LOCAL WARMING

ROBERT J. VANDERBEI

ABSTRACT. Using 55 years of daily average temperatures from a local weather station, I made a least-absolute-deviations regression model that accounts for three effects: seasonal variations, the 11-year solar cycle, and a linear trend. The solution correctly identifies the period and phase of the solar cycle. It also indicates that temperatures have gone up by about 2 °F over the 55 years covered by the data. A similar least-squares model failed to characterize the solar cycle correctly. The paper has two goals: (1) to demonstrate that median statistics and their generalizations provide a better statistical approach to analyzing climate data and (2) to show how easy it is to analyze at least some climate data for oneself so one can draw his or her own conclusions.

1. INTRODUCTION.

Most research on climate change aims to produce a high-fidelity model of climate that spans centuries [13] if not millennia [11]. These models must resort to proxy climate indicators as direct measurements only go back a century or so (see, e.g., [8]). In this paper, actual temperature readings from a single undisturbed location spanning a time horizon of 55 years are analyzed using a *least-absolute deviations (LAD)* regression model that robustly extracts a small linear trend from the much larger seasonal variations. An analogous least-squares regression model generates results that are much less believable than those obtained with the LAD model.

Least absolute deviations regression [1] belongs to a class of statistical techniques called *robust statistics* [10]. The sample median is the simplest and most widely used example of a robust statistic. Sample medians have played an important role in a wide range of scientific fields including astrophysics [7], medicine [3], and signal processing [2] to name a few.

The purpose of this paper is not to attempt to improve on any of the global warming estimates that exist in the literature. Rather, its purpose is to bring a new alternative statistical approach to these analyses that should allow one to extract meaningful climate information from noisy data in a more robust and reliable manner than before.

Key words and phrases. Global warming, climate change, local warming, solar cycle, least absolute deviations, regression, least squares, median, percentile, quantile, confidence interval, nonparametric statistics, linear programming, parametric simplex method.

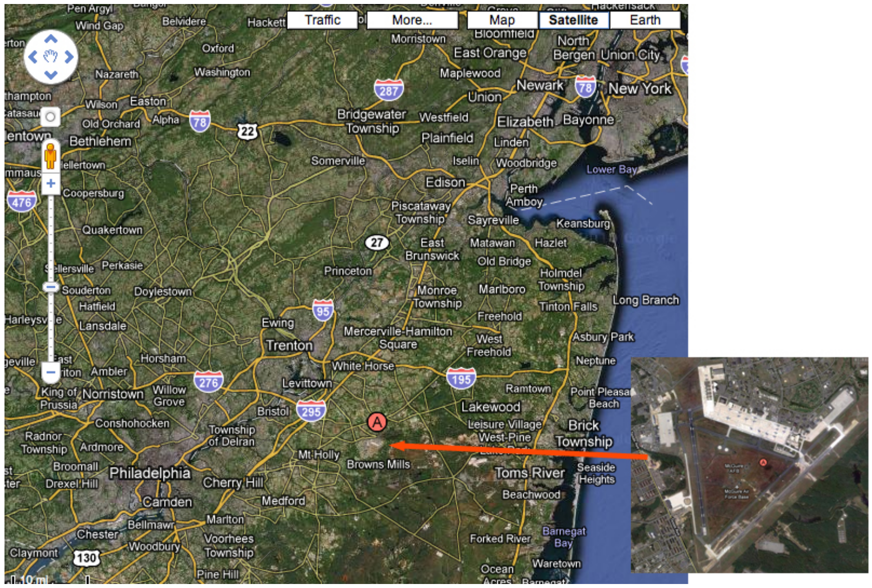

FIGURE 1. McGuire Air Force Base

2. THE DATA

The *National Oceanic and Atmospheric Administration* (NOAA) collects and archives weather data from thousands of collection sites around the globe. The data format and instructions for downloading the data can be found on the NOAA website [14] as can a list of the roughly 9000 weather stations [15]. McGuire Air Force Base, located not far from Princeton NJ, is one of the archived weather stations. This particular weather station seemed good for a number of reasons. First, it is about 50 miles from New York City and 30 miles from Philadelphia. It is in a rather undeveloped part of the state. It was established in 1937 and has been a major airbase since 1942. Finally, it is only about 25 miles from the Atlantic Ocean and therefore its climate should be moderated somewhat by its proximity to an ocean.

