

A Simple Deconstruction of the HadCRU Global-Mean Near-Surface Temperature Observations

Michael E. Schlesinger, Daniela Lindner, Michael J. Ring, Emily F. Cross

Climate Research Group, Department of Atmospheric Sciences, University of Illinois at Urbana-Champaign, Urbana, USA
Email: schlesin@illinois.edu

Received May 12, 2013; revised June 15, 2013; accepted June 23, 2013

Copyright © 2013 Michael E. Schlesinger *et al.* This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT

Previously we have used Singular Spectrum Analysis (SSA) to deconstruct the global-mean near-surface temperature observations of the Hadley Centre—Climate Research Unit that extend from 1850 through 2012. While SSA is a very powerful tool, it is rather like a statistical “black box” that gives little intuition about its results. Accordingly, here we use the simplest statistical tool to provide such intuition, the Simple Moving Average (SMA). First we use a 21-year SMA. This reveals a nonlinear trend and an oscillation of about 60 years length. Second we use a 61-year SMA on the raw observations. This yields a nonlinear trend. We subtract this trend from the raw observations and apply a 21-year SMA. This yields a Quasi-periodic Oscillation (QPO) with a period and amplitude of about 62.4 years and 0.11°C. This is the QPO we discovered in our 1994 Nature paper, which has come to be called the Atlantic Multidecadal Oscillation. We then subtract QPO-1 from the detrended observations and apply an 11-year SMA. This yields QPO-2 with a period and amplitude of about 21.0 years and 0.04°C. We subtract QPO-2 from the detrended observations minus QPO-1 and apply a 3-year SMA. This yields QPO-3 with a period and amplitude of about 9.1 years and 0.03°C. QPOs 1, 2 and 3 are sufficiently regular in period and amplitude that we fit them by sine waves, thereby yielding the above periods and amplitudes. We then subtract QPO-3 from the detrended observations minus QPOs 1 and 2. The result is too irregular in period and amplitude to be fit by a sine wave. Accordingly we represent this unpredictable part of the temperature observations by a Gaussian probability distribution (GPD) with a mean of zero and standard deviation of 0.08°C. The sum of QPOs 1, 2 and 3 plus the GPD can be used to project the natural variability of the global-mean near-surface temperature to add to, and be compared with, the continuing temperature trend caused predominantly by humanity’s continuing combustion of fossil fuels.

Keywords: Climate Change; Global Warming; Natural Variability

1. Introduction

Since 1994 we have published four scientific papers wherein we used Singular Spectrum Analysis (SSA) to analyze up to four observational records of global-mean near-surface temperature. In those papers, the SSA results were obtained as if from a statistical “black box”, that is, all at once and not intimately connected to the observed temperatures. Here we remedy this by connecting the SSA results to the observed temperatures. We will do this systematically, step by step. In particular we will use Simple Moving Averages (SMAs) of different time periods to reveal an aspect of the observed temperatures. We will compare this aspect to what SSA shows more completely. In this way we will link the features revealed by SSA directly to the temperature observations. As we do this with each SSA feature, we will subtract it

from the observed temperatures and then apply another SMA to reveal the next aspect of the observed temperatures.

In our two most-recent papers on this matter, “Causes of the Global Warming Observed Since the 19th Century” ([1]; hereafter Causes) and “A Fair Plan to Safeguard Earth’s Climate: 3. Outlook for Global Temperature Change Throughout the 21st Century” ([2]; hereafter FP3), we applied SSA to four observational datasets of global-mean near-surface temperature: 1) the Hadley Centre-Climate Research Unit (HadCRU) located in the United Kingdom, with data starting in 1850 [3]; 2) the National Climate Data Center of the US National Oceanographic and Atmospheric Administration (NOAA) located in Asheville, North Carolina, with data starting in 1880 [4]; 3) the Goddard Institute of Space Studies of the US National Aeronautics and Space Administration (NA-

SA) located in New York City, with data starting in 1880 [5]; and 4) the Japanese Meteorological Agency (JMA) located in Tsukuba, Japan, with data starting in 1891 [6, 7]. In FP3 we showed that the starting years of datasets 2-4 are too late to properly characterize the structure of the first Quasi-periodic Oscillation, QPO-1, revealed by the earlier starting date of dataset 1. Accordingly, here we restrict attention to the HadCRU temperature dataset alone.

