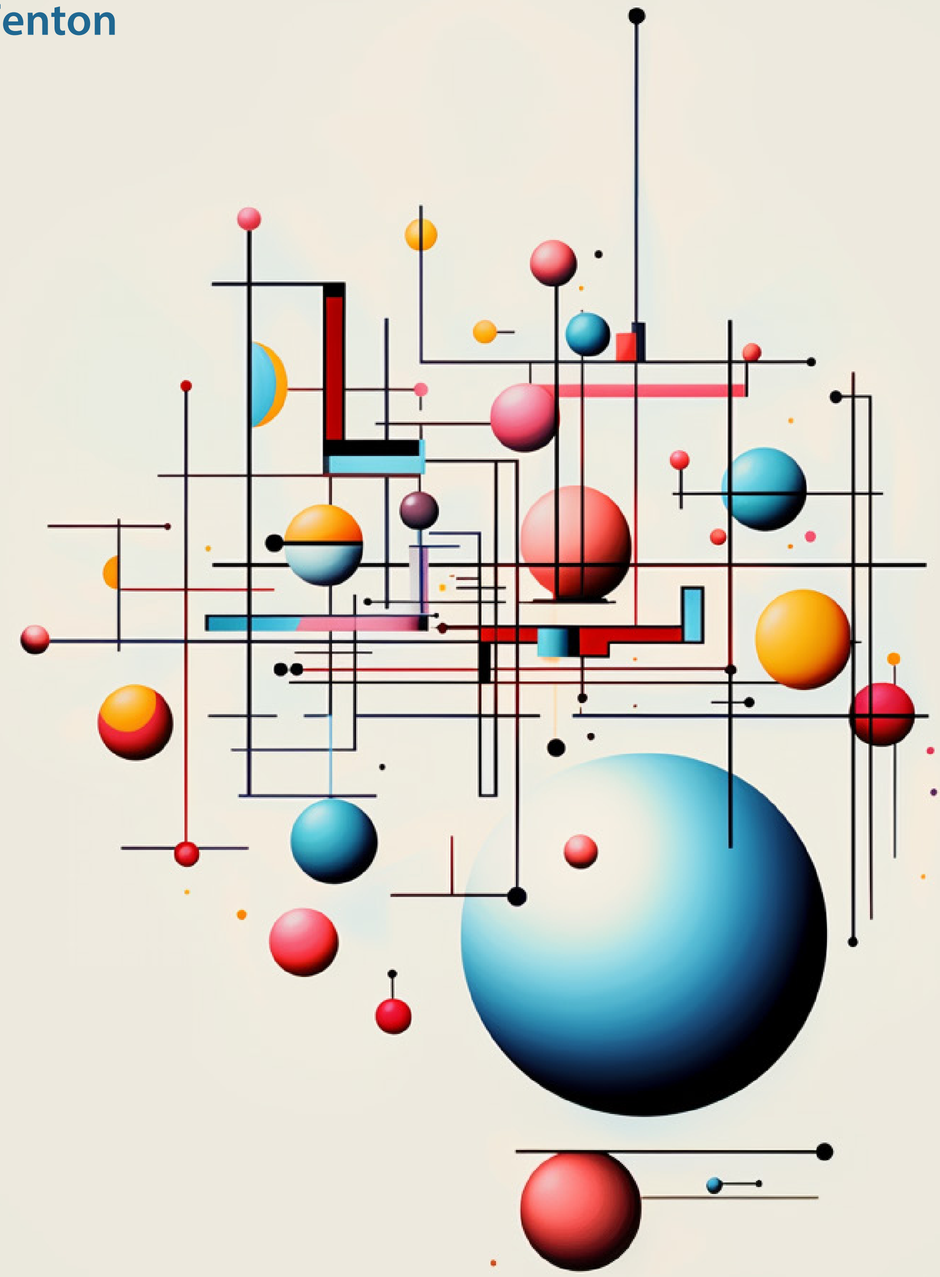
THE PROSECUTOR'S FALLACY AND THE IPCC REPORT