3. THE MODEL

Let T_d denote the average temperature in degrees Fahrenheit on day $d \in D$ where D is the set of days from January 1, 1955, to August 13, 2010 (that's 20,309 days).

We wish to model the average temperature as a *constant* x_0 plus a *linear trend* $x_1 d$ plus a sinusoidal function with a one-year period representing *seasonal changes*,

$$x_2 \cos(2\pi d/365.25) + x_3 \sin(2\pi d/365.25),$$

plus a sinusoidal function with a period of 10.7 years to represent the *solar cycle*,

$$x_4 \cos(2\pi d/(10.7 \times 365.25)) + x_5 \sin(2\pi d/(10.7 \times 365.25)).$$

The parameters $x_0, x_1, \ldots, x_5$ are unknown regression coefficients. We wish to find the values of these parameters that minimize the sum of the absolute deviations:

$$\begin{aligned}\min_{x_0,\ldots,x_5} \sum_{d \in D} \quad & \left| x_0 + x_1 d \right. \\ & +x_2 \cos(2\pi d/365.25) + x_3 \sin(2\pi d/365.25) \\ & +x_4 \cos(2\pi d/(10.7 \times 365.25)) + x_5 \sin(2\pi d/(10.7 \times 365.25)) \\ & \left. -T_d \right| .\end{aligned}$$

We use the usual trick of introducing a new variable for each absolute value term and then adding a pair of constraints that say that this new variable dominates the expression that was inside the absolute values and its negative ([17], Chapter 12). The linear programming problem, expressed in the AMPL modeling language [6], is shown in Figure 2. AMPL models, with their associated user-supplied data sets, can be solved online using the *Network Enabled Optimzation Server (NEOS)* at Argonne National Labs [4].

A *least-absolute-deviations* (LAD) model was chosen instead of a least-squares model because LAD regression, like the median statistic, is insensitive to "outliers" in the data.

4. The Results

The linear programming problem can be solved in only a few minutes on a modern laptop computer. The optimal values of the parameters are

```
set DATES ordered;
param avg {DATES};
param day {DATES};
param pi := 4*atan(1);

var x {j in 0..5};
var dev {DATES} >= 0, := 1;

minimize sumdev: sum {d in DATES} dev[d];
subject to def_pos_dev {d in DATES}:
    x[0] + x[1]*day[d] + x[2]*cos( 2*pi*day[d]/365.25)
                       + x[3]*sin( 2*pi*day[d]/365.25)
                       + x[4]*cos( 2*pi*day[d]/(10.7*365.25))
                       + x[5]*sin( 2*pi*day[d]/(10.7*365.25))
        - avg[d]
    <= dev[d];
subject to def_neg_dev {d in DATES}:
    -dev[d] <=
    x[0] + x[1]*day[d] + x[2]*cos( 2*pi*day[d]/365.25)
                       + x[3]*sin( 2*pi*day[d]/365.25)
                       + x[4]*cos( 2*pi*day[d]/(10.7*365.25))
                       + x[5]*sin( 2*pi*day[d]/(10.7*365.25))
        - avg[d];

data;

set DATES := include "data/Dates.dat";
param: avg := include "data/McGuireAFB.dat";
let {d in DATES} day[d] := ord(d,DATES);

solve;
```

FIGURE 2. The model expressed in the AMPL modeling language.

$$
\begin{aligned}
x_0 &= 52.6\ {}^\circ\mathrm{F} \\
x_1 &= 9.95 \times 10^{-5}\ {}^\circ\mathrm{F/day} \\
x_2 &= -20.4\ {}^\circ\mathrm{F} \\
x_3 &= -8.31\ {}^\circ\mathrm{F} \\
x_4 &= -0.197\ {}^\circ\mathrm{F} \\
x_5 &= 0.211\ {}^\circ\mathrm{F}
\end{aligned}
$$

4.1. **Linear Trend.** From x_0, we see that the nominal temperature at McGuire AFB was 52.56 °F (on January 1, 1955).

