Searching for Acceleration of Global Warming in the Global Temperature Dataset

William Menke, October 19, 2023

(After reading Hausfather's Op-ed and after a discussion with Alberto Malinverno)

Motivation. A recent NY Times Op-ed by Zeke Hausfather (Ref. 1) contained this graphic

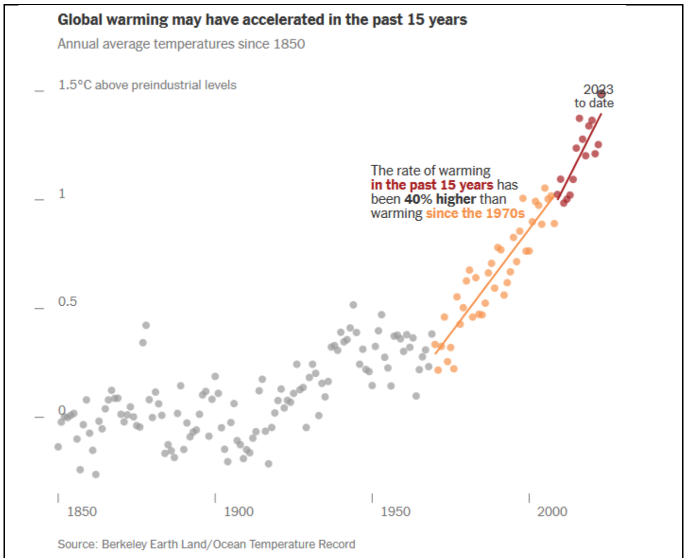

Fig. 1. Global temperature as a function of time, from Ref. 1, showing fits from different time intervals with different slopes. From Ref. 1.

which the author uses to argue that there is evidence that the increase in global temperatures is accelerating. The purpose of my analysis is to determine whether this increase is statistically significant to the 95% confidence level.

Data Preparation. I download a fresh copy of the GISS global temperature dataset (file GLB.Ts+dSST.csv, Ref. 3 and 4) from the GISS Website and extracted a table of global means:

Col A (Year) and Col N (J-D), Rows 93-145 (1970 thru 2022)

I note that this is a different source of data than cited in the Op-ed, which is given as Berkeley Earth Land/Ocean Temperature Record. I have overlaid the Op-ed's plot of the Berkely data with my plot of the GISS data (as shown in Fig. 2) and determined that they identical (or at least nearly so).

Methodology. The least squares and statistical test methodologies used here are completely standard. The notation follows that used in my textbook (Ref. 2).

Method A. Testing for a steady acceleration in the rate of temperature increase. I first scaled the 1970-2022 time range to the interval $(-1,1)$ and then least squares fit the temperature data, $d(z)$, to the three-parameter model:

$$d^{obs}(z_i) = m_0 P_0(z_i) + m_1 P_1(z_i) + m_2 P_2(z_i)$$

which defines the matrix equation:

$$\mathbf{d} = \mathbf{G}\mathbf{m}$$

Here, the ms are unknown model parameters and the P's are Legendre polynomials (LPs). I use LPs because they are mutually orthogonal on the $(-1,1)$ interval, which simplifies the interpretation of variances. The test is based on the posterior covariance matrix, $\mathbf{C}_m = \sigma_d^2[\mathbf{G}^T\mathbf{G}]^{-1}$, where σ_d^2 is the posterior covariance of the data, determined by scatter about the best-fit model. These LPs represent constant, linear and quadratic curves, with the quadratic term representing acceleration. The fit (Fig. 2)

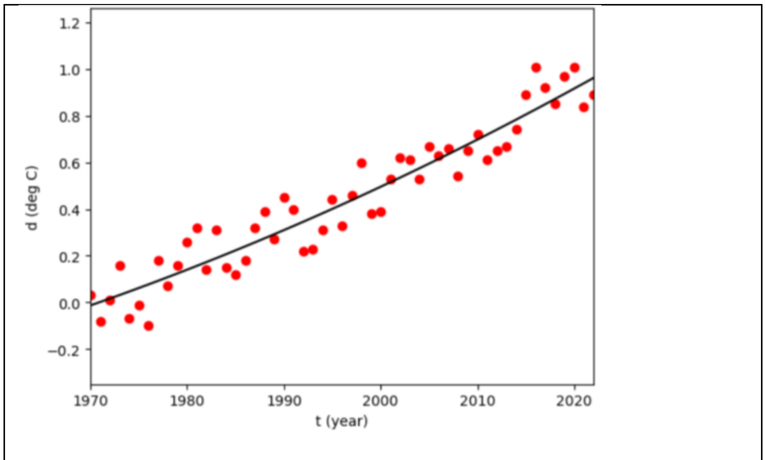

Fig. 1. GISS global temperature data (red dots) and LP fit (black curve). Data from GISS Website (Refs. 3 & 4).

The best-fit curve has positive curvature that is consistent with accelerated warming. However, the coefficient of P_1 is not positive to 95% certainty:

```
Least squares solution:
  m0: 0.4376 +/- 0.0358 (95 pct) (significantly different from zero)
  m1: 0.4874 +/- 0.0607 (95 pct) (significantly different from zero)
  m2: 0.0368 +/- 0.0770 (95 pct) (not significantly different from zero)
```

(Because of the time transformation, the units of these measurements are not "per year", but rather "per $52/2$ years").

Thus, the possibility that the curvature arises from random variation cannot be ruled out.