2. Results

Here we present results for the trend in the observations, the three Quasi-periodic Oscillations (QPOs) that are predictable year to year, and the remaining variations that are not predictable year to year. We represent the latter by a Gaussian probability distribution (GPD).

2.1. The Trend

The HadCRU observed global-mean near-surface temperatures from 1850 through 2012 are shown in **Figure 1(a)**. It can be seen that there is considerable information in this record, both with long and short periods. To reveal this more clearly, let's use a Simple Moving Average (SMA) given mathematically by

$$\bar{T}_i^N = \frac{1}{N} \sum_{j=i-M}^{i+M} T_j, i = M+1, \dots, 163-M, M = \frac{N-1}{2} \quad (1)$$

with N an odd number. We restrict N to be odd so that the year at which the SMA is located is an integer. For example, for $N = 21$, the year at which the SMA is located is 11, with 10 data points to both the left and right thereof. If N were 20, then the year of the SMA would be 10.5. Since this is not a year in the dataset, we would have to place the SMA at either 10 or 11. Because we will calculate the Pearson Coefficient of Determination (R^2) of the SMA with results from the SSA, our using an even N would result in slightly different values for the two possible locations for the SMA, and thus two possible different values of R^2 . We avoid this by restricting N to be odd. Of course it is well known that such SMA has deficiencies in terms of its frequency response [8], but we use it here anyway because of its simplicity and, thus, transparency.

The result of using SMA with $N = 21$ is shown by the blue line in **Figure 1(b)**. It can be seen that the short-period fluctuations have been removed by the 21-year SMA, thereby revealing the trend in the observed temperatures plus an oscillation of approximately 60 years duration. We first focus on the trend and then on the os-

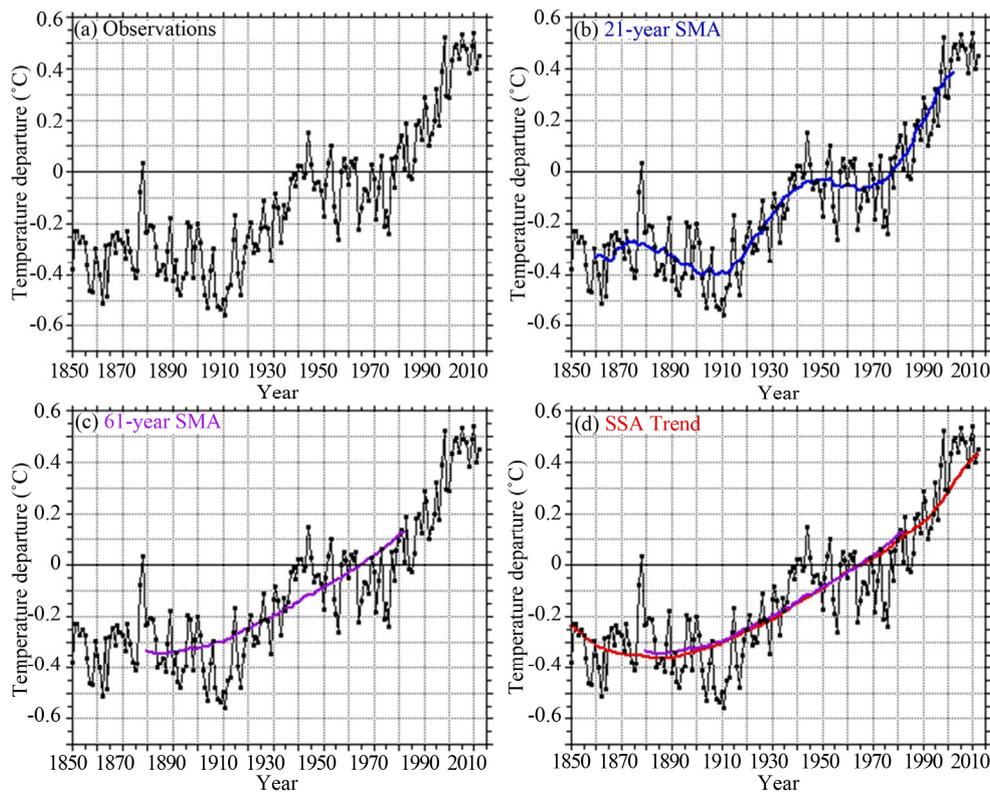


Figure 1. (a) The observed global-mean near-surface temperature departures from the 1961-1990 average; (b) As in (a), but with the 21-year Simple Moving Average shown by the blue line; (c) As in (a), but with a 61-year Simple Moving Average shown by the purple line; (d) As in (c), but with the SSA trend shown by the red line.