Norman Fenton

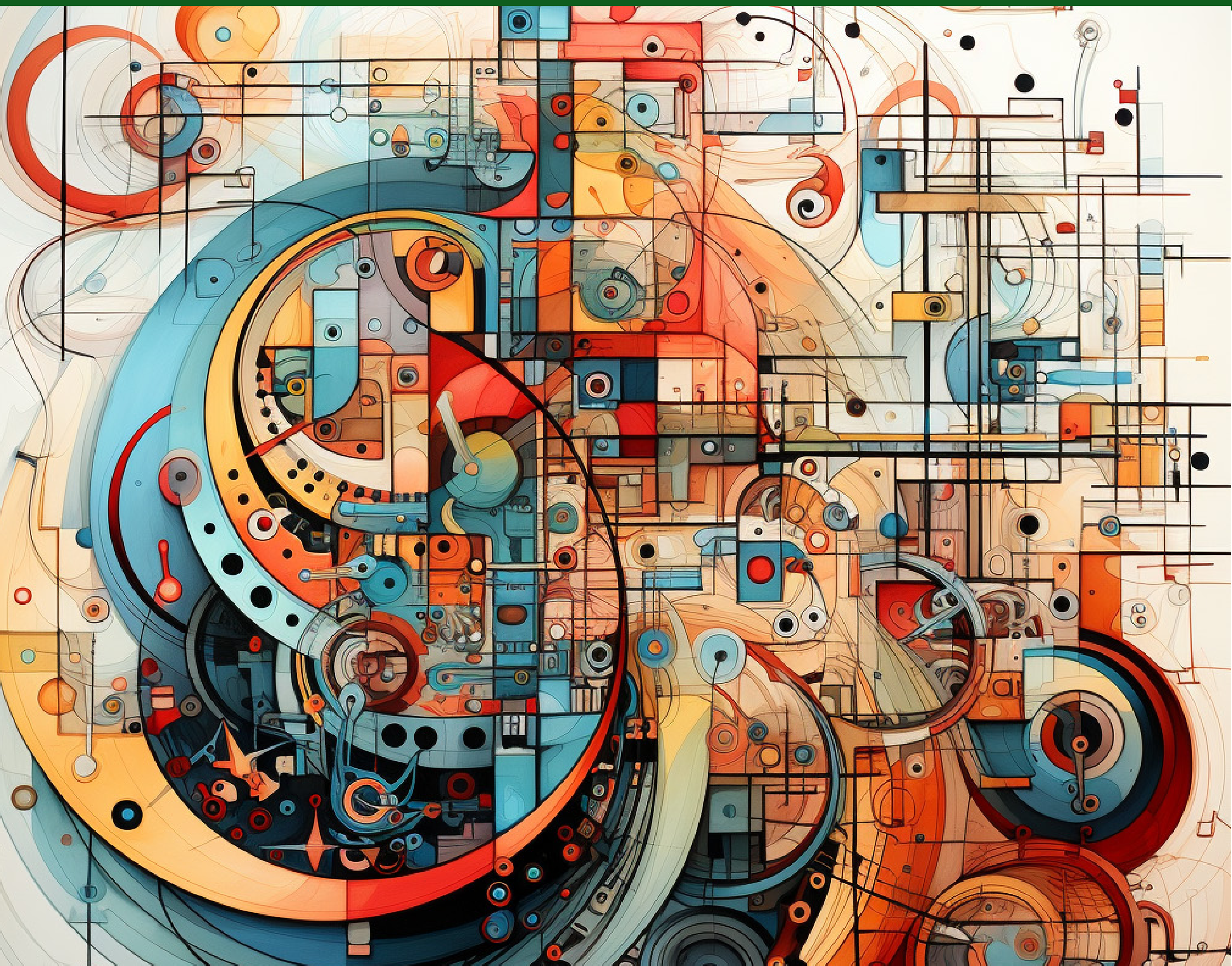


The Global Warming Policy Foundation
Note 45

The Prosecutor's Fallacy and the IPCC Report

Norman Fenton

Note 45, The Global Warming Policy Foundation



Contents

About the author

Norman Fenton is Professor Emeritus of Risk at Queen Mary University of London (retired as Full Professor, December 2022) and a Director of Agena, a company that specialises in artificial intelligence and Bayesian probabilistic reasoning. He is a mathematician by training, with a current focus on quantifying risk and uncertainty using causal, probabilistic models that combine data and knowledge (Bayesian networks). He has published 7 books and over 350 peer reviewed articles. His works covers multiple domains, including law and forensics, health, and system safety. Since 2020 he has been active in analysing data related to Covid risk.

