure to account for the actual degrees of freedom, lacking similarities, implausible physics, circular reasoning, incorrect interpretation of mathematical concepts, differences in data processing, selective use of data, ignoring relevant scales, and contamination by external factors. Some common traits of unpersuasive papers may include speculations about cycles, and it is apparent from Table 2 that there are various claims which together involve 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.

The merit of replication, by re-examining old publications in order to asses their veracity, is obvious. Published results in particular should be replicable, and access to open source codes and data should be regarded as a scientific virtue that facilitates more reliable knowledge. Results are far more persuasive if one can reproduce them oneself, although replication of published results requires scientific training, numerical skill, and mastering of statistics. One concern is that modern research is veering away from the scientific virtue of replication and transparency. Open-source methodology often does not belong to categories of articles in the scientific journals, and data is often inaccessible due to commercial interests and political reasons. Such limitations are especially unfortunate if the society at large has to make difficult choices depending on nontransparent knowledge, information, and data. It is widely recognised that climate sciences have profound implications for society (Solomon et al., 2007), and that the communication of misleading claims is a case of agnotology. Furthermore, the agnotology does not involve just one question or issue, but is embodied in a range of aspects that at first sight may seem unrelated. However, these all contribute to agnotology for instance through media (Rahmstorf, 2012; Theissen, 2011).

According to Sherwood (2011), unjustified claims and harsh debates are not new; history shows that they have been part of the scientific scene for a long time. Science is about trial and error, where errors and mistakes may be seen as an inevitable part of the process of learning, and may be valuable as a pedagogic tool to improve understanding of the science (Bedford, 2010). Academic societies and institutions could take a greater role in ensuring that the society gets the best information and knowledge that can be derived from science, by acting as trusted, respected, neutral and independent bodies replicating the results of high-profile papers or those influencing policy-makers, cited in the media and blogs. They could take science to the society (“domesticating science”), and show what kind of detective work lies behind the results. Hence demonstrate how the conclusions are reached, based on the scientific virtues of transparency, replicability, testability, and objectivity. The interest in resolving contested issues by sharing analysis codes, however, has sometimes been low (Le Page, 2009), and the demonstrations provided here may serve as an example of how past work can be re-assessed (Pebesma et al., 2012). The IPCC could also have played such a role; however, its target group has not been the general public, and critics argue that it has failed to correct the myths about climate research (Pearce, 2010). Another issue is that the source code used and data for producing the figures and tables in the IPCC’s assessment reports could be made openly available, in the same vein as the replicationDemos software and as proposed by Pebesma et al. (2012). There are already some examples where there is free access to climate data (e.g. Lawrimore et al., 2013).

A1 Method – further details

All the data sets contained in replicationDemos are provided with the attribute “url” which identifies their data source on the Internet. The R-package contains examples and syntax descriptions in addition to the source code for the R-scripts and data. The numbers from various tables discussed here have been copied electronically from the PDF-version of the respective papers.

The installation of the R-package can be done through a few command lines in R (here the R-prompt ">" is shown) – start R and then write these lines in the R window:

```
> install.packages("replicationDemos")
> library(replicationDemos)
```

The installation of the R-package requires Internet access. Some of the functions also read data directly over the Internet. Table 1 gives an overview of the cases which can be replicated through the replicationDemos package. Some of these may take some time if a large set of Monte-Carlo simulations is carried out. The replications should be possible on different platforms (Linux, Mac, and Windows). The replications of case 1 can be implemented through typing the command line in R:

```
> Humlum.et.al.2011()
```

The source-code is produced with the following line (the name of the functions without the parentheses "()"):

```
> Humlum.et.al.2011
```

The manual page for the R-function replicating the predictions by Humlum et al. (2011) are provided through:

```
> ?Humlum.et.al.2011
```

Likewise, cases 2–3 can be replicated and studied in further detail through the functions:

```
> LoehleScafetta2011()
> Scafetta2011()
```

User guides for R are freely available as PDF documents, e.g. from the CRAN site. The R-packages replicationDemos is hosted by official R-web site (http://cran.r-project.org),

which provides an archive of up-to-date as well as a history of past version (older versions of replicationDemos are already stored there). Although this archive is away from where the journal archives its papers, it ensures a wider visibility of the package among users of R. Any errors in the code may be fixed in new versions of the R-package that will be uploaded to the archive, however, the version control of the submitted R-packages ensures traceability.

A2 Additional examples

Here an extended list of examples is provided, both which are included in the replicationDemos package and cases which do not need replication in order to assess. All of these relate to agnotology as they have been used to support arguments presented in the media and on the Internet. Many of these have been compiled in reports such as the NIPCC (Idso and Singe, 2009) and several blogs.

A2.1 Case 3: unclear physics and misappropriate curve-fitting

Loehle and Scafetta (2011; L&S2011) purported that 20 and 60 yr natural cycles in the global mean temperature estimates were due to natural cycles which they explained in terms of solar and astronomical influences. Furthermore, they claimed there was only a weak linear trend in the global mean temperature, and explained this in terms of a slight negative feedback in the climate system to CO_2. At the same time, it is easy to reproduce the analysis and demonstrate why the conclusions drawn by L&S2011 are at variance with most of the climate research community, which also makes this case a good agnotological example.

The problem with the L&S2011 includes both a lack of clear physical basis and the analytical setup. L&S2011 assumed some kind of selective and potent resonance to solar and astronomical forcing while a negative feedback was acting for CO_2. Resonance is inherent to the system, and it is difficult to conceive what it would entail that differed for the different types of forcings. The analytical problem involved a curve-fit, and 60 yr cycles were estimated from a mere 160 yr of data. In a complex and non-linear system such as Earth's climate, such an exercise is prone to be non-robust and non-stationary. Furthermore, the analytical setup was not validated against independent data, and the skill of the model was not properly assessed.

L&S2011 assumed similar resonance as Scafetta (2012a), with the same weaknesses. Furthermore, the methods used in L&S2011 suffered from many of the similar flaws as those in Humlum et al. (2011a), even though L&S2011 employed a different strategy for spectral analysis. The type of analysis in L&S2011 was repeated in Scafetta (2012a). Again, it is important to keep in mind that all curves (finite time series) can be represented as a sum of sinusoids describing cycles with different frequencies (Table 1). Furthermore, Fourier transforms are closely related to spectral analysis, but these concepts are not exactly the same. Spectral analysis also tries to account for mathematical artifacts, such as "spectral leakage" (Press et al., 1989), attribute probabilities that some frequencies are spurious, and estimate the significance of the results. There is a number of different spectral analysis techniques, and some are more suitable for certain types of data. Sometimes, one can also use regression to find the best-fit combination of sinusoids for a time series, as in L&S2011's "empirical decomposition" (Table 1). It is typical, however, that geophysical time series, such as the global mean temperature, are not characterised by one or two frequencies. In fact, if we try to fit other sinusoids to the same data as L&S2011, we get many other frequencies which fit equally well, and we see that the frequencies of 20 and 60 yr are not the most dominant ones. A trial with a range of periodicities for harmonic fitting in a similar regression analysis suggested that periodicities of 65.75 and 21.5 yr gave a better fit in terms of value explained (R^2) than 60 and 20 yr respectively (Table 1).

Fitting sinusoids with long time scales compared to the time series is careless, which can be demonstrated through constructing a synthetic time series that is much longer than the one we just looked at. It is important that the synthetic series is constructed from a combination of sinusoids for the entire period but with random amplitude and phase. We can divide this synthetic time series into sequences with the same length as

that L&S2011 used to fit their model, and compare the fits for each segment (Table 1). The 20 and 60 yr amplitude estimates vary substantially from sequence to sequence when we adopt the same strategy as in L&S2011 (Table 1), and the amplitude for the fits to the shorter sequences will typically be 4 times greater than a similar fit gives for the original 10 000 yr long series. This is because there is a band of frequencies present in random, noisy and chaotic data, which brings us back to our initial point: any number or curve can be split into a multitude of different components, most of which will not have any physical meaning.

The analysis presented in L&S2011 can be described as a curve-fitting exercise based on two periods and that assumed that cycles with constant frequency in a non-linear and chaotic system. The paper also failed to provide a persuasive account of the physics behind the purported links.

A2.2 Case 4: ignoring negative tests

Solheim et al. (2011; SSH2011) argued that 60 % of the annual and winter temperature variations at Svalbard are related to the solar cycle length (SCL). The basis for their conclusions was a high correlation estimated between SCL and the temperature estimates, and results from a Durbin-Watson test. The highest correlation reported were −0.82 for the winter mean of the decade lagging one solar cycle. Repeating their analysis with our open-source agnotological toolkit gave different answers. The conclusions from this paper has been disseminated by the organisation "klimarealistene", who also try to reach the Norwegian schools. In order to shed light on the agnotological aspects of this case, we need to replicate their work.

The conclusion of the paper lacked clear physical basis, as the chain of processes linking the solar cycle length and temperatures in the Arctic over the subsequent decade is not understood. Furthermore, the analysis was not objective, inflating the significance of the results. A more subtle aspect of this study was the number of attempts to find a correlation, and the lack of accounting for all the tests in the evaluation

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

of the significance of the results. There is a good chance of seeing false fortuitous correlations if one examines enough local temperature records.

When we reconstructed their Table 1 we got nearly the same results, albeit not identical. SSH2011 stated that they based their method for estimating SCL on a publication from 1939 (Waldmeier, 1961), however, more recent work on the estimation of SCL account for uncertainties in estimating the true SCL as the sunspot record exhibits stochastic variations around the slow Schwabe cycle. Rather than estimating the SCL from the few data points around the solar minima, Benestad (2005) proposed to use a Fourier truncation to fit the sunspot record and hence use the entire data sample to estimate the SCL.

