Calibrating Statistical Tools: Improving the Measure of Humanity's Influence on the Climate

Corey Dethier

Leibniz Universität Hannover
Philosophy Department
corey.dethier@gmail.com

June 29, 2022

Intro

Attribution

Attribution: study of how factors such as greenhouse gasses have contributed to climate change.

Attribution studies often involve what looks like a kind of *measurement*.

Over the last thirty years, the statistical techniques used in attribution have changed dramatically.

Questions

Some questions:

- How should we understand changes in statistical techniques?
- What makes one statistical technique "better" than another in the context of climate science?
- How does the general desire for better statistical techniques get translated into specific changes?

See Dethier (forthcoming).

Questions

Some questions:

- How should we understand changes in statistical techniques?

Questions

Some questions:

- How should we understand changes in statistical techniques?

Changes should be understood as a kind of *calibration*.

Specifically in the sense of Bokulich (2020): a change to either a physical process or the means of inferring outcomes from said process that aims making the measurement procedure as a whole deliver more reliable results.

The plan

1. Late 80s, early 90s: detecting climate change.
2. Late 90s, early 00s: the introduction of fingerprinting.
3. Late 00s, early 10s: refining fingerprinting.
4. What does it mean to calibrate statistics?

Late 80s, early 90s: detecting climate change

The problem

Question: is the earth warming, and if so, why?

Data: decades of temperature readings from around the globe.

Problem: the data is *extremely* noisy.

First pass

We can think of our data as being comprised of two parts:

$$\text{Observations } (\mathbf{O}) = \text{Signal } (\mathbf{O_S}) + \text{Noise } (\mathbf{O_N})$$

Assume that changes in temperature will be *randomly distributed*.

This (false) assumption allows us to massively simplify the problem.

Resulting model

The resulting statistical model:

$$\bar{\mathbf{O}} \sim \mathcal{N}(\mathbf{O_S}, \sigma^2)$$

which says, roughly, that we should expect the average temperature to be normally distributed around our signal.

We can now use standard χ^2 tests and/or least-square analyses to determine $\mathbf{O_S}$.

"Detection"

Roughly the statistical assumptions just outlined were standard through the early 90s (see Santer, Wigley, et al. 1996).

"Detection"

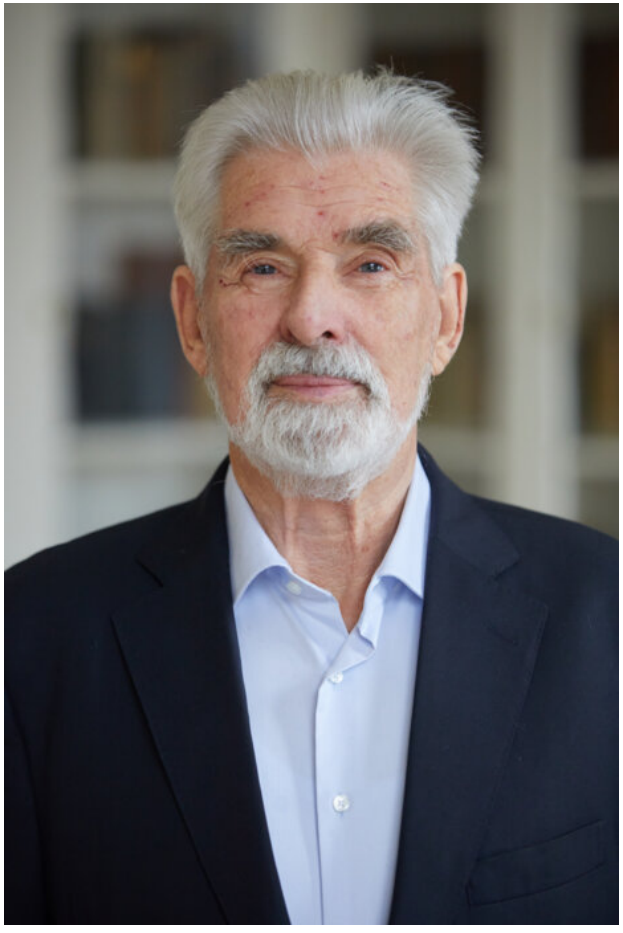
Roughly the statistical assumptions just outlined were standard through the early 90s (see Santer, Wigley, et al. 1996).