cillation. To do this we apply a 61-year SMA to the observations of panel a, that is, we use Equation (1) with $N = 61$ years rather than 21 years. Of course in our so doing we “lose” the first and last 30 years of the observations, just as we “lost” the first and last 10 years of the observations when we used the 21-year SMA. This is another drawback of SMA, a drawback that is not suffered by SSA, as we shall see.

The purple curve in **Figure 1(c)** shows the result of the 61-year SMA. Here all the oscillations have been removed to reveal the trend, albeit only over the $163 - 2 \times 30 = 103$ central years of the record. The red line in **Figure 1(d)** shows the trend revealed by SSA. This trend, which comes from the SSA “black box”, is essentially identical to the trend revealed by the 61-year SMA. In fact we chose the 61-year period by calculating the Pearson Coefficient of Determination, R^2 , of the SMA data with the SSA trend for several odd values of N , and selecting the one that gives the largest R^2 . As shown in **Table 1**, $R^2 = 0.999$ for $N = 61$ years. In other words, the SSA trend is the same as the trend given by the 61-year SMA, but extended backward in time by 30 years to the beginning of the observations in 1850, and forward in time to the latest observation in 2012. Accordingly, there is nothing mysterious about the SSA trend.

As we showed in Causes, the SSA trend is due to hu-

manity—not nature—as a result of our emissions of greenhouse gases (carbon dioxide, methane, nitrous oxide and the chlorofluorocarbons), aerosol precursors (sulfur dioxide, black carbon and organic carbon), and land-use changes (predominantly through deforestation).

2.2. QPO-1

We now investigate the long-period Quasi-periodic Oscillation in the observations. We do so by subtracting the SSA trend shown by the red curve in **Figure 1(d)** from the observed temperature departures shown by the black curve therein. The result is shown in **Figure 2(a)**. These detrended observations reveal not only the long-period QPO, but the shorter period QPOs and stochastic noise as well. Accordingly, we apply a 21-year SMA given by Equation (1) with $N = 21$ to the detrended observations shown in **Figure 2(a)**. The result is the blue curve shown in **Figure 2(b)**, wherein the first and last 10 years of the detrended observations are “lost” due to the 21-year SMA. The long-period variation is now clearly seen. In **Figure 2(c)** the SSA QPO-1 is shown by the red curve. It is seen that SSA QPO-1 is almost identical to the blue curve also shown in **Figure 2(c)**. We chose the length of the SMA to be 21 years because it maximizes the R^2 with QPO-1, yielding a value of 0.991, as shown in **Table 1**. Thus SSA QPO-1 is what one obtains from an SMA of