This is in contrast to the m_1 (linear) coefficient, which is extremely significant. Global Warming is definitely occurring.

Method B. Testing for a sudden acceleration in the rate of temperature increase using the F-test. I examined two models

$$d^{obs}(z_i) = m_0^A + m_1^A t_i$$

and

$$d^{obs}(z_i) = \begin{Bmatrix} m_0^B + m_1^B(t_i - t_0) & t < t_0 \\ m_0^B + m_2^B(t_i - t_0) & t \geq t_0 \end{Bmatrix}$$

In the case of Model B, the two line-segments meet at time t_0, with an intercept of m_0^B.

The first model is just a linear increase with time. The second model is a continuous curve consisting of two straight lines of different slopes, on either side of a point, t_0. This point is taken to be a known parameter. I focus here on $t_0 = 2008$, but I repeated the test for all t_0s in the range $1975 - 2018$ and obtained similar results.

The best-fit Model B (Fig. 3) has an increase in slope that is consistent with accelerated warming.

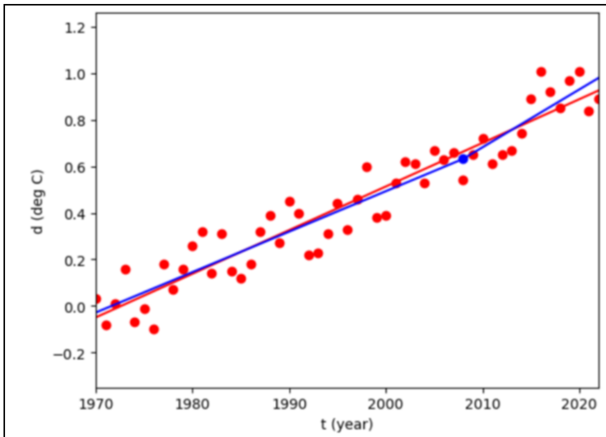

Fig. 3. GISS global temperature data (red dots) and linear fit (red line) and linked line fit (blue curve), with kink in 2008 (blue dot). Data from GISS Website (Refs. 3 and 4).

The decrease in mean-squared error, compared to the linear fit, is relatively small, only 4%, and the value of $F = 1.03$ is very close to unity, something that can be expected to happen 91.6% of the time due to random variation alone.

```
t0 2008, relative error 0.048346, F 1.0302, P 0.9162, significant no
```

Although the introduction of a kink improves the fit, it does not do so at the 95% confidence level. None of the other choices of t_0 are significant either. (The year 2009 has the lowest probability, 91.3%, of being caused by random error, just a hair lower than 2008).

Method C. Testing for a sudden acceleration in the rate of temperature increase by examining the difference slopes on either side of the kink. This method used the same model as in B but analyzes it differently. For a fixed t_0, the estimated model parameters, $\mathbf{m}^B$, and their posterior covariance, $\mathbf{C}_m$, are computed using least squares.

The difference in slopes $\Delta m = m_2^B - m_1^B = \mathbf{M}\mathbf{m}^{est}$ (with $\mathbf{M} = [0 \quad -1 \quad 1]$) and its variance, $\sigma_{\Delta m}^2 = \mathbf{M}\mathbf{C}_m\mathbf{M}^T$ are computed from the least squares solution $\mathbf{m}^{est}$. The ratio, $r = 2\sigma_{\Delta m}/|\Delta m|$ is examined; only when $r < 1$ can one exclude the possibility that $\Delta m = 0$ to 95% confidence. The results for $t_0 = 2008$ indicate that the possibility that that the difference in slopes is due to random variation cannot be excluded.

```
t0 2008, Dslope 0.0074 +/- F 0.0093 (95 pct), ratio 1.2549, significant no
```

The results for $t_0 = 2011$ has a slightly smaller ratio, $r = 1.22$, which is also greater than unity.

Effect of Correlation. My methodology assumes that the errors in the data are uncorrelated and with uniform variance, σ_d^2. In reality, the error in the data ought to be construed as "model error" (as contrasted to "observational error"), as "random" but real climate processes, such as ENSO probably are the main source of the scatter, and are likely correlated from year to year (as is ENSO). However, although the results would be different is we assumed positive correlation of the error for neighboring years, they would indicate more likelihood that differences were due to random variation, and not less, for correlation reduces the effective number of degrees of freedom of the dataset.

Conclusions. Although it is true that the rate of global temperature increase is higher for 2008-2022 than for 1970-2007, the difference is not statistically significant. If the rate is increasing, the current global temperature data are insufficient to demonstrate it at the 95% confidence level.

Caveat. Nothing in my analysis precludes the possibility that other climate data – meaning data that I have not analyzed here – demonstrates a statistically significant acceleration.

References

1. Hausfather, Zeke, I Study Climate Change. The Data Is Telling Us Something New, New York Times, Op-ed Section, Oct. 13, 2023. https://www.nytimes.com/2023/10/13/opinion/climate-change-excessive-heat-2023.html

2. Menke, W. (2022), Environmental Data Analysis with MATLAB or Python, Third Edition (textbook), Elsevier, ISBN 9780323955768.

3. GISS Global Mean Temperature Dataset:, https://data.giss.nasa.gov/gistemp/tabledata_v4/T_AIRS/GLB.Ts+dSST.csv

4. Hansen, J., R. Ruedy, Mki. Sato, and K. Lo, 2010: Global surface temperature change. Rev. Geophys., 48, RG4004, doi:10.1029/2010RG000345 (last accessed 10/19/2023).