Arguably good enough for the reliable *detection* of global warming.

But not reliable enough to determine what was causing it.

Late 90s, early 00s: the introduction of fingerprinting

Introducing fingerprinting



Technical papers in the late 80s outlined a better approach: "fingerprinting."

Most influential were those of Hasselmann (1993, 1997).

These began to be put into practice in the mid-90s.

See, e.g., North et al. (e.g. 1995) and Santer, Taylor, et al. (1995).

How does it work?

What we want is a replacement for $\bar{\mathbf{O}}$—that is, we want some manageable subset of the data.

Ideally, we'd like to find that part of the data that maximizes the ratio between the signal ($\mathbf{O_S}$) and noise ($\mathbf{O_N}$).

How do we find and isolate that part of the data?

Fingerprinting: intuitive idea

Climate models tell us both (a) where CO_2 (say) will have a large effect and (b) where we will find high amounts of variability.

So we just look at the places that appear on the first list but not on the second list.

If the signal appears in those places, then we have good reason to believe that it is real—and if it doesn't, then we have good reason to believe it isn't.

Fingerprinting: (only) slightly more technically

Call the predicted signal $\mathbf{M_S}$, and the predicted covariance matrix for the noise $\mathbf{\Sigma_M}$.

If we assume that these predictions are perfectly accurate, then

$$\mathbf{\Sigma_M}^{-1}\mathbf{M_S}\mathbf{O}$$

maximizes the signal-to-noise ratio.

Basically: we weight the original data so that the areas with more (predicted) signal will be more emphasized and those with more (predicted) noise will be de-emphasized.

Calibration and de-idealizing

Replacing $\bar{\mathbf{O}}$ with $\mathbf{\Sigma_M}^{-1}\mathbf{M_S O}$ essentially involves a (partial) de-idealization of the statistical model.

Resulting statistical model expected to give better (but still imperfect) tests of attribution hypotheses.

But it brings with it its own complications similar to those described by Bokulich (2020) in the calibration of carbon dating.

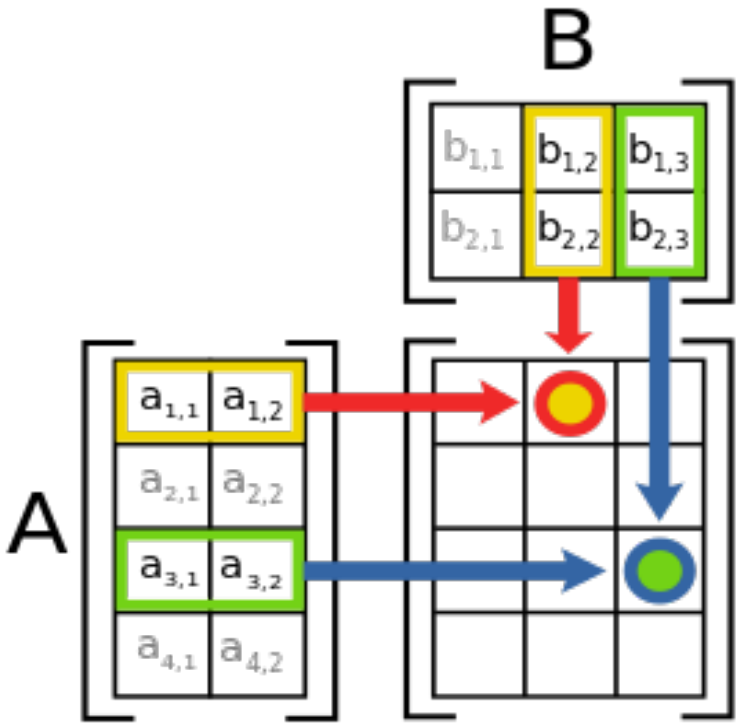
Late 00s, early 10s: refining fingerprinting

The problem

We know that $\mathbf{M_S}$ and $\mathbf{\Sigma_M}$ aren't perfectly accurate (Hasselmann 1993).