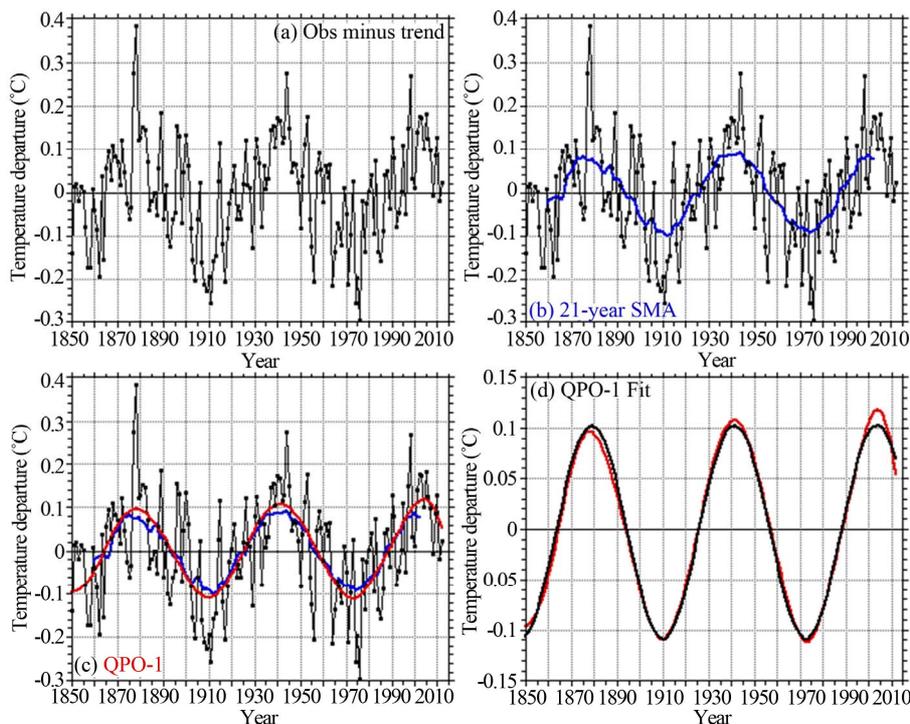


Figure 2. (a) The detrended observed global-mean near-surface temperature departures from the 1961–1990 average; (b) As in (a), but with a 21-year Simple Moving Average shown by the blue line; (c) As in (b), but with QPO-1 shown by the red line; (d) QPO-1 shown by the red line and its fit by $y(t) = C + A \sin [2\pi(t - 1850)/P - \phi]$, with C , A , P and ϕ shown in **Table 1**, shown by the black line.

Table 1. Averaging period of the Simple Moving Average that yields the largest coefficient of determination (R^2) with the SSA trend and QPOs 1, 2 and 3, together with the R^2 thereof. Values of the parameters of the sine-wave representations, $y(t) = C + A \sin [2\pi(t - 1850)/P - \phi]$, and coefficients of determination (R^2), for QPO's 1, 2 and 3 for the HadCRU observed temperature dataset.

Parameter	Trend	QPO-1	QPO-2	QPO-3
Averaging period of SMA that yields the largest R^2 with the trend or QPO/ R^2	61 years/ 0.999	21 years/ 0.991	11 years/ 0.763	3 years/ 0.567
C		-3.2735e-3	-1.2687e-3	-3.7248e-5
A		-0.10597	-0.039488	0.028353
P		62.416	21.043	9.112
ϕ		10.778	16.283	-0.1946
R^2		0.993	0.884	0.703

21-years applied to the detrended observations, except that SSA extends backward to 1850 and forward to 2012. Accordingly, there is nothing mysterious about QPO-1 given by the SSA “black box”. In fact, it is the QPO that we discovered in our 1994 paper “An Oscillation in the Global Climate System of Period 65-70 Years” ([9]; hereafter Discovery) using SSA; it has come to be known as the Atlantic Multidecadal Oscillation, or AMO. QPO-1 is most likely caused by the natural variation of the thermohaline circulation [10].

As shown in **Figure 2(d)**, QPO-1, aka the AMO, is sufficiently regular in period and amplitude that it can be represented quite well by a sine wave, with a Coefficient of Determination $R^2 = 0.993$. The characteristics of this sine-wave fit are shown in **Table 1**. It is seen that the period and amplitude are 62.4 years and 0.11°C, respectively. Accordingly, QPO-1 is the dominant natural variation in the HadCRU observations.

2.3. QPO-2

Subtracting QPO-1 from the detrended temperature observations yields the record shown in **Figure 3(a)**. It is evident that this record contains at least one additional QPO. We can show this by applying Equation (1) to the data with $N = 11$. The result is shown by the blue line in **Figure 3(b)**. Here an irregular oscillation is seen with a period close to 20 years. The result given by SSA is shown by the red curve of **Figure 3(c)**. This is QPO-2. The 11-year length of the SMA was chosen to maximize the Coefficient of Determination with QPO-2, yielding a value of $R^2 = 0.763$, as shown in **Table 1**. Thus the 11-year SMA is a reasonably good indicator of QPO-2.

As for QPO-1, QPO-2 determined by SSA is smoother and more regular than the result given by SMA with $N = 11$. Moreover, SSA defines QPO-2 for all 163 years of the observed temperature record, while SMA loses the first and last 5 years thereof.

QPO-2 was discovered by Ghil and Vautard [11] and also found by us in our Discovery paper. No definitive cause for QPO-2 has yet been found [12].