We also see, from x_1, that there is a positive trend of 0.000099 °F/day. That scales to 3.63 °F per century. This result agrees rather well with results from global climate change models, which predict a per century warming rate of between 2.0 °C and 2.4 °C ([12, 9]).

4.2. **Magnitude of the Sinusoidal Fluctuations.** From the sine and cosine terms x_2 and x_3, we can compute the amplitude of annual seasonal changes in temperatures:

$$\sqrt{x_2^2 + x_3^2} = 22.02 \text{ °F}.$$

In other words, on the hottest summer day we should expect the temperature to be 22.02 degrees warmer than the nominal value of 52.56 degrees; that is, 77.58 degrees. Of course, this is a daily average—daytime highs can be expected to be higher and nighttime lows lower by about the same amount.

Similarly, from the x_4 and x_5 sine and cosine terms, we can compute the amplitude of the temperature changes brought about by the solar-cycle:

$$\sqrt{x_4^2 + x_5^2} = 0.2887 \text{ °F}.$$

The effect of the *solar cycle* is real but relatively small.

4.3. **Phase of the Sinusoidal Fluctuations.** Close inspection of the output shows that January 22 is nominally the coldest day in the winter and July 24 is the hottest day of summer. It is perhaps worth noting that the coldest days in the winter of 2011 turned out to be January 23 and 24.

According to the LAD model, February 12, 2007, was the day of the last minimum in the 10.78 year solar cycle. It is well-known that the solar cycle had its last minimum in 2007 [16]. The correct extraction of the phase (and, in a later section, the period) of the solar cycle, which is a small effect having an amplitude of only 0.29 °F, is strong support for fidelity of the LAD model.

4.4. **Visualizing the Results.** Figure 3 shows a plot of all 20,309 data points. Overlaid on these data points is the solution of the LAD regression model. Seasonal fluctuations completely dominate other effects. It is impossible to “see” any linear warming trend or the solar cycle.

Figure 4 has the seasonal and solar-cycle variations removed. Even this plot is noisy. Of course, there are many days in a year and some days are unseasonably warm while others are unseasonably cool. It is not uncommon for there to be an “unseasonably warm” day that is 20 or even sometimes 30 degrees above seasonally adjusted averages.

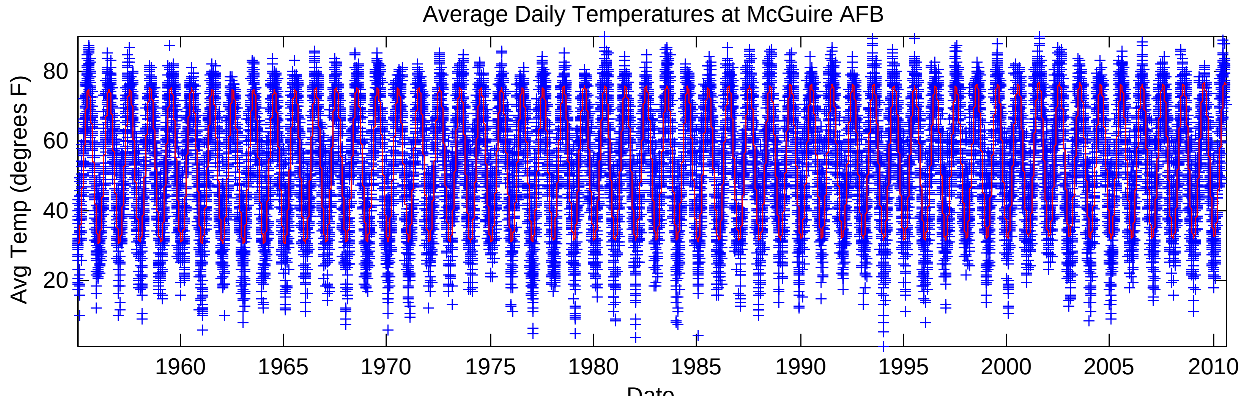

FIGURE 3. Plot Showing Actual Data and Regression Curve. *Blue:* Average daily temperatures at McGuire AFB from 1955 to 2010. *Red:* Output from least absolute deviation regression model.