In particular, SSH2011's estimate of the SCL for cycle 23 (12.2 yr) was substantially longer than the estimate of 10.5 yr reported by the Danish Meteorological Institute (based on Friis-Christensen and Lassen (1991) and follow-up studies) and 10.8 yr estimated by Benestad (2005) (Table 1). Such a long cycle is the basis for their projected cooling (a decrease from −11.2 to −17.2 °C with a 95 % confidence interval of −20.5 to −14 °C) at Svalbard over solar cycle 24 (starting 2008). The observed mean over 2008–2011 suggests a continued warming that reached −9.17 °C as an average for the 4 yr, which means that the mean winter temperature of 2012–2018 (the next 7 winter seasons) must be −21.8 °C for a good prediction. An analysis of 7-season running mean values of the Svalbard temperature reveals that it is rarely below −15 °C and has never been as low as −21 °C since the measurements began.

SSH2011 used a weighted regression to account for errors of the mean temperature estimates over the periods corresponding to solar cycles. Hence they accounted for errors in the mean estimate, but neglected the errors associated with the SCL, which are more substantial than the errors in the mean seasonal or annual temperature over 10 yr segments. They also applied a bootstrapping approach to estimate the errors in the correlation coefficients (between −0.52 and −0.97), as they argued that there is no analytical expression to do so. When we computed the correlation (using "R"s "cor.test") between SCL and the winter temperature listed in their Table 1, we obtained

a correlation for the winter of 0.37 with 95 % confidence interval between −0.39 and 0.83, and when using the previous cycle SCL, we got −0.84 with a confidence interval between −0.39 and −0.96 (as opposed to −0.52 and −0.97 reported by SSH2011). The "cor.test" test statistic is based on Pearson's product moment correlation coefficient "cor(x, y)" and follows a t-distribution with "length(x)-2" degrees of freedom, and an asymptotic confidence interval is given based on Fisher's Z transform. Hence, the claim made by SSH2011 that there is no analytical expression for estimating confidence intervals for correlation is false.

Their estimate of the errors in the correlation involved 1000 picks of random paired sub-samples from the SCL and temperatures, where the same pair sometimes were picked more than once. A more appropriate strategy would be to carry out a set of Monte-Carlo simulations accounting for the errors due to the SCL (η_S) and mean temperature estimates (η_T). Here the symbol (subscripts) "S" refers to SCL and "T" refers to the local winter mean temperature. We estimated the error in SCL from the standard deviation of the difference between the SCL estimates from SSH2011 and Benestad (2005): $\eta_S = \sigma_S$, where σ_S is the standard deviation of the SCL difference: $S_{HSS2011}$–S_{B2005}. Then we re-calculated the 95 % confidence interval of the correlation estimates by adding white noise to temperature and SCL with standard deviations of σ_S for SCL, and for temperature we took the error of the mean estimate to be $\sigma_T/n^{1/2}$ where n is 10 for each 10 yr long segment. The Monte-Carlo simulation of the correlations between temperature and SCL were then estimated as: cor($T + \eta_T$, $S + \eta_S$), and was repeated 30 000 times with different random realisations of the error terms η_T and η_S (Table 1). The Monte-Carlo simulations gave a 95 % confidence interval for the correlation between −0.85 to 0.08, substantially wider than both "cor.test" and SSH2011. However, the latter two did not account for the uncertainties in the SCL estimates, which amplify the real uncertainties. Due to substantial uncertainties in the SCL, the Monte-Carlo simulations that we propose represent the most appropriate approach assessing the confidence intervals.

The Monte-Carlo simulation also revealed that the SSH2011 correlation estimate was not centered in the simulated correlation error distribution, but was biased towards higher absolute values. The correlation estimate based on the Benestad (2005) SCL, on the other hand, gave a better match with the mean correlation from the Monte-Carlo simulation, although this too had a greater absolute value than the mean error estimate. Furthermore, the bootstrapping approach adopted by SSH2011 seemed to give a biased error distribution, and we did not get the same 95 % confident limits as they did (we made 30 000 iterations). From just 9 data points, we find it quite incredible that the magnitude of their lower confidence limit was higher than 0.5. These results therefore suggest that the choice made in SSH2011 of SCL was indeed "fortunate" within the bounds of error estimates by getting correlations in the high end of the spectrum.

Since SSH2011 made at least 10 different tests (zero and one SCL lag and for 4 seasons plus the annual mean), the true significance can only be estimated by a field significance test, e.g. the Walker test: $p_W = 1-(1-\alpha_{global})^{1/K}$ (Wilks, 2006). The reason is that from 100 random tests, about 5 % are expected to achieve scores that are at the 5 % significance level. Another question is how many other temperature series that have been examined, as the appropriate number of tests to use in the Walker test should include all (also any unreported) tests in order to avoid a biased selection or lucky draw. When we estimate the p-value of their correlation from the null-hypothesis derived from the Monte-Carlo simulations, we find that all the p-values exceed p_W, and hence their results are not statistically significant at the 5 %-level.

Solheim et al. (2012) expanded the correlation exercises between SCL and temperature to include several locations in the North Atlantic region. The fact that several of these give similar results can be explained from the spatial correlation associated with temperature anomalies on time scales greater than one month. Their analysis involved 6–11 degrees of freedom, depending on the length of the available record, but since they applied their analysis to both SCL with zero and one-period lag, in addition to a number of locations, they would need to account for the problem of multiplicity,

and apply e.g. the Walker test. The failure to do so will give misleading results. The replication of SSH2011 is implemented with the following command lines in R:

```
> library(replicationDemos)
> Solheim.et.al.2011()
```

The main problem with the analysis presented by SSH2011 was the lack of a convincing physical basis, inappropriate hypothesis testing, the inflation of significance, and a small data sample insufficient to support the conclusions.

A2.3 Case 5: presumed dependencies and no model evaluation

Scafetta and West (2007, 2006a, b, 2005) argued that the recent increases in the global mean temperature were influenced by solar activity rather than increased GHG concentrations.

The analysis, on which Scafetta and West based their conclusions, assumed that the global mean temperature was not influenced by factors other than solar variability on decadal to multi-decadal time scales. Furthermore, they dismissed the role of increased concentrations of GHG, based on the model fit to the solar trend, assuming that a solar influence excludes the effect from increased CO2-levels.

Scafetta and West assumed that all the climate variability over wide frequency bands spanning 11 and 22 yr were due to changes in the Sun. They developed a model which was not evaluated against independent data, and hence they had no information about its skill. Benestad and Schmidt (2009) demonstrated that the strategies employed in Scafetta and West (2005, 2006a, b, 2007) were unsuitable for analysing solar-terrestrial relationships, and the source code for replicating these studies is included in replicationDemos (Table 1). Scafetta and West's strategy failed to account for "spectral leakage", common trends, and the presence of a range of frequencies in chaotic signals. They applied a transfer function based on the ratio of the standard deviation for respective temperature and total solar irradiance (TSI) after having applied a broad band-pass filter (7.3–14.7 and 14.7–29.3 yr) to both. Moreover, their analysis a priori assumed that no other factor was affecting Earth's climate over these wide ranges of time-scales, and hence it is not surprising that they arrive at a misguided answer that seemed to suggest a strong solar influence.

The replication of the Scafetta and West papers, as done by Benestad and Schmidt (2009), is implemented with the following command lines in R:

```
> library(replicationDemos)
> Scafetta2006()
```

The Scafetta and West (2005, 2006a, b, 2007) papers demonstrate how potential misleading conclusions are drawn when the model has not been subject to careful evaluation. Furthermore, their conclusions hinged on a set of assumptions which were not justified.

A2.4 Case 6: misinterpretation of statistics

Douglass et al. (2007) claimed that upper air trends predicted by global climate models were inconsistent with the trends measured by radiosondes and satellites. This purported discrepancy has been echoed on various Internet sites, been promoted by the Norwegian organisation "*klimarealistene*", and included in the NIPCC report. The flaw in the analysis presented in this paper can easily be exposed through replication, and hence this is a perfect case in terms of agnotology.

The paper relied on an analysis which confused the confidence interval for the mean estimate with the spread of a statistical sample, leading to a conclusion which was inconsistent with the results presented in the paper itself.

Douglass et al.'s (2007) conclusions were based on an inappropriate analytical set-up, and the flaw in their paper was caused by a confusion about the interpretation of the error estimate $\pm 2\sigma_{SE}$ derived from the standard deviation σ of the different trend estimates associated with different global climate model simulations. Here, we use the

same notation as in the original paper, where the subscript "SE" denotes the error estimate, and $\sigma_{SE} = \sigma/(N\text{-}1)^{1/2}$, where N is the sample (model ensemble) size.

Their invalid definition was taken to be the confidence interval of the data sample, whereas the correct interpretation should the ~95 % confidence interval for the estimated mean value. The statistic that Douglass et al. (2007) wanted was an interval describing the range of the data sample representing the trends predicted by climate models, in order to test whether the observed trends was distinguishable to any of the model results. To do so, the observations must be compared with the range of model results. However, the range of data samples cannot possibly decrease with increased sample size N, which Douglass et al. (2007) implied when they used σ_{SE} to describe confidence interval. On the other hand, the estimated mean value from a data sample will become more accurate when estimated from a larger data sample, and the confidence interval for this mean estimate is proportional to σ_{SE}.