Introduction

In 2015, I was one of the three presenters of the award winning BBC documentary called *Climate Change by Numbers*.[1] The concept of the programme was to take a new look at the climate change debate by focusing on three key numbers from what was then the most recent IPCC report. The numbers were:

- 0.85 degrees Centigrade – the amount of warming the planet has undergone since 1880;
- 95% – the degree of certainty climate scientists have that at least half the warming in the last 60 years is man-made;
- one trillion tonnes – the cumulative amount of carbon that can be burnt, ever, if the planet is to stay below 'dangerous levels' of climate change.

The idea was to get mathematicians/statisticians who had not been involved in the global warming debate to explain in lay terms how and why climate scientists had arrived at these three numbers. The other two presenters were Professor Hannah Fry (UCL) and Professor Sir David Spiegelhalter (Cambridge), and we were each assigned approximately 25 minutes on one of the numbers. My number was 95%.

On Bayesian and classical statistics

Being neither a climate scientist nor a classical statistician (my research uses Bayesian probability rather than classical statistics to reason about uncertainty), I have to say that I found the complexity of the climate models and their underlying assumptions to be daunting. The relevant sections in the IPCC report were extremely difficult to understand and they use assumptions and techniques that are very different to the Bayesian approach, in which we build causal models that combine prior expert knowledge with data.

In attempting to understand and explain how the climate scientists had arrived at their 95% figure, I used a football analogy, both because of my life-time interest in the sport and because – along with my colleagues Anthony Constantinou and Martin Neil – I have worked extensively on models for football prediction.

The climate scientists had performed what is called an 'attribution study' to understand the extent to which different factors – such as human carbon dioxide emissions – contributed to changing temperatures. This was analogous to my work in football, where we try to understand the extent to which different factors contribute to changing success of premiership football teams, as measured by the total number of points they achieve season-by-season. So, for the programme, we drew out this analogy, generating a model of football success in terms of various factors (but using classical rather than Bayesian statistical methods, so as to be consistent with the climate scientists' approach).

Unlike the climate models, which involve thousands of variables, we had to restrict ourselves to a very small number (due to a combination of time limitations and lack of data). Specifically, for each team and each year we considered:

- wages (this was the single financial figure we used)
- total days of player injuries
- manager experience
- squad experience
- number of new players.

The statistical model generated from these factors produced, for most teams, a good fit of success over the years for which we had the data. Our 'attribution study' showed wages was by far the major influence; when it was removed from the study, the resulting statistical model was not a good fit. This was analogous to what the climate scientists' models were showing when the human carbon dioxide emissions factor was removed from their models; the previously good fit to temperature was no longer evident. And, analogous to the climate scientists' 95% derived from their models, we were able to conclude there was a 95% chance that an increase in wages of 10% would result in at least one extra Premiership point. This football model was extremely crude and, after the programme, we produced a more realistic peer-reviewed attribution model using Bayesian networks, which showed that wages offered was one of the many factors (and not even the greatest) influencing performance.[2]

The issues with my number

Obviously, there was no time in the programme to explain either the details or the limitations of my hastily-put-together football attribution study, but the programme also did not have the time or scope to address the complexity of some of the broader statistical issues involved in the climate debate (including issues that lead some climate scientists to claim the 95% figure is underestimated and others to believe it is overestimated). In particular, the issues that were not covered were:

- Any real details of the underlying statistical methods and assumptions. For example, there has been controversy about the way a method called Principal Component Analysis was used to create the famous 'Hockey Stick' graph, which appeared in previous IPCC reports.
- Assumptions about the accuracy of historical temperatures. Much of the climate debate (such as that concerning the exceptionalness of the recent rate of temperature increase) depends on assumptions about historical temperatures dating back thousands of years. There has been some debate about whether sufficiently large variances were used.
- Variety and choice of models. There are many common assumptions in all of the statistical models used by the IPCC

and it has been argued that there are alternative models not considered by the IPCC which provide an equally good fit to climate data, but which do not support the same conclusions.