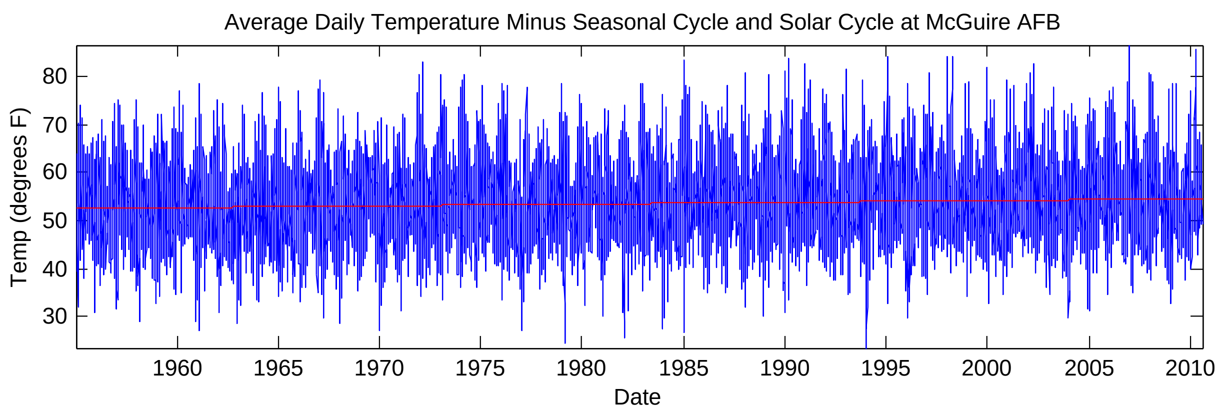

FIGURE 4. As before but with sinusoidal seasonal variation removed and sinusoidal solar-cycle variation removed as well.

Figure 5 is derived from Figure 4 by applying a 101 day rolling average to each data point. The rolling average reduced the extreme fluctuations to about $1/10$-th their original amplitude thus making the linear warming trend quite apparent. In NJ we have *local warming*.

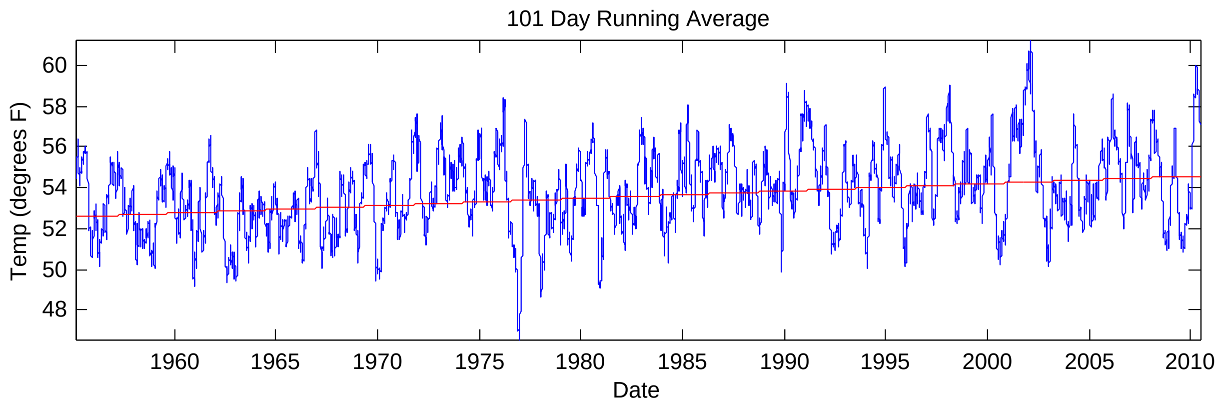

FIGURE 5. Smoothed Seasonally Subtracted Plot. To smooth out high frequency fluctuations, we use 101 day rolling averages of the data.

5. MEANS, MEDIANS, AND OPTIMIZATION

Point estimates of the regression coefficients without some indication of a confidence interval for these estimates are almost useless. However, with least absolution deviation regression the notion of a confidence interval is a bit subtle. As it is related to the confidence intervals for simple median statistics, we digress briefly in this section to review for the reader how one determines confidence intervals for sample medians (see, e.g., Chapter 10 of [**?**] for more details).