In other words, Douglass et al. (2007) used the confidence interval for the mean value rather than for the sample and their test constituted an evaluation of how many models were consistent with the mean of the ensemble. Hence, for some vertical levels their misconceived confidence limit excluded up to 59 % of the models from which it was derived (Table 1). Furthermore, if many more models with similar trend estimates had been added to the ensemble, the confidence interval for the mean value would diminish, however the spread would not necessarily be sensitive to the number of models, and a larger sample would not imply that almost all the models fall outside the model spread. The replication of the Douglass et al. (2007) is implemented with the following command lines in R:

```
> library(replicationDemos)
> Douglass2007()
```

The internal inconsistency and logical flaw of the analysis indicate that the conclusions of Douglass et al. (2007) could have been dismissed, even before Santer et al. (2008) highlighted additional flaws in the Douglass et al. (2007) paper. More recent analysis on the observations of the lower atmosphere also suggest even greater trends (Foster and Rahmstorf, 2011) and even better agreement between models and observations.

A2.5 Case 7: failure to account for the actual degrees of freedom

A paper by McKitrick and Michaels (2004a; MM2004) claimed that much of the historical temperature trends could be explained from local economic activity, level of literacy, and the heat island effect. The analysis was based on a regression analysis between local temperature trends and a set of economic co-variates.

MM2004 did not take into account the real degrees of freedom, as pointed out by Benestad (2004) who replicated their results. The economic co-variates would contain the same data within the border of each country, and temperature trends are smooth functions in space. The analysis neither involved a proper validation of the regression model against independent data.

A simple test by splitting the data according to latitude and using one part for calibration and the other for independent evaluation, demonstrated that the analysis of MM2004 was flawed (Table 1). Such a split-sample test had to make sure that there were no dependencies between the samples used for calibration and evaluation, and hence these samples would involve data from different regions.

A follow-up paper, McKitrick and Michaels (2007; MM07) involved similar flaws, revealed in Schmidt (2009) who concluded that the basis of their results was a set of correlations for a small selection of locations mainly from western Europe, Japan, and the USA. Schmidt found that these projected strongly onto naturally occurring patterns of climate variability and their spatial auto-correlation implied reduced real degrees of freedom. Corresponding correlations from GCMs were found to vary widely due to the chaotic weather component in any short-term record, and the results of MM2007 did not fall outside the simulated distribution. There was therefore no evidence of any "large-scale contamination" in the temperature record from economic activity. The replication of the MM04, as done by Benestad (2004), is implemented with the following command lines in R:


```
> library(replicationDemos)
> MM2004()
```

MM2004 drew their conclusion based on inappropriate statistics, not recognising that the temperature trends vary slowly over space, and their regression analysis misapplied weights to the different covariates resulting in poor predictions of independent data. Furthermore, recent studies indicates that the analysis for the ground temperatures is in accordance with the satellite-based analyses (Foster and Rahmstorf, 2011).

A2.6 Case 8: missing similarities

Veizer (2005) argued that galactic cosmic rays (GCR) are responsible for the most recent warming. This conclusion assumes that the GCR affects cloudiness and hence the planetary albedo, and provides a support for the purported dependencies by Svensmark (1998) and Courtillot et al. (2007). The GCR have been introduced to the general society through popular science books (e.g. Svensmark, 2007) and videos (e.g. "The Cloud Mystery"), and have represented an important feature of agnotology in northern Europe. The influence of GHG has often been dismissed on grounds of the speculated correlation between GCR and climate, assuming that the GCR-connection excludes the effect of changes in the GHG concentrations.

Veizer (2005) failed to present any resemblance between the GCR-proxies discussed in the paper and a proxy for temperature, and he provided no quantitative statistical analysis on the correspondence between these quantities. Furthermore, the purported dependency involved a neglect of the fact that many other factors may be more important in terms of generating cloud condensation nuclei.

The GCR are known to be modulated by solar activity through its influence on the inter-planetary magnetic field (IMF). In replicationDemos estimates based on Be-10 and temperature from the Vostoc ice cores can be shown together (Table 1), and any correlation between the two seems to be due to long-term trend over 40 000 yr rather than more "ephemeral" fluctuations on thousand year time scales. The correlation between Be-10 and the temperature proxy over the last 40 000 yr was −0.78 but this number reflected the long time scales (greater than 5000 yr). The high frequency component was estimated by subtracting a low-pass filtered record, using a Gaussian window with a width of 5000 yr. The correlation between the high-frequency components were only −0.23 with a 95 % confidence interval of −0.65 to +0.30. The replication of the Veizer (2005) is implemented with the following command lines in R:

```
> library(replicationDemos)
> paleaoproxy()
```

In other words, it is difficult to discern any credible evidence linking GCR and recent climate change, due to lacking correlation and the number of other factors present. Veizer (2005) did not exclude other possibilities, but assumed that the other factors would be weak if there were a strong connection between GCR and climate.

A2.7 Case 9: looking at wrong scales

Humlum et al. (2013) argued that changes in CO_2 follow changes in the temperature, and that this implies that the increases seen in the Keeling curve are not man-made. Their claims implicitly support the CO_2-curve presented by Beck (2008), and the meme that the increase in the CO_2 concentrations seen in the Keeling curve is not due to the burning of fossil fuels, has long been an aspect of agnotology surrounding the global warming issue. It is also acknowledged in Humlum et al. (2013) that their paper had received inputs from "*klimarealistene*" and people with documented connections to organisations such as *The Science & Environmental Policy Project* (SEPP[3]) and *The Heartland Institute*[4].

[3] http://web.archive.org/web/20070215190653/http://www.sepp.org/Archive/NewSEPP/ipccreview.htm

[4] http://www.viddler.com/v/6471cf6e

The analysis on which Humlum et al. (2013) based their conclusions removed the long-term signal through a correlation between the annual time differences in CO_2 and temperature, This procedure removes the long time scales, and emphasises the short-term variations. Hence, Humlum et al. (2013) found the well-known link between El Niño Southern Oscillation and CO_2. They then incorrectly assumed that this link excludes the effect of anthropogenic emissions.

Humlum et al. chose to analyse a short series from 1980 describing the global analysis of the CO_2 concentrations rather than the almost identical series from Mauna Loa going back to 1958. This, in addition to applying a differencing operator to the data effectively removed all trends and long time scales. It is easy to demonstrate that the method Humlum et al. used is unable to pick up the longer time scales, as shown in replicationDemos. In other words, the analysis emphasised the short time scales, and the analytical set-up was pre-disposed to ignore the anthropogenic component to the CO_2 concentrations. Another problem was that their study did not account for the carbon-budget, such as sources and sinks. It is not clear whether the increased CO_2 was assumed to originate from the ocean surface or the deep ocean, and their discussion ignored the vast literature concerning diffusion of trace gases in the oceans. They also neglected the work documented in the fourth assessment report of the IPCC (Solomon et al., 2007) regarding changes in the O_2/N_2 ratios, the acidification of the world oceans, and isotope ratios. The replication of the Humlum et al. (2013) is implemented with the following command lines in R:

```
> library(replicationDemos)
> Humlum.et.al.2012()
> diff12demo().
```


Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

A2.8 Case 10: circular reasoning

Cohn and Lins (2005) observed that tests for trends are sensitive to the expectations (the choice of the null-hypothesis), and argued that long-term persistence (LTP) makes standard hypothesis testing difficult. The implications of their conclusions was that the observed recent global warming is not extraordinary, but something which one should expect due to long-term persistence.

All processes involving a trend also exhibit some LTP, and the test by Cohn and Lins involved some degree of circular logic: forcings increase LTP and so an LTP derived from the data already contains the forcings and is not a measure of the intrinsic LTP of the system. Thus, in order to be physically consistent, arguing for the presence of LTP also implies an acknowledgement of past radiative forcing in the favour for an enhanced greenhouse effect. This point can be demonstrated by applying the function "testLTP" in replicationDemos to compare the auto-correlation function (ACF) estimated from a set of results produced by the global climate model ECHAM5 (Demuzere et al., 2009; Keenlyside and Latif, 2002) with constant boundary conditions and with historic GHG forcing respectively.

While it is true that statistical tests do depend on the underlying assumptions, it is not given that chosen statistical models such as auto-regressive moving-average (ARMA), auto-regressive integrated moving-average (ARIMA), or auto-regressive fractionally integrated moving-average (FARIMA) provide an adequate representation of the null-hypothesis. It is important to avoid interpreting part of the signal as "noise", as all these statistical models do represent a type of structure in time, be it as simple as a serial correlation, persistence, or more complex recurring patterns. Thus, the choice of model determines what kind of temporal pattern one expects to be present in the process analysed, and it is important to keep in mind that these models are not necessarily representative of nature. The statistical LTP models employed by Cohn and Lins were just convenient models which to some degree mimic the empirical data (tuned for several parameters), and are arguably far inferior compared to the physics-based

general circulation models (GCMs) for providing appropriate null-distributions (long control simulations). No GCM reproduces the observed global warming unless an enhanced greenhouse effect is taken into account (Solomon et al., 2007), and there is a well-known physical reasoning for why it has to be so (Weart, 2004).

Another difficulty with the notion that the global mean temperature varies randomly with substantial long-term departures from its mean, is that it then would imply a more unstable system with similar warming as we now observe throughout our history. However, the indications are that the historical climate has been fairly stable during the Holocene (Solomon et al., 2007). Cohn and Lins ignored all physical considerations in their analysis, and a serious problem with the idea that departures (such as the recent global warming) is random and natural, is that such changes in the global surface temperature would have physical implications in terms of energy conservation and the climate sensitivity.