Problem: "inverting" $\mathbf{\Sigma_M}$—that is, turning it into $\mathbf{\Sigma_M}^{-1}$—tends to multiply the error.

(Slightly) more technically



- Wikimedia Commons

B is the inverse of **A** when $\mathbf{AB} = \mathbf{I}$.

One error in **A** lead to errors throughout **B**.

Potential solutions

More models runs $\Rightarrow$ better control over $\mathbf{\Sigma_M}^{-1}$.

Potential solutions

More models runs $\Rightarrow$ better control over $\mathbf{\Sigma_M}^{-1}$.

- We'd need 10-100x more (Ribes, Planton, and Terray 2013).

Potential solutions

More models runs $\Rightarrow$ better control over $\mathbf{\Sigma_M}^{-1}$.

- We'd need 10-100x more (Ribes, Planton, and Terray 2013).

"Truncation": cut out many variables (and thus much of the data)—again, more control over $\mathbf{\Sigma_M}^{-1}$.

Potential solutions

More models runs $\Rightarrow$ better control over $\mathbf{\Sigma_M}^{-1}$.

- We'd need 10-100x more (Ribes, Planton, and Terray 2013).

"Truncation": cut out many variables (and thus much of the data)—again, more control over $\mathbf{\Sigma_M}^{-1}$.

- Arbitrary and less informative.

Regularization

Introduced by Ribes, Azaïs, and Planton (2009).

Roughly: rather than use $\mathbf{\Sigma_M}$ itself, introduce a prior distribution and then update on $\mathbf{\Sigma_M}$.

As before: not getting rid of all errors or creating a genuinely accurate picture of the "true variation."

Only address the problem of inversion.

What does it mean to calibrate statistics?

Statistical models

"Statistical models" are representational tools.

These tools can be tweaked to make the system as a whole more reliable.

And in fact they are!

Evaluating statistical models

Note, however: improving the statistical model doesn't necessarily mean making the model more accurate qua representation of "how the data are [actually] generated."

Instead, we need an "adequacy-for-purpose" (Parker 2020) picture of what makes "statistical models" better or worse.

- Bokulich, Alisa (2020). Calibration, Coherence, and Consilience in Radiometric Measures of Geologic Time. *Philosophy of Science* 87.3: 425–56.
- Dethier, Corey (forthcoming). Calibrating Statistical Tools: Improving the Measure of Humanity's Influence on the Climate. *Studies in the History and Philosophy of Science*.
- Hasselmann, Klaus (1993). Optimal Fingerprints for the Detection of Time-dependent Climate Change. *Journal of Climate* 6.10: 1957–71.
- — (1997). Multi-pattern Fingerprint Method for Detection and Attribution of Climate change. *Climate Dynamics* 13: 601–11.
- North, Gerald R. et al. (1995). Detection of Forced Climate Signals. Part I: Filter Theory. *Journal of Climate* 8.3: 401–8.
- Parker, Wendy S. (2020). Model Evaluation: An Adequacy-for-Purpose View. *Philosophy of Science* 87.3: 457–77.


Ribes, Aurélien, Jean-Marc Azaïs, and Serge Planton (2009). Adaptation of the optimal Fingerprint Method for Climate Change Detection using a Well-conditioned Covariance Matrix Estimate. *Climate Dynamics* 33.5: 707–22.


Ribes, Aurélien, Serge Planton, and Laurent Terray (2013). Application of Regularised Optimal Fingerprinting to Attribution. Part I: Method, Properties and Idealised Analysis. *Climate Dynamics* 41.11-12: 2817–36.


Santer, Benjamin D., Karl E. Taylor, et al. (1995). Towards the Detection and Attribution of an Anthropogenic Effect on Climate. *Climate Dynamics* 12: 79–100.


Santer, Benjamin D., Tom M. L. Wigley, et al. (1996). Detection of Climate Change and Attribution of Causes. In: *Climate Change 1995: The Science of Climate Change*. Ed. by John T. Houghton et al. Cambridge: Cambridge University Press.