Let $b_1, b_2, \ldots, b_n$ denote a set of measurements. Solving

$$\text{argmin}_x \sum_i (x - b_i)^2$$

computes the *mean* of the b_i's, whereas solving

$$\text{argmin}_x \sum_i |x - b_i|$$

computes their *median*, which is also called the 50-th percentile. The $100p$*-th percentile* $(0 \le p \le 1)$ can be computed by solving the following optimization problem:

$$\text{argmin}_x \sum_i \left(|x - b_i| + (1 - 2p)(x - b_i) \right). \tag{1}$$

To see that this is correct, let $f(x)$ denote the function being minimized. To find its minimum, we set the derivative to zero:

$$f'(x) = \sum_i \text{sgn}(x - b_i) + n(1 - 2p) = 0, \tag{2}$$

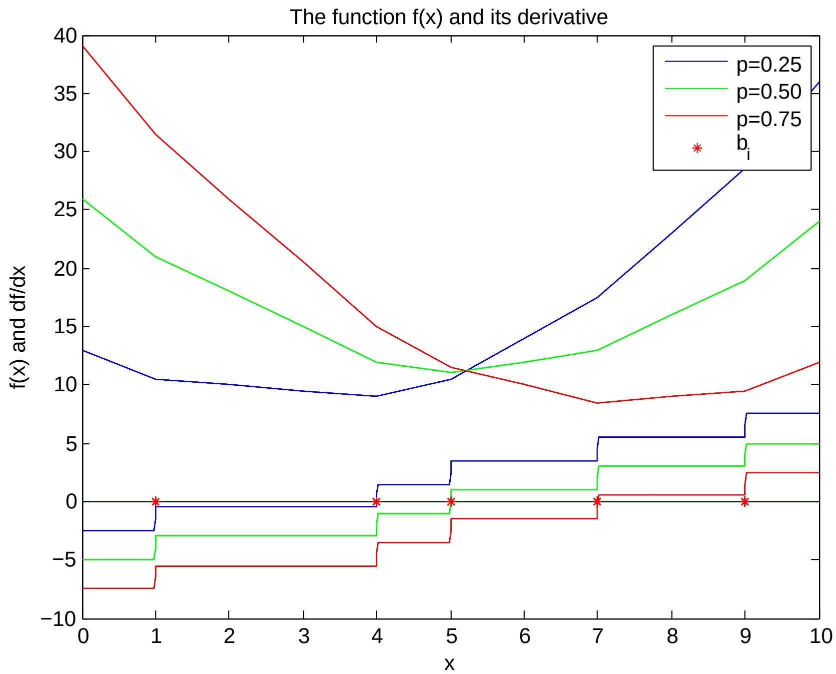

FIGURE 6. An example. Here we plot the function $f(x)$ being minimized in (1) and its derivative given in (2) for three different values of p. The raw data are the b_i's. There are 5 of them plotted along the x-axis. Changing p causes the function $f'(x)$ to slide up or down thereby changing where it crosses zero.

where $\mathrm{sgn}(x)$ is the function that is 1 if $x > 0$, -1 if $x < 0$, and (by convention) 0 if $x = 0$ (see Figure 6). Of course, the absolute value function is, strictly speaking, not differentiable at the origin but this detail turns out not to matter—we leave it to the reader to ponder the subtleties. The summation computes the number of data points to the left of x minus the number to the right. But, since the number to the left is n minus the number to the right, the summation can also be written as $2k(x) - n$, where $k(x)$ denotes the number of data points smaller than x. Therefore, equation (2) reduces to

$$k(x) = np,$$

which says that x is a point at which np data points lie to the left, i.e., x is the $100p$-th percentile.

Assume now that $b_1, b_2, \dots, b_n$ are specific values of a sequence of independent identically distributed random variables $B_1, B_2, B_3, \dots, B_n$. Let m denote their common unknown median and let

$$B_{(1)} < B_{(2)} < B_{(3)} < \cdots < B_{(n)}$$

denote the *order statistics*, i.e., the original variables rearranged into increasing order. Note: $B_{(k)}$ is the $100(k/n)$-th *sample percentile*. Then,

$$\begin{aligned}\mathbf{P}(B_{(k)} \leq m \leq B_{(k+1)}) &= \mathbf{P}(B_j \leq m \text{ for } k \text{ indices and} \\ &\qquad B_j \geq m \text{ for the remaining } n-k \text{ indices}) \\ &= \binom{n}{k}\left(\frac{1}{2}\right)^n.\end{aligned}$$