Similarly, ARFIMA-type models, auto-correlation functions (ACF) and phase scrambling (Franzke, 2012) are sensitive to embedded long-term trends which may not be part of the noise. Hence, such models are not suitable for testing trend hypotheses when it is not known a priori what fraction is part of long-memory noise and what is really the signal. The function "testLTP" in replicationDemos demonstrates how the ACF differs when applied to model simulations of temperature, with a constant forcing and with historic GHG forcings respectively.

The difficulty with the analysis presented by Cohn and Lins was distinguishing between noise and signal, and treating both as noise resulted in misguided conclusions. Another problem with the idea that the climate is highly sensitive to variations in its own state, is that this implies a high climate sensitivity. We know that there is a forcing present associated with increases in CO_2 concentrations, and Cohn and Lins could not show that the climate sensitivity discriminate against some types of forcings and not others. The replication of the Cohn and Lins (2005) is implemented with the following command lines in R:

> library(replicationDemos)

> testLTP().

A3 Additional examples not in replicationDemos

There are some good examples which have not been replicated in replicationDemos, but nevertheless are useful in terms of agnotology. It is also important to include these to show that the small selection provided in replicationDemos does not represent a few isolated cases, but is part of a larger pattern and pose a challenge to the scientific community.

A3.1 Case 11: lack of plausible physics

Scafetta (2010) assumed that changes in the earth's rotation rate, which he somehow associated with climate variability, is entirely due to planetary forcing, neglecting other factors such as changes in the circulation in the earth's interior, which may be more important (Appell, 2012). There is no known mechanism explaining how the climate responds to minute changes in the planet's rotation rate, and Scafetta offered no estimates for the Coriolis force or sensitivity tests with different values for the Coriolis coefficient.

A3.2 Case 12: incorrect interpretation of mathematics

McIntyre and McKitrick (2005; MM2005) claimed that the reconstruction carried out by Mann et al. (1999, 1998) resulted from inappropriate data processing before a principal component analysis (PCA). They attributed the shape of the curve describing the reconstruction ("hockey stick shape") to the leading principal component (PC), and argued that since it had a "hockey stick shape" the results were likely an artifact. They argued that red noise processes tend to produce such shapes if the data were not 'centred' before computing anomalies.

MM2005 neglected the regression involved in the process of reconstructing the past temperatures. The important question is of how many PCs were included in the regression model and how much of the variance they could describe. The shape of each individual PC, on the other hand, is less relevant as the regression analysis weights the different PCs according to how well they match the calibration data. Another point is that the actual "blade" of the hockey stick graph were not a result of the PCA, but consisted of instrumental data which had been added to the reconstructions (Mann, 2012).

PCA is a common way of transforming a data matrix (X) into a new set of basis functions in data space, while keeping its information intact. The purpose is often to reorganise the data in a way that makes use of the redundancy in the data and makes subsequent analysis faster and less prone to incorrect weighting. The PCA can for instance be done through singular vector decomposition, where $X = \mathrm{UWV}^T$ (Press et al., 1989; Strang, 1988), where U and V contain sets of orthogonal vectors. Hence, the shape of the leading PC is not really relevant, but the question that matters is how many components are included in the subsequent weighting of these components, and what information are embedded in these components.

The arguments presented in MM2005 were irrelevant for the question they wanted to address. Furthermore, the general features of the Mann et al. (1998, 1999) reconstruction were also found in other independent analyses (Solomon et al., 2007, and the work has been further evaluated by the Committee on Surface Temperature Reconstructions for the Last 2000 yr, National Research Council (2006). Furthermore, the MM2005-paper were criticised by Wahl and Ammann (2007), Huybers (2005), and Von Storch and Zorita (2005). These criticisms, however, did not convince McIntyre and McKitric, and further exchange followed in the literature (Mann et al., 2009; McIntyre, 2005a, b; McIntyre and McKitrick, 2009); there is also a long Wikipedia entry on this topic: "Hockey stick controversy" http://en.wikipedia.org/wiki/Hockey_stick_controversy.

The source code for the Mann et al. (1998) analysis has been available on-line since 2005[5], although there have been accusations of not sharing the data and the code.

A3.3 Case 13: contamination by other factors

Beck (2008) described a curve for atmospheric CO_2-concentrations which is at variance with corresponding results presented in Solomon et al. (2007). He compiled measurements from different locations at different times, often in Europe near CO_2-sources. The implication is that the upward trend in the current CO_2-measurements (Keeling curve) is not extraordinary.

Modern satellite-based measurements (NASA/AIRS) show that the concentrations in these regions may be substantially higher than the background level because of their proximity to the emission sources. Beck presented dramatic changes in CO_2-concentrations, which cannot be explained in terms of the carbon cycle (exchange between air, sea, and surface, involving photosynthesis and ocean acidification). Hence, the ignorance and neglect of relevant context makes such analyses prone to misguided interpretations, as in Humlum et al. (2013).

The analysis carried out by Beck (2008) did not reflect the global background levels, but the results were affected by the contamination from local sources and suffered from a lack of homogeneity. His results were not corroborated by independent studies of related aspects, such as the carbon cycle and carbon budgets.

A3.4 Case 14: incomplete account of the physics

Miskolzi (2010) attempted to calculate the significance of greenhouse effect through estimating how much of the upwelling long infra-red radiation (IR) is absorbed in the atmosphere. He purported that the atmosphere is saturated with respect to CO_2. The purports made in this paper have been promoted by organizations such as "Friends of Science", been propagated through the Internet, and contributed to the misguided

[5] http://www.meteo.psu.edu/holocene/public_html/shared/research/MANNETAL98/

idea that the increases in the CO_2 concentrations have little effect on the global mean temperature.

The Miskolzi (2010) paper was published in same journal as Beck (2008), Energy & Environment, and arrived at wrong conclusions due to neglecting relevant physics, such as convection, latent heat of evaporation and sensible heat. The paper is also difficult to follow, as the manuscript has the character of being an unfinished draft with undefined terms, and making few references to relevant previous work; 6 of the 19 citations were to his own work while only 11 references could be considered as scientific journals.

His calculations for the atmospheric absorption of upwelling IR neglected latent and sensible heat fluxes, e.g. associated with vertical motions due to adjustment by hydrostatic stability. This negligence alone invalidates his results, as the time scale associated with hydrostatic adjustment is shorter than the time scale of reaching local radiative equilibrium. Miskolzi's calculations also assumed that the amount of absorbed upwelling IR from the ground equals the downwelling IR from the atmosphere, and that the height, from which the bulk of the outgoing long-wave radiation (OLR) emissions occur, is insensitive to the atmosphere's optical depth. The former is hard to justify if the re-emission from the atmosphere is isotropic, as a volume of air is expected to emit equal amount of IR radiation upward and downward. The total amount of IR emitted by the air is expected to balance the amount of IR received from the ground if the atmosphere is in equilibrium, transparent to sunlight, and has no other source of energy. The latter claim would mean that an observer viewing IR from space would see down to the same height level even if the IR optical thickness increases, which logically does not make sense. Miskolzi (2010) also argued that atmospheric moisture has decreased, in contrast to independent observations (e.g. see the NOAA climate indicators[6]).

The conclusion drawn by Miskolzi is difficult to consolidate with the situation on Venus, which has a heavy atmosphere that mostly consists of CO_2 and has a potent greenhouse effect (Pierrehumbert, 2011). Miskolzi's analysis failed on multiple accounts and his conclusions are invalid.

[6] http://www.ncdc.noaa.gov/bams-state-of-the-climate/2009-time-series/humidity

A3.5 Case 15: differences in pre-processing of data

The papers Friis-Christensen and Lassen (1991; FL1991), Lassen and Friis-Christensen (1995; LF2000), Svensmark (1998; S1998), and Svensmark and Friis-Christensen (1997; SF1997) claimed that changes in the sun can explain a large part of the recent global warming. These papers have been used by Scafetta (see the earlier examples) and others as a support for their purports. Furthermore, they have contributed to the GCR meme, that has been popularised through the media.

The conclusions from these papers rest on a curve-fitting exercise and are based on little physics. The data handling has also been questioned (Laut, 2003), and recent up-to-date replication has suggested that the predictions diverge from the observations. These analyses are similar to the classical studies on the relationship between sunspots and climate performed over the centuries and that eventually have failed to stand up to new data (Benestad, 2002). Another point is that there is no trend in the solar proxies over the last 50 yr (Benestad, 2005; Lockwood and Frölich, 2008).

Damon and Laut (2004; DL2004) pointed out several flaws in the FL1991, LF2000, S1998, and SF1997, and argued that the apparent good match in FL1991 were obtained by "adding to a heavily smoothed ("filtered") curve, four additional points covering the period of global warming, which were only partially filtered or not filtered at all". Stauning (2011) took advantage of two additional solar cycles to recalculate the relationship between sunspot and temperature data. The trends in temperature and solar cycle length showed a strong divergence after 1976. Another question is whether filtering solar cycle lengths could be justified, as each epoch lasted approximately 11 yr, and hence implied that very slow changes in the sun would correlate directly with short term variations on Earth in a warped fashion. It is hard to conceive how the mean temperature in the period from 5 yr ago to the next 5 yr will be influenced by the solar activity from 25 yr in the past to 25 yr in the future. DL2004 also found trivial arithmetic errors in LF2000, being responsible for an incorrect curve.