But the most important flaw concerns the real probabilistic meaning of the 95% figure. This is considered in the next section.

What the number does and does not mean

Recall that the particular 'climate change number' that I was asked to explain was the number 95: specifically, relating to the assertion made in the IPCC 2013 Report of 'at least 95% degree of certainty that more than half the recent warming is man-made'. The 'recent warming' related to the period 1950–2010. So, the assertion is about the probability of humans causing most of this warming.

Before explaining the problem with this assertion, we need to make clear that (although superficially similar) it is very different to another more widely known assertion (still promoted by NASA) that '97% of climate scientists agree that humans are causing global warming and climate change'. That assertion was simply based on a flawed survey of authors of published papers and has been thoroughly debunked.[3,4]

The 95% degree of certainty is a more serious claim. But the case made for it in the IPCC report is also flawed. To explain why, it is useful to illustrate the flaw with a simple motivating example.

The fundamental flaw: coin tossing example

Imagine that there are known to be some double-headed coins in circulation. Suppose a coin is randomly selected and, without inspecting it, it is tossed five times. Each time the result is heads. What is the probability that the coin is double-headed? Most people intuitively believe it is very likely to be one of the double-headed coins. But that is a fallacy.

In classical statistical hypothesis testing, it is not possible to make any direct conclusions about the hypothesis that the coin is double-headed. Instead, the observation of the five consecutive heads is used to either accept or reject the 'null hypothesis' (the 'opposite' statement to the one believed to be true, so in this case that the coin is *not* double-headed) at some agreed level of significance. Specifically, we compute the probability that we would have observed five consecutive heads if the coin was not double-headed. In this case the probability is 1/32 which is about 3%. So that is indeed very unlikely. Typically, a 5% level of significance (also called the p-value) is used, meaning that we 'reject' the null hypothesis in this case because the probability is less than 5%.

Note that we can equivalently conclude that there is a very high probability (97%) that we would not have observed five consecutive heads if the coin was not double headed.

Unfortunately, people often conclude (wrongly, as we will show) that rejecting the null hypothesis at the 5% significance level means there is less than 5% probability that the coin is not double-headed. And hence they further conclude that we can be at least 95% confident that the coin is double-headed. But that is wrong.

While the evidence of the five consecutive heads certainly provides support for the hypothesis that the coin is double-headed, it tells us nothing about the probability that it *really is* double-headed. The only way we can make any firm conclusion about that probability is if we have some knowledge of the 'prior probability' that the coin was double-headed; in this case that means knowing what proportion of coins in circulation are double-headed. It will make a big difference if is 1 in 2, 1 in 100, 1 in 1000, 1 in a million etc.

If we know the proportion of double headed coins in circulation, then Bayes' theorem can be used to calculate the answer we seek. Let's suppose, for example, that we know there are 1 in 500 double headed coins in circulation (so the prior probability a coin is double-headed is 1 in 500 which is 0.2%). The formal calculation is set out in the appendix, but we can give an intuitive explanation without resorting to the Bayes formula:

- Imagine a bag of 500 coins in which exactly one is double-headed (i.e. a typical bag of coins in this case). Suppose we test each coin by tossing it five times. Then we are certain that the (one) double-headed coin will result in 5 heads.
- But, 1 in every 32 of the other 499 fair coins - that is about 16 fair coins - will also result in five consecutive heads.
- So for every 17 coins recording five consecutive heads, there is only one which is double-headed.
- So, if we know that a coin has recorded five consecutive heads what we can conclude is that there is a 1 in 17 chance (that is about 6%) that it is double headed, i.e. about a 94% chance it is not double-headed.

So, whereas it is very unlikely to observe 5 consecutive heads if the coin is not double headed (probability 3%), it is still very likely that the coin is not double headed (probability 94%).

The fallacy of concluding that there was only a small probability that the coin is not double headed is called the 'fallacy of the transposed conditional' (or the 'prosecutor fallacy') because we have assumed that:

- the probability of an assertion E given an assertion '$\text{not } H$'

is the same as

- the probability of '$\text{not } H$' given E.