While QPO-2 is less regular than QPO-1, it too can be fit by a sine wave as shown in **Figure 3(D)** and **Table 1**. It is seen that the period and amplitude are 21.0 years and 0.04°C, respectively, with a Coefficient of Determination $R^2 = 0.884$.

2.4. QPO-3

Subtracting QPO-2 from the detrended temperature observations minus QPO-1 yields the record shown in **Figure 4(a)**. It appears therefrom that there is least one additional QPO therein. We can show this by applying Equation (1) to the data with $N = 3$. The result is shown by the blue line in **Figure 4(b)**. Here an irregular oscillation is seen with a period close to 10 years. The result given by SSA is shown by the red curve of **Figure 4(c)**. This is QPO-3. The 3-year length of the SMA was chosen to maximize the Coefficient of Determination with QPO-3, yielding a value of $R^2 = 0.567$, as shown in **Table 1**. Thus the 3-year SMA is an indicator of QPO-3, but it is not as good an indicator as the SMA with $N = 11$ years is of QPO-2, and it is much worse than the SMA with $N = 61$ years is of QPO-1. This indicates that we have gone as far as we can go in representing the QPOs of the HadCRU temperature record by an SMA. Nevertheless, we can represent QPO-3 by a sine wave as shown in **Figure 4(d)**. As shown in **Table 1**, this yields a period of 9.1 years, an amplitude of 0.03°C and $R^2 = 0.703$.

QPO-3 was discovered by Ghil and Vautard [11] and also found by us in our Discovery paper. As for QPO-2, no definitive cause for QPO-3 has yet been found [12].

2.5. The Unpredictable Natural Variability

Figure 5(a) shows the detrended observations minus QPOs 1, 2 and 3. While there are additional QPOs in these data [12], they are too irregular in amplitude and/or period to be represented by a sine wave. Thus, these QPOs are not predictable on a year-to-year basis. Accordingly we represent them and the additional stochastic noise therein not by a sine wave, but rather by a probability distribution. **Figure 5(b)** shows the Cumulative Distribution Function (CDF) for the data in **Figure 5(a)**. The vertical axis on this plot is the Error Function of the horizontal axis. Since the integral of a Gaussian probability distribution (GPD) is an Error Function, a straight line on this plot reveals a probability distribution that is

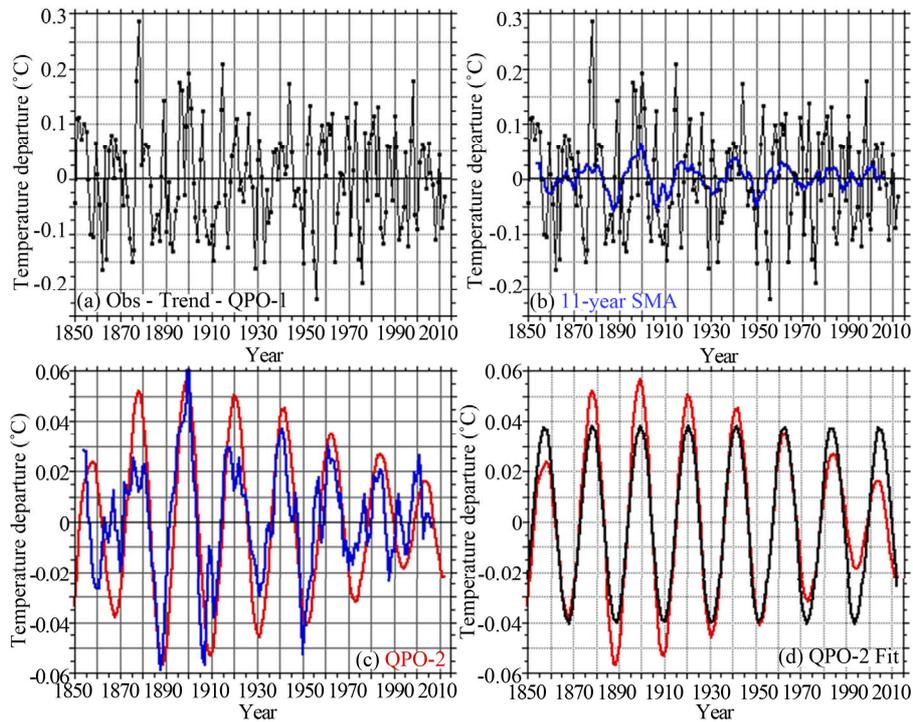


Figure 3. (a) The detrended observed temperature departures minus QPO-1; (b) As in (a), but with an 11-year Simple Moving Average shown by the blue line; (c) The 11-year Simple Moving Average shown by the blue line and QPO-2 shown by the red line; (d) QPO-2 shown by the red line and its fit by $y(t) = C + A \sin [2\pi(t - 1850)/P - \phi]$, with C , A , P and ϕ shown in Table 1, shown by the black line.