For the analysis by S1998, DL2004 argued that the use of data from the US Defence Meteorological Satellite Program in SL1997 and S1998 was inappropriate as they did not represent total global cloud cover. More appropriate data from the International Satellite Cloud Climatology Program (ISCCP) were inconsistent with their hypothesis (Laken et al., 2012), and DL2004 observed that the more recent and conflicting part of the ISCCP data were shown in the SF1997 article but were omitted from the S1998. Independent investigation of the solar cycle lengths is in line with DL2004 (Benestad, 2005), and open-source replication method is available from an R-package called "cyclones":

```
> install.packages("cyclones")
> library(cyclones)
> Benestad2005()
```

Although a connection between solar activity and Earth's climate is plausible, there is no trend in the recent solar indices that can explain the current global warming. The way that these papers have handled and selected the data have been questioned, but due to lack of openness and transparency, it has been difficult to pinpoint the exact reason for the differences. These papers have also had an influence on Svensmark (2007), Shaviv (2002), Courtillot et al. (2007) and Veizer (2005).

A3.6 Case 16: selective use of data

The papers S1998, Svensmark (2007), Shaviv (2002), and Courtillot et al. (2007) argued that GCR affect earth's global cloud cover which subsequently modulates the planetary albedo. They also assumed that a strong connection between GCR and climate implies a weak role for GHG such as CO_2. Their ideas have pervaded the public minds through books (Svensmark, 2007), and videos ("The Cloud Mystery", "The Great Climate Swindle").

The analysis was based on data from various sources, stitched together with no testing of their quality, ignoring part of the data, and no proper verification. They also neglected the role that other processes may play in generating cloud condensation nuclei (CCN). The cloud data that Svensmark used were not universally accepted, as he had made changes to the data to correct for purported errors.

Svensmark (2007) made selective references without answering the serious criticism forwarded by Damon and Laut (2004) and Laut (2003). The original analysis presented by Svensmark was based on total cloud cover, which later turned out to provide a poor fit, and he then replaced these with data describing low-level cloudiness. He then used a different version of the cloud data to others, claiming that the original data were incorrect due to calibration problems and that the recent global warming was caused by GCR (Laken et al., 2012).

On a similar note, Courtillot et al. (2007) presented an analysis between solar irradiance and geomagnetic field, but ignored part of the data record for which the data diverged (Bard and Delaygue, 2008). Other errors included a confusion between the interpretation of solar irradiance changes and net forcing. They also argued that periods with high GCR-flux, found in cosmogenic isotope records, coincided with periods with high ice raft debris in the North Atlantic and assumed that high iceberg drift activity implies cold global conditions, which has not yet been established. Icebergs tend to originate from calving of ice sheets and glaciers, and should not be confused with sea-ice. As opposed to sea-ice, calving activity may not necessarily increase with lower temperatures.

Shaviv (2002) considered extreme time-scales of millions of years. He argued that our solar system takes about 250 million years to circle the Milky Way galaxy and that our solar system crosses one of the spiral arms about every ~ 150 million years. This number was arrived at by measuring the rotational velocity of stars in the Milky Way disk or other spiral galaxies. The pattern speed of the spiral arm in the Milky Way has not been firmly established, and a number of values are listed in Table 3 of Shaviv (2002) for the pattern speed of the spiral arms, taken from other publications ranging from

1969 to 2001. However, he selectively disregarded most of these results and derived "period for spiral arm crossing" of $p = 134 \pm 25$ Myr for four spiral arms in the upper extreme of the published range. Nevertheless, such astronomical considerations are a far shot from present state-of-the-art measurements and understanding of cloud physics here on earth. The distant past of the solar system and our galaxy is known to a far lesser extent than modern climate science.

The connection between these papers and agnotology is the implication that the GCR-mechanism explains the recent global warming. However, it can easily be shown that neither the GCR nor other solar indices exhibit any long-term trend over the last 50 yr which can explain the global warming (Benestad, 2005), and open-source code to show this is available from the "cyclones" R-package:

```
> library(cyclones)
> Benestad2005()
```

These papers neglected cloud condensation nuclei (CCN) from other sources, and their implications concerning GCR and the recent global warming cannot explain why the warming has been greatest during night (Solomon et al., 2007): the albedo mechanisms would be more important for the day side of the earth. It has also been established that there is no significant trend in GCR and other solar activity proxies in the last ~ 50 yr (e.g. Richardson et al., 2002; Benestad, 2002, 2005) and that in the most recent decades, there even has been a small trend in opposite direction to what is expected for solar forcing to cause a warming (Lockwood and Frölich, 2008). A review of the study of GCR-climate links suggest that the findings suggesting a link have not been supported through subsequent investigation (Laken et al., 2012).

A3.7 17: misinterpretation of spectral methods

In addition to claims that the giant planets exert influence on earth's climate, there is a paper by Yndestad (2006) that claims to identify a lunar "nodal" cycle (18.6 yr) in a selection of Arctic measurements and that the Arctic is a forced oscillating system controlled by the pull of gravity from the moon. This study too is based on harmonic analysis, as those described in cases 1–3, and hence the same type of criticism applies. Sea-ice and the local Arctic climate are strongly affected by winds and ocean currents, in addition to being closely coupled (Benestad et al., 2002). Furthermore, the Arctic climate involves dynamics with a pronounced non-linear chaotic character (Benestad et al., 2010), and the tides tend to propagate as coastal Kelvin waves rather than in the ocean interior. Hence, it would be problematic from a pure statistical analysis of a few measurements alone to attribute celestial causes to Arctic secular variations.

A4 Link between the cases and agnotology

Common to all these purported celestial influences is the lack of clear physical reasoning, which also applies to the papers claiming to report a link through SCL. For instance, it is hard to explain why SCL should affect the climate, although the notion is based on the idea that short SCL is associated with more intense solar activity. It is questionable whether this link has been established and it is a conundrum why there is a correlation between SCL that is stronger than corresponding correlations with the number of sunspots or the total solar irradiance.

The message from several of these papers has been picked up by a number of organisations, blogs, and been turned into videos. The Canadian organisation called "Friends of Science" embraced the work done by Miskolzi to argue that the atmospheric greenhouse effect is "saturated"[7]. Videos with the title "*The Global Warming Swindle*"[8] and "*The cloud mystery*"[9] have combined the ideas presented in the papers of Shaviv, Singer, Svensmark, and Friis-Christensen, and these have targeted the lay public who

[7] http://www.friendsofscience.org/assets/documents/The_Saturated_Greenhouse_Effect.htm

[8] http://www.youtube.com/watch?v=YtevF4B4RtQ

[9] http://www.youtube.com/watch?v=ANMTPF1blpQ

have been left with the impression that GHG are playing a minor role, CO_2 in the atmosphere is unproblematic, and that the recent warming has been caused by changes in the Sun.

The problem has been exacerbated by the lack of appropriate fora for debating these issues at a more profound depth. There are some exceptions as noted above, such as attempts in the science literature to correct a myth regarding an alleged recent "slowdown" in the global warming (Easterling and Wehner, 2009; Foster and Rahmstorf, 2011). Other exceptions include attempts to correct such myths by scientific communities through blogs such as SkepticalScience.com and RealClimate.org (Rapley, 2012). However, the criticism is not necessarily heeded, as MM2007 repeated the claim of made in MM2004 without acknowledging the criticism presented in Benestad (2004). While McKitrick and Michaels (2004b) responded to this, they defended their original positions by dismissing the criticism stating they were "*unaware of any paper in the refereed applied climatology literature that has performed the test [splitting the sample, using one for model calibration and the other for validation] suggested by Dr. Benestad... if he has ever seen such a test applied anywhere in a published atmospheric science paper he should have provided an example, which he did not*". Split sample tests are often the norm for testing statistical models, however (Benestad et al., 2008, 2007; Wilks, 1995).

Another problem was the lack of openness and transparency, which prevented finding out why the conclusions in some of these cases differed to attempts to replicate (Le Page, 2009). However, in the absence of real replication of the analysis on which these rest and recognised journals in which to publish and share the replicated results, the discourse has been superficial and dogmatic.

The current situation for climate sciences can be described as a struggle about the truth about the state of climate ("Climate of fear," 2010), and the discourse around the question of climate change has influenced the general society. Somerville and Hassol (2011) made a call for a badly needed voice of rational scientists in the modern society, and Theissen (2011) argued that many US undergraduate students are confused by a number of myths concerning climate change, propagated by blogs and media. Rahmstorf (2012) argued that the well-being of our societies depends on the wider public being well informed about the state of scientific knowledge and discourse. Books such as Oreskes and Conway (2010), Gelbspan (1998), Hoggan and Littlemore (2009), and Mooney (2005) also describe how climate science has been introduced to the society in a distorted way. Many of these myths can be traced to misleading papers or arise from implied false dichotomy that different causes exclude each other.

The objective of science should be to search for true answers as far as possible, and the question for a scientist should not be what the answer is, but how credible it is. In other words, how the results were derived and whether the approach is robust and sound. The virtue of science and science papers should also be to get to the bottom of disputed issue whenever possible, and resolve the differences in order to arrive at the most credible answer. The concept of agnotology brings all these cases together by providing replication exercises that demonstrate the shortcomings of these analyses. The selection made here included papers for which the demonstrations could be done in a simple manner, however, there are more papers which too are unpersuasive.