Hence,

$$\mathbf{P}(B_{(k)} \leq m \leq B_{(n-k+1)}) = \sum_{j=k}^{n-k}\binom{n}{j}\left(\frac{1}{2}\right)^n.$$

For any given n, it is easy to choose k so that

$$\sum_{j=k}^{n-k}\binom{n}{j}\left(\frac{1}{2}\right)^n \approx 0.95.$$

A 95% confidence interval for the median is the interval $[B_{(k)}, B_{(n-k)}]$. Given actual data, such as the b_i's, we can sort the numbers and then pick the k-th and $(n-k)$-th value from the sorted list to give a confidence interval based on actual data. As an alternative to sorting the data, we could compute the sample confidence interval by computing the $p_{\min}$-th and $p_{\max}$-th percentiles obtained by solving (1) with $p = p_{\min} = k/n$ and with $p = p_{\max} = 1 - p_{\min}$.

For large n, we can use the Gaussian distribution as an approximation to the binomial distribution. In this case, it is convenient to use an interval around $p = 1/2$, with half-width of two standard deviations, which gives roughly a 95% confidence interval. Hence, for this we choose:

$$p_{\min} = \frac{1}{2} - \frac{1}{\sqrt{n}} \quad \text{and} \quad p_{\max} = \frac{1}{2} + \frac{1}{\sqrt{n}}. \tag{3}$$

6. CONFIDENCE INTERVALS FOR LEAST ABSOLUTE DEVIATION REGRESSION

Before returning to our six-parameter local warming model, let us first look at an example of a simple two-parameter least absolute deviation model.

6.1. A Simple 2-D Example.

Suppose that we have n pairs of measurements (a_i, b_i), $i = 1, 2, \ldots, n$, (to connect with our warming model, we can think of the a_i's as days and each corresponding b_i as the average temperature on day a_i) and we posit that there is an affine relationship between the pairs:

$$b_i = x_1 + x_2 a_i + \varepsilon_i,$$

where the ε_i's are independent, identically distributed, and have median zero. We don't know the coefficients x_1 and x_2 and wish to find an estimator and an associated confidence "interval" for these two parameters.

For medians, confidence intervals are given by percentiles and can be computed by solving an optimization problem. The analogous optimization problem for this regression model is:

$$\min_{x_1,x_2} \sum_i \left(|x_1 + x_2 a_i - b_i| + (1-2p)\left(x_1 + x_2 a_i - b_i\right)\right). \tag{4}$$

Using a standard trick in optimization, it is easy to convert this problem into a linear programming problem by introducing new variables δ_i, $i = 1, \dots, n$:

$$\begin{array}{lll} \text{minimize} & \sum_i \left(\delta_i + (1-2p)\left(x_1 + x_2 a_i - b_i\right)\right) & \\ \text{subject to} & \quad\ x_1 + x_2 a_i - b_i \le \delta_i & i = 1, \dots, n \\ & -\delta_i \le x_1 + x_2 a_i - b_i & i = 1, \dots, n. \end{array}$$

Using the simplex method, it is straight-forward to find the pair (x_1^*, x_2^*) that achieves the minimum for any given p, say $p = 1/2$. Better yet, using the parametric simplex method (see e.g. Chapter 7 in [17]) with p as the "parameter", one can solve this problem for every value of p in about the same time as the standard simplex method solves one instance of the problem. Starting at $p = 1$ and sequentially pivoting toward $p = 0$, the parametric simplex method gives a set of thresholds $1 = p_0 \ge p_1 \ge p_2 \ge \cdots \ge p_K = 0$, at which the optimal solution changes. In other words, over any interval, say $p \in [p_k, p_{k-1}]$, there is a certain fixed optimal solution, call it $(x_1^{(k)}, x_2^{(k)})$. At the intersection of two intervals, say $[p_{k+1}, p_k]$ and $[p_k, p_{k-1}]$, both solutions $(x_1^{(k+1)}, x_2^{(k+1)})$ and $(x_1^{(k)}, x_2^{(k)})$ are optimal as are all convex combinations of these two solutions.