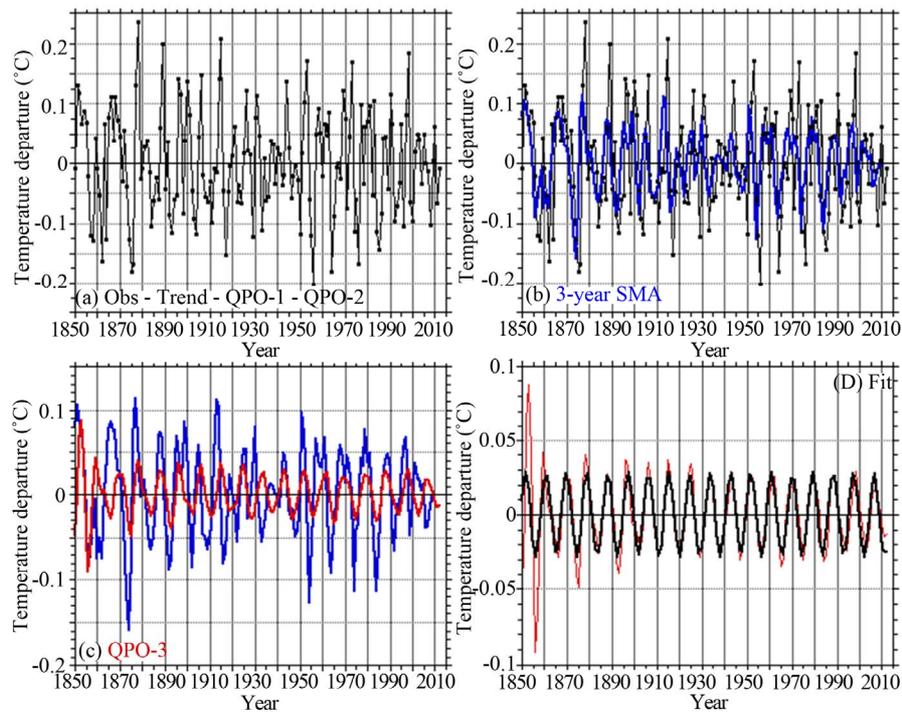


Figure 4. (a) The detrended observed temperature departures minus QPO-1 and QPO-2; (b) As in (a), but with a 3-year Simple Moving Average shown by the blue line; (c) The 3-year Simple Moving Average shown by the blue line and QPO-3 shown by the red line; (d) QPO-3 shown by the red line and its fit by $y(t) = C + A \sin [2\pi(t - 1850)/P - \phi]$, with C , A , P and ϕ shown in Table 1, shown by the black line.