In this case:

- H is the hypothesis: 'selected coin is double-headed'
- E is the evidence: '5 consecutive heads tossed'

And we have shown that:

- Probability of (E given not H) = 3%

whereas:

- Probability of ('not H' given E) = 94%

The flaw in the IPCC summary report

It turns out that the assertion that 'at least 95% degree of certainty that more than half the recent warming is man-made' is based on the same fallacy. In my article about the programme, I highlighted this concern as follows:

> The real probabilistic meaning of the 95% figure.
>
> In fact it comes from a classical hypothesis test in which observed data is used to test the credibility of the 'null hypothesis'. The null hypothesis is the 'opposite' statement to the one believed to be true, i.e. 'Less than half the warming in the last 60 years is man-made'. If, as in this case, there is only a 5% probability of observing the data if the null hypothesis is true, the statisticians equate this figure (called a p-value) to a 95% confidence that we can reject the null hypothesis. But the probability here is a statement about the data given the hypothesis. It is not generally the same as the probability of the hypothesis given the data (in fact equating the two is often referred to as the 'prosecutors fallacy', since it is an error often made by lawyers when interpreting statistical evidence).

The claim that there was at least 95% probability that more than half the warming was man-made was made in the 'Summary for Policymakers' section of the 2013 IPCC Report.

> It is extremely likely [defined as 95–100% certainty] that more than half of the observed increase in global average surface temperature from 1951 to 2010 was caused by the anthropogenic [human-caused] increase in greenhouse gas concentrations and other anthropogenic forcings together.

They defined 'extremely likely' as at least 95% probability.

The basis for the claim is found in Chapter 10 of the detailed Technical Summary, which describes various climate change simulation models, which reject the null hypothesis (that more than half the warming was not man-made) at the 5% significance level. Specifically, in the simulation models, if you assumed that there was little man-made impact, then there was less than 5% chance of observing the warming that has been measured. In other words, the models do not support the null hypothesis of little man-made climate change. The problem is that, even if the models were accurate (and it is unlikely that they are) we cannot conclude that there is at least a 95% chance that more than half the warming was man-made, because doing so is the fallacy of the transposed conditional.

All we can conclude is that there is at least a 95% probability we would not observe the warming we have seen based on the climate change model simulations and their multiple assumptions. Just like there was a 97% probability we would not observe five consecutive heads on a coin that was not double-headed.

The illusion of confidence in the coin example comes from ignoring (the 'prior probability') of how rare the double-headed coins are. Similarly, in the case of climate change there is no allowance made for the prior probability of man-made climate change, i.e. how likely it is that humans rather than other factors such as solar activity cause most of the warming. After all, previous periods of warming certainly could not have been caused by increased greenhouse gases from humans, so it seems reasonable to assume – before we have considered any of the evidence – that the probability humans caused most of the recent increase in temperature to be very low; only the assumptions of the simulation models are allowed, and other explanations are absent. In both of these circumstances, classical statistics can then be used to deceive you into presenting an illusion of confidence when it is not justified.

Conclusion

Although I obviously have a bias, my enduring impression from working on the programme is that the scientific discussion about the statistics of climate change would benefit from a more extensive Bayesian approach. Recently, some researchers have started to do this, but it is an area where I feel causal Bayesian network models could shed further light, and this is something that I would strongly recommend.

Appendix: Bayes' Theorem calculation for double-headed-coin example

- H is the hypothesis: 'selected coin is double-headed'
- E is the evidence: '5 consecutive heads tossed'

We are assuming $\mathrm{P}(H) = 1/500$, so $\mathrm{P}(\text{not } H) = 499/500$, and we we know $\mathrm{P}(E|\text{not } H) = 1/32$ and $\mathrm{P}(E|H) = 1$.