Acknowledgements. Part of the material presented here has been written up based on posts written for www.realclimate.org. This initiative was also carried out in support for the COST-action ES 1005 TOSCA in the sense of underlining the importance of scientific virtues such as transparency, evaluation of methods, replication of results, testing hypotheses, and assessment of the objectivity of the analytical design. We acknowledge valuable comments from Kristoffer Rypdal.

References

Anderegg, W. R. L., Prall, J. W., Harold, J., and Schneider, S. H.: Expert credibility in climate change, Proc. Natl. Acad. Sci., 107, 12107–12110, 2010.

Appell, D.: A pressing matter, Physics World, 37–41, 2012.

Bard, E. and Delaygue, G.: Comment on “Are there connections between the Earth’s magnetic field and climate?” by V. Courillot, Y. Gallet, J. L. Le Mouël, F. Fluteau, A. Genevey, EPSL 253, 382, 2007. EPSL 265, 2008.

Beck, E.-G.: 50 Years of continuous measurement of CO_2 on mauna loa, Energ. Environ., 19, 1017–1028, doi:10.1260/095830508786238288, 2008.

Bedford, D.: Agnotology as a Teaching Tool: Learning Climate Science by Studying Misinformation, J. Geogr., 109, 159–165, 2010.

Benestad, R. E.: Solar Activity and Earth’s Climate, Praxis-Springer, Berlin and Heidelberg, 2002.

Benestad, R. E.: Are temperature trends affected by economic activity? Comment on McKitrick & Michaels, Clim. Res., 27, 171–173, 2004.

Benestad, R. E.: A review of the solar cycle length estimates, Geophys. Res. Lett. 32, L15714, doi:10.1029/2005GL023621, 2005.

Benestad, R. E. and Schmidt, G. A.: Solar Trends and Global Warming, J. Geophys. Res., 114, D14101, doi:10.1029/2008JD011639, 2009.

Benestad, R. E., Hanssen-Bauer, I., and Førland, E. J.: Empirically downscaled temperature scenarios for Svalbard, Atmos. Sci. Lett., 3, 71–93, 2002.

Benestad, R. E., Hanssen-Bauer, I., and Førland, E. J.: An Evaluation of Statistical Models for Downscaling Precipitation and Their Ability to Capture Long-Term Trends, Int. J. Climatol., 27, 649–665, 2007.

Benestad, R. E., Chen, D., and Hanssen-Bauer, I.: Empirical-Statistical Downscaling. World Scientific Publishing, Singapore, 2008.

Benestad, R. E., Senan, R., Balmaseda, M., Ferranti, L., Orsolini, Y., and Melsom, A.: Sensitivity of summer 2-m temperature to sea ice conditions, Tellus A, 63, p. 324, doi:10.1111/j.1600-0870.2010.00488.x, 2010.

Bye, E., Humlum, O., and Stordahl, K.: Fortsatt dominans av naturlige klimaendringer, Teknisk Ukeblad, Climate of fear, 2010, Nature, 464, 141–141, 2011.

Cohn, T. A. and Lins, H. F.: Nature’s style: Naturally trendy, Geophys. Res. Lett., 32, L23402, 10.1029/2005GL024476, 2005.

Committee on Surface Temperature Reconstructions for the Last 2,000 Years, National Research Council: Surface Temperature Reconstructions for the Last 2,000 Years, The National Academies Press, Washington, D.C., 2006.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

Cook, J., Nuccitelli, D., Green, S. A., Richardson, M., Winkler, B., Painting, R., Way, R., Jacobs, P., and Skuce, A.: Quantifying the consensus on anthropogenic global warming in the scientific literature, submitted, Environ. Res. Lett., accepted, 2013.
Courtillot, V., Gallet, Y., Le Mouël, J.-L., Fluteau, F., and Genevey, A.: Are there connections between the Earth's magnetic field and climate? Earth Planet. Sci. Lett., 253, 328–339, 2007.
Damon, P. E. and Laut, P.: Pattern of Strange Errors Plagues Solar Activity and Terrestrial Climate Data, Eos, 85, 370–374, 2004.
Demuzere, M., Werner, M., van Lipzig, N. P. M., and Roeckner, E.: An analysis of present and future ECHAM5 pressure fields using a classification of circulation patterns, Int. J. Climatol., 29, 1796–1810, 2009.
Doran, P. T. and Zimmerman, M. K. Examining the Scientific Consensus on Climate Change. Eos, Transactions American Geophysical Union, 90, 22–23, doi:10.1029/2009EO030002, 2009.
Douglass, D. H., Christy, J. R., Pearsona, B. D., and Singer, S. F.: A comparison of tropical temperature trends with model predictions. Int. J. Climatol., 28, 1693–1701, 2007.
Easterling, D. R. and Wehner, M. F.: Is the climate warming or cooling?, Geophys. Res. Lett., 36, L08706, doi:10.1029/2009GL037810, 2009.
Foster, G. and Rahmstorf, S.: Global temperature evolution 1979–2010, Environ. Res. Lett., 6, 044022, doi:10.1088/1748-9326/6/4/044022, 2011.
Franzke, C.: On the statistical significance of surface air temperature trends in the Eurasian Arctic region, Geophys. Res. Lett., 39, L23705, doi:10.1029/2012GL054244, 2012.
Friis-Christensen, E. and Lassen, K.: Length of the solar cycle: an indicator of solar activity closely associated with climate, Science, 254, 698–700, 1991.
Gelbspan, R.: The Heat Is On: The Climate Crisis, The Cover-up, The Prescription, Upd Sub. Ed. Basic Books, 1998.
Hoggan, J. and Littlemore, R.: Climate Cover-Up: The Crusade to Deny Global Warming, Greystone Books, 2009.
Humlum, O., Solheim, J.-E., and Stordahl, K.: Identifying natural contributions to late Holocene climate change, Global Planet. Change, 79, 145–156, 2011a.
Humlum, O., Solheim, J.-E., and Stordahl, K.: Spectral Analysis of the Svalbard Temperature Record 1912–2010, Adv. Meteorol., 2011, 543146, doi:10.1155/2011/543146, 2011b.
Humlum, O., Stordahl, K., and Solheim, J.-E.: The phase relation between atmospheric carbon dioxide and global temperature, Global Planet. Change, 100, 51–69, 2013.

Huybers, P.: Comment on "Hockey sticks, principal components, and spurious significance" by S. McIntyre and R. McKitrick, Geophys. Res. Lett., 32, L20705, doi:10.1029/2005GL023395, 2005.
Idso, C. and Singer, F.: Climate Change Reconsidered: 2009 Report of the Nongovernmental Panel on Climate Change (NIPCC). The Heartland Institute, 2009.
Keenlyside, N. and Latif, M.: Enso Hindcast Performance of The Mpi-om/echam5 Coupled Model, Nice, 2002.
Kinne, O.: Climate Research: an article unleashed worldwide storms, Clim. Res., 24, 197–198, 2003.
Laken, B. A., Pallé, E., Čalogović, J., and Dunne, E. M.: A cosmic ray-climate link and cloud observations, J. Space Weather Space Clim., 2, A18, doi:doi:1310.1051/swsc/2012018, 2012.
Lassen, K. and Friis-Christensen, E.: Variability of the solar cycle length during the past five centuries and the apparent association with terrestrial climate, J. Atmos. Terr. Phys., 57, 835–845, 1995.
Laut, P.: Solar activity and terrestrial climate: an analysis of some purported correlations, J. Atmos. Sol.-Terr. Phys., 65, 801–812, 2003.
Lawrimore, J., Rennie, J., and Thorne, P.: Responding to the Need for Better Global Temperature Data, Eos, Transactions American Geophysical Union, 94, 61–62, 2013.
Le Page, M.: Sceptical climate researcher won't divulge key program, New Scientist, 2009.
Lockwood, M. and Frölich, C.: Recent oppositely directed trends in solar climate forcings and the global mean surface air temperature. II. Different reconstructions of the total solar irradiance variation and dependence on response time scale, Proc. R. Soc. A, 464, 1367–1385, 2008.
Loehle, C. and Scafetta, N.: Climate Change Attribution Using Empirical Decomposition of Climatic Data, The Open Atmospheric Science Journal, 5, 74–86, 2011.
Mann, M. E.: The hockey stick and the climate wars: dispatches from the front lines. Columbia University Press, New York, 2012.
Mann, M. E., Bradley, R. S., and Hughes, M. K.: Global-scale temperature patterns and climate forcing over the past six centuries, Nature, 392, 779–787, 1998.
Mann, M. E., Bradley, R. S., and Hughes, M. K.: nothern Hemisphere Temperatures During the Past Millennium: Inferences, Uncertainties, and Limitations, Geophys. Res. Lett., 26, 759–762, 1999.