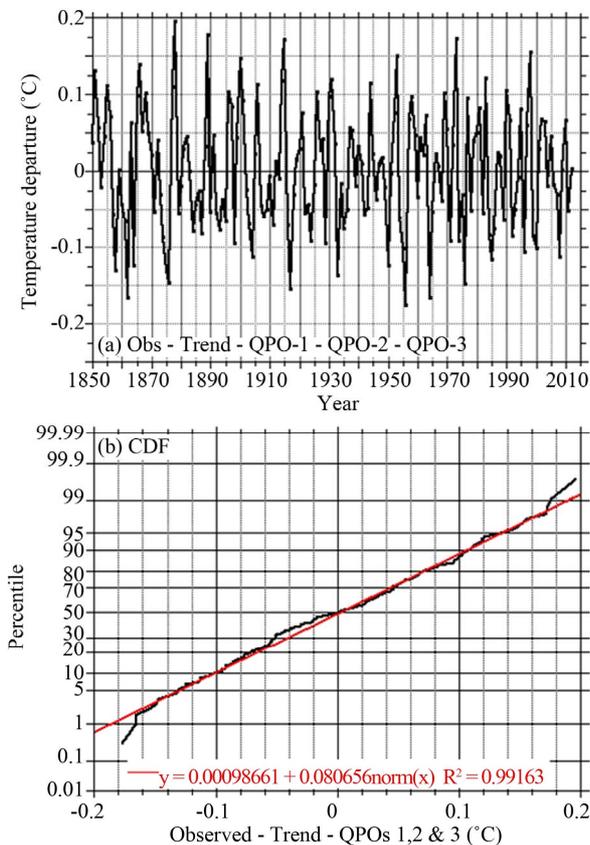


Figure 5. (a) The detrended observed temperature departures minus QPO-1, QPO-2 and QPO-3; (b) The cumulative distribution function of the data in panel a (black line) and its fit by the integral of a Gaussian distribution (Error Function; red line).

Gaussian. **Figure 5(b)** shows a linear fit (red line) to the CDF (black line) in the form of $m + s \text{ norm}(x)$, where m is the mean and s the standard deviation of the linear fit, norm (Gaussian). It is seen that the CDF is very well represented by a GPD with a mean of essentially zero and a standard deviation of 0.08°C . The Coefficient of Determination for this fit is $R^2 = 0.992$, the deviation from unity due to the small deviation of the CDF from the Error Function at the tails of the distribution.

3. Discussion and Conclusion

We have simply and systematically extracted the signals within the HadCRU observations of the global-mean near-surface temperatures from 1850 through 2012 by sequentially taking the Simple Moving Average (SMA) of the data that best represents the results obtained by Singular Spectrum Analysis (SSA), a method of spectrum analysis that is rather opaque to understanding because it is rather like a statistical “black box”.

We used an SMA of 61-years duration to reveal the trend found by SSA. Thus there is nothing mysterious

about this trend—it could have been discovered by SMA alone without the use of SSA. However, the SMA with a 61-year duration “loses” the first and last 30 years of the 163-year temperature record, while SSA “loses” none of this record. Thus SSA can be seen as the optimal form of SMA, one that loses none of the observational record. As we have shown in our 2000 and 2012 Causes papers [1, 13], the trend in the global-mean near-surface temperature record is predominantly due to humanity’s burning of fossil fuels—coal, oil and natural gas—to release the chemical energy therein to run our civilization.

We subtracted the SSA trend from the observations to reveal the natural variability therein. We used an SMA of 21-years duration to reveal the first Quasi-periodic Oscillation (QPO-1) obtained by SSA of the observed temperature record, with a period and amplitude of about 62.4 years and 0.11°C . Accordingly there is nothing mysterious about QPO-1—it could have been found by SMA alone, as we demonstrated here by applying a 21-year SMA to the original non-detrended temperature observations. QPO-1 is the oscillation we discovered in 1994 using SSA on the temperature observations then available [9]. QPO-1 has come to be called the Atlantic Multidecadal Oscillation. The AMO is most likely due to the natural variability of the Thermohaline Circulation [10].

We subtracted QPO-1 from the detrended temperature observations and applied an 11-year SMA thereto. This revealed the second QPO found by SSA, with a period and amplitude of about 21.0 years and 0.04°C . We then subtracted QPO-2 from the detrended observations minus QPO-1 and applied a 3-year SMA to the result. This revealed the third QPO found by SSA, with a period and amplitude of about 9.1 years and 0.03°C . Thus neither QPO-2 nor QPO-3 is mysterious in that they could have been found by a suitable application of SMAs with different temporal lengths. QPOs 2 and 3 are mysterious though because no definitive explanation of their causes has been put forth. Accordingly, it would be useful to apply SSA to simulations of contemporary climate by Atmosphere/Ocean Global Climate Models to determine if they simulate QPOs 2 and 3, and if so, the causes (physics) thereof [12].

Although SSA does find additional QPOs [12], we have not attempted here to find them by SMA. We have not done so because, unlike QPOs 1, 2 and 3, the higher-order QPOs are too irregular in period and/or amplitude to be represented by sine waves, as we have done here for QPOs 1, 2 and 3. Instead, we have represented these higher-order QPOs and the stochastic noise in the observed temperature record by a probability distribution. We have found that a Gaussian probability distribution works very well for this purpose.