Then:

$$\mathrm{P}(\text{not } H|E) = \frac{\mathrm{P}(E|\text{not } H) \times \mathrm{P}\,(\text{not } H)}{\mathrm{P}(E|\text{not } H) \times \mathrm{P}(\text{not } H) + \mathrm{P}(E|H) \times \mathrm{P}(H)}$$

$$= \frac{(1/32) \times (499/500)}{(1/32) \times (499/500) + 1 \times (1/500)}$$

$$= \quad 0.94 = 94\%$$

Acknowledgements

I would like to thank the BBC team (especially Jonathan Renouf, Alex Freeman, Eileen Inkson, and Gwenan Edwards) for their professionalism, support, encouragement, and training; and my colleagues Martin Neil and Anthony Constantinou for their technical support and advice.

Notes

1. First screened on BBC4 on 2 March 2015.
2. http://bayesknowledge.blogspot.com/2017/03/explaining-and-predicting-football-team.html
3. http://richardtol.blogspot.com/2013/08/open-letter-to-vice-chancellor-of.html
4. https://www.econlib.org/archives/2014/02/david_friedman_14.html.

About the Global Warming Policy Foundation

People are naturally concerned about the environment, and want to see policies that protect it, while enhancing human wellbeing; policies that don't hurt, but help.

The Global Warming Policy Foundation (GWPF) is committed to the search for practical policies. Our aim is to raise standards in learning and understanding through rigorous research and analysis, to help inform a balanced debate amongst the interested public and decision-makers. We aim to create an educational platform on which common ground can be established, helping to overcome polarisation and partisanship. We aim to promote a culture of debate, respect, and a hunger for knowledge.

RECENT GWPF NOTES

9	Anthony Kelly	Global Warming and the Poor
10	Susan Crockford	Healthy Polar Bears, Less Than Healthy Science
11	Andrew Montford	Fraud, Bias and Public Relations
12	Harry Wilkinson	UK Shale Developments
13	Peter Lilley	The Helm Review and the Climate-Industrial Complex
14	Constable and Hughes	Bubble or Babble?
15	Patricia Adams	The Road from Paris: China's Climate U-Turn
16	Mikko Paunio	Save the Oceans: Stop Recycling Plastic
17	John Christy	The Tropical Skies: Falsifying Climate Alarm
18	Gordon Hughes	Who's the Patsy? Offshore Wind's High-stakes Poker Game
19	J. Ray Bates	What Caused Hurricane Lorenzo?
20	Andrew Montford	Reducing Emissions Without Breaking the Bank
21	Andrew Montford	£3 Trillion and Counting: Net Zero and the National Ruin
22	David Whitehouse	The Next Solar Cycle and Why it Matters for Climate
23	Gautam Kalghatgi	The Battery Car Delusion
24	John Constable	Who Are They Fooling? The UK Energy White Paper 2020
25	William M Briggs	The Climate Blame Game: Are We Really Causing Extreme Weather?
26	Ross McKitrick	Suboptimal Fingerprinting
27	William M Briggs	How the IPCC Sees What Isn't There
28	J. Ray Bates	Polar Sea Ice and the Climate Catastrophe Narrative
29	Susan Crockford	The State of the Polar Bear 2021
30	Michael Kelly	Achieving Net Zero: A Report from a Putative Delivery Agency
31	Michael Kelly	Net Zero for New Zealand
32	Andrew Montford	Ten Things Everyone Should Know About Climate Models
33	Peter Ridd	The Good News on Coral Reefs
34	William Kininmonth	Rethinking the Greenhouse Effect
35	Alexander Stahel	The Crisis of the European Energy System
36	Ralph Alexander	Chinese Whispers: How Climate Science Gets Lost in Translation
37	Wade Allison	Confidence in Nuclear Energy
38	Bryan Leyland	Do We Need A New Electricity Market?
39	Michael Kelly	The Feasibility of a Net-zero Economy for the USA by 2050
40	Wade Allison	The Inadequacy of Wind Power
41	Wade Allison	The Provision of Energy Without Fear
42	Mikko Paunio	Microplastics: The Environmental Harms of the Circular Economy
43	Andrew Montford	Heat Pumps: Mythology and Actuality
44	Tim Worstall	Taxing the North Sea
45	Norman Fenton	The Prosecutor's Fallacy and the IPCC Report

For further information about the Global Warming Policy Foundation, please visit our website at www.thegwpf.org.
The GWPF is a registered charity, number 1131448.