Mann, M. E., Bradley, R. S., and Hughes, M. K.: Reply to McIntyre and McKitrick: Proxy-based temperature reconstructions are robust, Proc. Natl. Acad. Sci., 106, E11–E11, 2009.
McIntyre, S.: Reply to comment by von Storch and Zorita on “Hockey sticks, principal components, and spurious significance”, Geophys. Res. Lett., 32, L20714, doi:10.1029/2005GL023089, 2005a.
McIntyre, S.: Reply to comment by Huybers on “Hockey sticks, principal components, and spurious significance”, Geophys. Res. Lett., 32, L20713, doi:10.1029/2005GL023586, 2005b.
McIntyre, S. and McKitrick, R.: Hockey sticks, principal components, and spurious significance, Geophys. Res. Lett., 32, L03710, doi:10.1029/2004GL021750, 2005.
McIntyre, S. and McKitrick, R.: Proxy inconsistency and other problems in millennial paleoclimate reconstructions, Proc. Natl. Acad. Sci., 106, E10–E10, 2009.
McKitrick, R. and Michaels, P. J.: A test of corrections for extraneous signals in gridded surface temperature data, Clim. Res., 26, 159–173, 2004a.
McKitrick, R. and Michaels, P. J.: Are temperature trends affected by economic activity? Reply to Benestad (2004), Clim. Res., 27, 175–176, 2004b.
McKitrick, R. and Michaels, P. J.: Quantifying the influence of anthropogenic surface processes and inhomogeneities on gridded global climate data, J. Geophys. Res., 112, D24S09, doi:10.1029/2007JD008465, 2007.
Meehl, G. A., Covey, C., Delworth, T., Latif, M., McAvaney, B., Mitchell, J. F. B., Stouffer, R. J., and Taylor, K. E.: The WCRP CMIP3 Multimodel Dataset: A New Era in Climate Change Research, B. Am. Meteorol. Soc., 88, 1383–1394, 2007.
Miskolzi, F.: The stable stationary value of the earth’s global average atmospheric planck-weighted greenhouse-gas optical thickness, Energ. Environ., 21, 243–262, doi:10.1260/0958-305X.21.4.243, 2010.
Mooney, C. C.: The Republican War on Science, Basic Books, 2005.
National Research Council (US): Climate change science: an analysis of some key questions, The compass series, National Academy Press, Washington, D.C., 2001.
Newt, M. and Wiik, H.: Når akademikere villeder om klimaendringer, UTDANNING, 15, 2012.
Oreskes, N.: The scientific consensus on climate change, Science, 306, 5702, doi:10.1126/science.1103618, 2004.
Oreskes, N. and Conway, E. M.: Merchants of Doubt: How a Handful of Scientists Obscured the Truth on Issues from Tobacco Smoke to Global Warming. Bloomsbury Publishing, 2010.
Pain, H. J.: The Physics of Vibrations and Waves, 3rd Edn., John Wiley & Sons Ltd. 1983.

Pearce, F.: The climate files: the battle for the truth about global warming, Guardian Books, London, 2010
Pebesma, E., Nüst, D., and Bivand, R.: The R Software Environment in Reproducible Geoscientific Research, Eos, 93, 163–164, 2012.
Pierrehumbert, R. T.: Infrared radiation and planetary temperature, Phys. Today, 64, 33–38, 2011.
Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T.: Numerical Recipes in Pascal. Cambridge University Press, Cambridge, UK, 1989.
Proctor, R. and Schiebinger, L. (Eds.): Agnotology: The Making and Unmaking of Ignorance, Stanford University Press, 2008.
Rahmstorf, S.: Is journalism failing on climate?, Environ. Res. Lett., 7, 041003, doi:10.1088/1748-9326/7/4/041003, 2012.
Rapley, C.: Climate science: Time to raft up, Nature, 488, 583–585, 2012.
R Development Core Team, 2004. R: A language and environment for statistical computing, Vienna, Austria.
Richardson, I. G., Cliver, E. W., and Cane, H. V.: Long-term trends in interplanetary magnetic field strength and solar wind structure during the twentieth century, J. Geophys. Res., 107, 1304, doi:10.1029/2001JA000507, 2002.
Santer, B. D., Thorne, P. W., Haimberger, L., Taylor, K. E., Wigley, T. M. L., Lanzante, J. R., Solomon, S., Free, M., Gleckler, P. J., Jones, P. D., Karl, T. R., Klein, S. A., Mears, C., Nychka, D., Schmidt, G. A., Sherwood, S. C., and Wentz, F. J.: Consistency of modelled and observed temperature trends in the tropical troposphere, Int. J. Climatol., 28, 1703–1722, 2008.
Scafetta, N.: Empirical evidence for a celestial origin of the climate oscillations and its implications, J. Atmos. Sol.-Terr. Phys., 72, 951–970, 2010.
Scafetta, N.: Testing an astronomically based decadal-scale empirical harmonic climate model versus the general circulation climate models, J. Atmos. Sol.-Terr. Phys., 80, 124–137, 2012a.
Scafetta, N.: A shared frequency set between the historical mid-latitude aurora records and the global surface temperature, J. Atmos. Sol.-Terr. Phys., 74, 145–163, 2012b.
Scafetta, N.: Multi-scale harmonic model for solar and climate cyclical variation throughout the Holocene based on Jupiter–Saturn tidal frequencies plus the 11-year solar dynamo cycle, J. Atmos. Sol.-Terr. Phys., 80, 296–311, 2012c.

Scafetta, N. and West, B. J.: Estimated solar contribution to the global surface warming using the ACRIM TSI satellite composite, Geophys. Res. Lett. 32, L18713, doi:10.1029/2005GL023849, 2005.
Scafetta, N. and West, B. J.: Phenomenological solar signature in 400 years of reconstructed Northern Hemisphere temperature record, Geophys. Res. Lett., 33, L17718, doi:10.1029/2006GL027142, 2006a.
Scafetta, N. and West, B. J.: Phenomenological solar contribution to the 1900–2000 global surface warming, Geophys. Res. Lett., 33, L05708, doi:10.1029/2005GL025539, 2006b.
Scafetta, N. and West, B. J.: Phenomenological reconstructions of the solar signature in the Northern Hemisphere surface temperature records since 1600, J. Geophys. Res., 112, D24S03, doi:10.1029/2007JD008437, 2007.
Scafetta, N. and West, B. J.: Is climate sensitive to solar variability?, Phys. Today, 61, 50–51, 2008.
Schmidt, G. A.: Spurious correlations between recent warming and indices of local economic activity, Int. J. Climatol., 29, 2041–2048, doi:10.1002/joc.1831, 2009.
Shaviv, N. J.: Cosmic Ray Diffusion from the Galactic Spiral Arms, Iron Meteorites, and a Possible Climatic Connection, Phys. Rev. Lett., 89, 051102, doi:10.1103/PhysRevLett.89.051102, 2002.
Sherwood, S.: Science controversies past and present, Phys. Today, 64, 39–44, 2011.
Solheim, J.-E., Stordahl, K., and Humlum, O.: Solar Activity and Svalbard Temperatures, Adv. Meteorol., 2011, 1–8, 2011.
Solheim, J.-E., Stordahl, K., and Humlum, O.: The long sunspot cycle 23 predicts a significant temperature decrease in cycle 24, J. Atmos. Sol.-Terr. Phys., 80, 267–284, 2012.
Solomon, S., Quin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignotr, M., and Miller, H. L. (Eds.): 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, United Kingdom and New York, NY, USA, 2007.
Somerville, R. C. J. and Hassol, S. J.: Communicating the science of climate change, Phys. Today, 64, 48–53, 2011.
Soon, W. and Baliunas, S.: Proxy climatic and environmental changes of the past 1000 years, Clim. Res., 23, 89–110, 2003.

Spencer, R. W. and Braswell, W. D.: On the diagnosis of radiative feedback in the presence of unknown radiative forcing, J. Geophys. Res., 115, D16109, doi:10.1029/2009JD013371, 2010..
Stauning, P.: Solar activity–climate relations: A different approach, J. Atmos. Sol.-Terr. Phys., 73, 1999–2012, 2011.
Stephenson, G.: Mathematical methods for science students, 2nd Edn., Longman Scientific & Technical, London, 1973.
Strang, G.: Linear Algebra and its Application, Harcourt Brace & Company, San Diego, California, USA, 1988.
Svensmark, H.: Influence of Cosmic Rays on Earth's Climate, Phys. Rev. Lett., 81, 5027–5030, 1998.
Svensmark, H.: The chilling stars: a new theory of climate change, Icon, Thriplow, Cambridge, 2007.
Svensmark, H. and Friis-Christensen, E.: Variations of cosmic ray flux and global cloud coverage – a missing link in solar-climate relationships, J. Atmos. Sol.-Terr. Phys. 59, 1225, 1997..
Theissen, K. M.: What do U.S. Students Know About Climate Change?, Eos, 92, 477–478, 2011.
Veizer, J.: Celestial Climate Driver: A Perspective from Four Billion Years of the Carbon Cycle, Geoscience Canada, North America, 32, March 2005, available at: http://journals.hil.unb.ca/index.php/GC/article/view/2691/3113 (last access: 26 April 2013), 2005.
Von Storch, H. and Zorita, E.: Comment on "Hockey sticks, principal components, and spurious significance" by S. McIntyre and R. McKitrick Geophys. Res. Lett., 32, L20701, doi:10.1029/2005GL022753, 2005.
Wagner, W.: 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, 1603–1613, Remote Sens., 3, 2002–2004, 2011.
Wahl, E. R. and Ammann, C. M.: Robustness of the Mann, Bradley, Hughes reconstruction of Northern Hemisphere surface temperatures: Examination of criticisms based on the nature and processing of proxy climate evidence, Clim. Change, 85, 33–69, 2007.
Waldmeier, M.: The sunspot-activity in the years 1610–1960, Zurich Schulthess & Company AG, 1961.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

Weart, S. R.: The Discovery of Global Warming, Harvard University Press, 2004.
Wilks, D. S.: Statistical Methods in the Atmospheric Sciences, Academic Press, Orlando, Florida, USA, 1995.
Wilks, D. S.: On “Field Significance” and the False Discovery Rate, J. Appl. Meteorol. Climatol., 45, 1181–1189, 2006.
Williams, R. W.: Fundamental Formulas of Physics, in: Dover books, edited by: Menzel, D. H., 544–562, 1960.
Yndestad, H.: The influence of the lunar nodal cycle on Arctic climate, ICES J. Mar. Sci., 63, 401–420, 2006.