Lastly, the results obtained herein can be used to project the natural variability of global-mean near-surface

temperature into the future to add to, and compare with, the continuing temperature trend caused predominantly by the continuing emission of greenhouse gases by humanity. This we have done in our companion paper “A Fair Plan to Safeguard Earth’s Climate: 3. Outlook for Global Temperature Change Throughout the 21st Century [2]”.

REFERENCES

- [1] M. J. Ring, D. Lindner, E. F. Cross and M. E. Schlesinger, “Causes of the Global Warming Observed Since the 19th Century,” *Atmospheric and Climate Sciences*, Vol. 2, No. 3, 2012, pp. 401-415. [doi:10.4236/acs.2012.24035](https://doi.org/10.4236/acs.2012.24035)
- [2] M. E. Schlesinger, D. Lindner, M. J. Ring and E. F. Cross, “A Fair Plan to Safeguard Earth’s Climate. 3: Outlook for Global Temperature Change throughout the 21st Century,” *Journal of Environmental Protection*, Vol. 4, No. 6, 2013, pp. 653-664. [doi:10.4236/jep.2013.46075](https://doi.org/10.4236/jep.2013.46075)
- [3] C. P. Morice, J. J. Kennedy, N. A. Rayner and P. D. Jones, “Quantifying Uncertainties in Global and Regional Temperature Change Using an Ensemble of Observational Estimates: The HadCRUT4 Dataset,” *Journal of Geophysical Research*, Vol. 117, No. D8, 2012, pp. 1984-2012. [doi:10.1029/2011JD017187](https://doi.org/10.1029/2011JD017187)
- [4] T. M. Smith, R. W. Reynolds, T. C. Peterson and J. H. Lawrimore, “Improvements to NOAA’s Historical Merged Land-Ocean Surface Temperature Analysis,” *Journal of Climate*, Vol. 21, No. 10, 2008, pp. 2283-2296. [doi:10.1175/2007JCLI2100.1](https://doi.org/10.1175/2007JCLI2100.1)
- [5] J. Hansen, R. Ruedy, M. Sato and K. Lo, “Global Surface Temperature Change,” *Reviews of Geophysics*, Vol. 48, No. 4, 2010, Article ID: RG4004. [doi:10.1029/2010RG000345](https://doi.org/10.1029/2010RG000345)
- [6] K. Ishihara, “Calculation of Global Surface Temperature Anomalies with COBE-SST,” (*Japanese*) *Weather Service Bulletin*, Vol. 73, 2006, pp. S19-S25.
- [7] K. Ishihara, “Estimation of Standard Errors in Global Average Surface Temperature,” (*Japanese*) *Weather Service Bulletin*, Vol. 74, 2007, pp. 19-26.
- [8] S. W. Smith, “Moving Average Filters,” *The Scientist and Engineer’s Guide to Digital Signal Processing*, 1999, pp. 277-284.
- [9] M. E. Schlesinger and N. Ramankutty, “An Oscillation in the Global Climate System of Period 65 - 70 years,” *Nature*, Vol. 367, No. 6465, 1994, pp. 723-726.
- [10] T. Delworth, S. Manabe, R. S. Stouffer, N. G. Andronova and M. E. Schlesinger, “Interdecadal Variations of the Thermohaline Circulation in a Coupled Ocean-Atmosphere Model,” *Journal of Climate*, Vol. 6, No. 11, 1993, pp. 1993-2011. [doi:10.1175/1520-0442\(1993\)006<1993:IVOTTC>2.0.CO;2](https://doi.org/10.1175/1520-0442(1993)006<1993:IVOTTC>2.0.CO;2)
- [11] M. Ghil and R. Vautard, “Interdecadal Oscillations and the Warming Trend in the Global Temperature Time Series,” *Nature*, Vol. 350, No. 6316, 1991, pp. 324-327. [doi:10.1038/350324a0](https://doi.org/10.1038/350324a0)
- [12] D. Lindner, “Quasi-Periodic Oscillations in Observed and Simulated Temperatures and Implications for the Future,” Ph.D. Thesis, University of Illinois, Urbana-Champaign, 2013, 135p.
- [13] N. G. Andronova and M. E. Schlesinger, “Causes of Temperature Changes during the 19th and 20th Centuries,” *Geophysical Research Letters*, Vol. 27, No. 14, 2000, pp. 2137-2140. [doi:10.1029/2000GL006109](https://doi.org/10.1029/2000GL006109)