Table 1. An overview of functions in the “replicateDemos” package, a short description of their demonstrations/replication and the reference to the paper for which they are relevant.

R function	Description	Reference
Humlum.et.al.2011()	Extends the curve-fit beyond the fitted interval for Greenland temperature.	Humlum et al. (2011a)
Humlum.et.al.2012()	Repeats lagged correlation between differentiated series – where trends are effectively removed.	Humlum et al. (2013)
Solheim.et.al.2011()	Replicates analysis based on data given in tables: SCL, temperature, correlations, and bootstrapping.	Solheim et al. (2011)
LoehleScafetta2011()	Some replication and some demonstrations showing how curves segments are made up of harmonics, and how a fit within an interval fails to describe the remaining part of the curve.	Loehle and Scafetta (2011)
Scafetta2010()	Some replication and some demonstrations showing how different temperatures change results, and how different noise processes may seem to contain long-term cycles.	Scafetta (2010)
ENSO.example()	Example showing how the El Niños during the 1980s resembled sinousoids, and that calibrating a curve-fit on one half will give a good match for the other half. But the fit fails outside this segment.	Scafetta (2012a)
resonance()	Simulation of how a damped oscillator forced with noisy signal produces oscillations with fixed frequencies.	Loehle and Scafetta (2011); Scafetta (2012a, 2010)
Scafetta2012()	Replicates Fig. 5b in Scafetta (2012a)	Scafetta (2012a)
Scafetta.tab1()	Replicates and visualises Table 1 in Scafetta (2012a).	Scafetta (2012a)
Scafetta2006()	The R-script used by Benestad and Schmidt (2008), modified to be part of the R-package.	Scafetta and West (2005, 2006a, b)
MM2004()	The R-script used to carry out the analysis of Benestad (2004)	McKitrick and Michaels (2004a)
Douglas2007()	Replication and evaluation of the GCM confidence interval based on data in tables.	Douglass et al. (2007)
paleaoproxy()	Comparison between cosmogenic Be-10 isotope proxies, CO_2, and temperature from the Vostoc ice core.	Veizer (2005)
DJF()	Examiine the winter temperature at Svalbard and the forecast made by SSH2011	Solheim et al. (2011)
testLTP()	Test the LTP assumption and the way long-term trends affect the auto-correlation.	(Cohn and Lins, 2005; Franzke, 2012)

Table 2. Overview of papers which have attempted to identify and attribute cycles in Earth's climate to external causes.

Author/yr	Journal	Periodicity (yr)	Comment
Humlum et al. (2011a)	Global Planet Change	16.8, 25.7, 68.4, 556, 1186, 2804	
Humlum et al. (2011b)	Adv. Meteorol.	71.7, 24.9, 15.3, (74.3, 24.5, 17.1)	
Scafetta and West (2006b)	Geophys. Res. Lett.	7.3–14.7, 14.7–29.3	35–60 % and 20–40 % of 1900–2000 and 1980–2000 warming respectively
Scafetta and West (2007)	J. Geophys. Res.	11	up to ~ 50 % of warming since 1950
Scafetta (2012a)	J. Astr. Terr. Phys	9.1, 10–10.5, 20–21,60–62	at least 60 % of warming since 1970
Loehle and Scafetta (2011)	Open Atmos. Sci. J.	20, 60	
Scafetta (2012c)	J. Astr. Terr. Phys	9.98, 10.9, 11.86, 61, 115, 130	
Yndestad (2006)	J. Mar. Sci.	6, 18, 74	

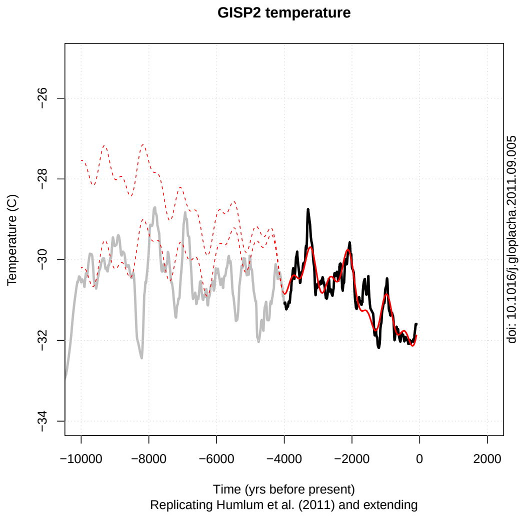

Fig. 1. A replication of Humlum et al. (2011a)'s model for the GISP2-record (solid red) and extensions back to the end of the last glacial period (red dashed). The two red dashed lines represent two attempts to extend the curve fit, one keeping the trend over the calibration interval and one setting the trend to zero. The black curve shows the part of the data showed in Humlum et al. (2011a) and the grey part shows the section of the data they discarded. The figure can be reproduced from "replicationDemos" (Table 1).

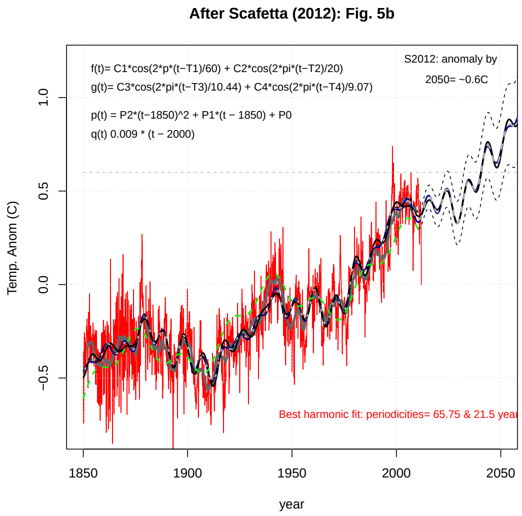

Fig. 2. A reproduction of Fig. 5 in Scafetta, 2012a (also available on-line from http://arxiv.org/pdf/1201.1301v1.pdf). The reproduction was done calling the function "Scafetta2012()" in replicationDemos. The grey horizontal dashed line marks the level where Scafetta's curve intersects year 2050. An interative search for periodicities in the vicinity of those suggested by Scafetta, gave a best fit to the pair of harmonics if they were 21.5 and 65.75 yr respectively.

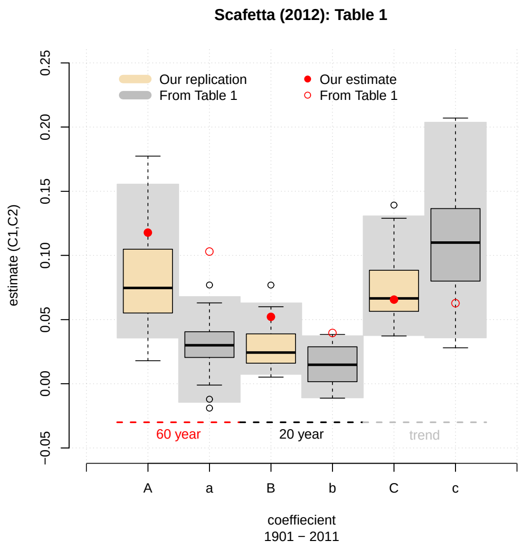

Fig. 3. This figure compares the amplitudes for the 60 **(a)** and 20 yr harmonics **(b)** and trends **(c)** from the GCMs (box) and the HadCRUT3v (red symbols). Capital letters on the x-axis refer to re-calculated regression coefficients whereas the lower-case letters refer to those in Scafetta (2012). The comparison shows that the amplitudes of the 20 and 60 yr variability found in the observational record (filled circles) are within the range simulated by the GCMs. The grey boxes are values copied from Table 1 in Scafetta (2012a), whereas the yellow boxes are re-computed twith no constraints on the phase. The boxes mark the middle 50 % of the GCM results (i.e. the interquartile range). See Table 1 for description of the functions used to generate this figure. Here the values for a, b, and c in Scafetta's Table 1 have been divided by 0.1, 0.04, and 0.1 respectively in order to provide comparable values. The grey backgrounds represent the 90 % confidence intervals. The whiskers extend to the most extreme data point which is no more than 1.5 times the interquartile range from the box.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

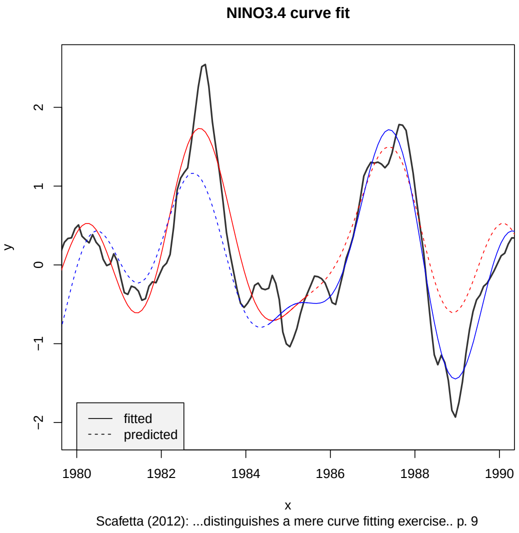

Fig. 4. An example demonstrating that for a small interval of the ENSO cycle, it is possible to find two cycles which seem to be part of a regular oscillation. This demonstration is produced with the call "ENSO.example()", which also shows a comparison between the fits and the rest of the data – for which they fail to track the ENSO evolution.