Let $f(x_1, x_2)$ denote the objective function in (4). We can minimize this function by finding x_1 and x_2 so that zero belongs to the subgradient of f. Taking the subdifferential with respect to x_1, we get

$$\begin{aligned} \frac{\partial f}{\partial x_1} &= \sum_{i=1}^n \left(\operatorname{sgn}(x_1 + x_2 a_i - b_i) + (1-2p)\right) \\ &= \sum_{i=1}^n 2\left(1_{x_1+x_2a_i>b_i} + [0,1]_{x_1+x_2a_i=b_i} - p\right). \end{aligned} \tag{5}$$

Hence, we require that

$$np \in \sum_{i=1}^n \left(1_{x_1+x_2a_i>b_i} + [0,1]_{x_1+x_2a_i=b_i}\right). \tag{6}$$

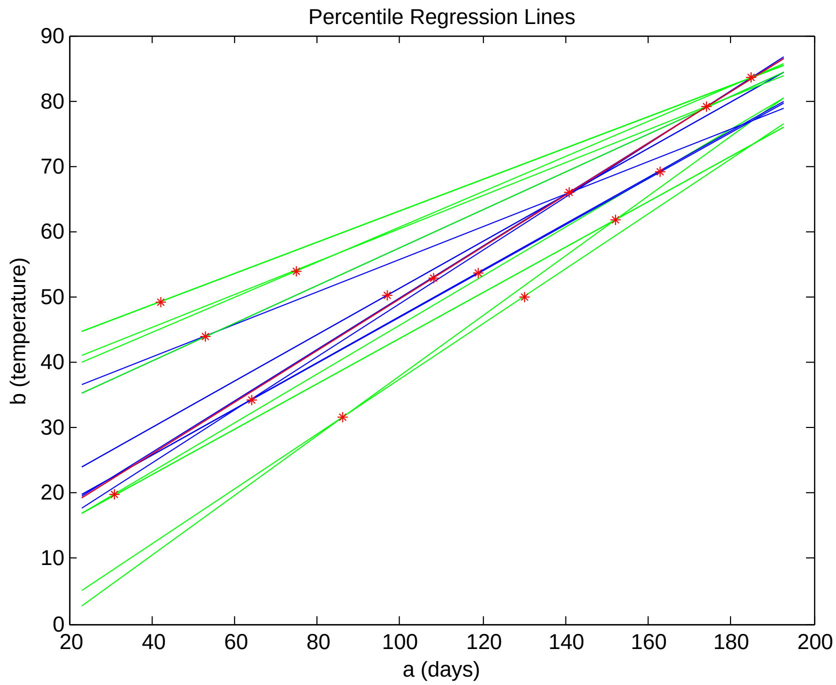

FIGURE 7. Fifteen pairs of points, shown as red stars, and all of the regression lines associated with different intervals of p-values from $p = 1$ at the top to $p = 0$ at the bottom. The line associated with the interval that covers $p = 1/2$ is red and the lines within the confidence interval, computed using all p values between $p_{\min}$ and $p_{\max}$ in (3), are shown in blue.

Similarly, taking the subgradient of f with respect to x_2 and requiring it to contain zero reduces to

$$p \sum_{i=1}^{n} a_i \in \sum_{i=1}^{n} a_i \left(1_{x_1+x_2 a_i > b_i} + [0,1]_{x_1+x_2 a_i = b_i}\right). \tag{7}$$

Figure 7 shows an example consisting of $n = 15$ data points. Also shown are the regression lines associated with each of the particular pairs of regression coefficients $(x_1^{(k)}, x_2^{(k)})$. Note that, unlike with order statistics, the successive regression lines as p changes are not *ordered*. Hence, it does not suffice to specify only the two regression lines corresponding to the extreme values of p. Rather, we must include all of the intermediate regression lines in what might be called a *confidence curve*.

6.2. **Back To The Full 6-D Regression Model.** We can compute a 6-dimensional confidence curve for the six regression coefficients in our local warming regression model. Figures 8 and 9 show a few 2-dimensional projections of this curve. Any one-dimensional projection of the 6-dimensional confidence curve defines a *confidence interval* for the associated quantity. The 95%

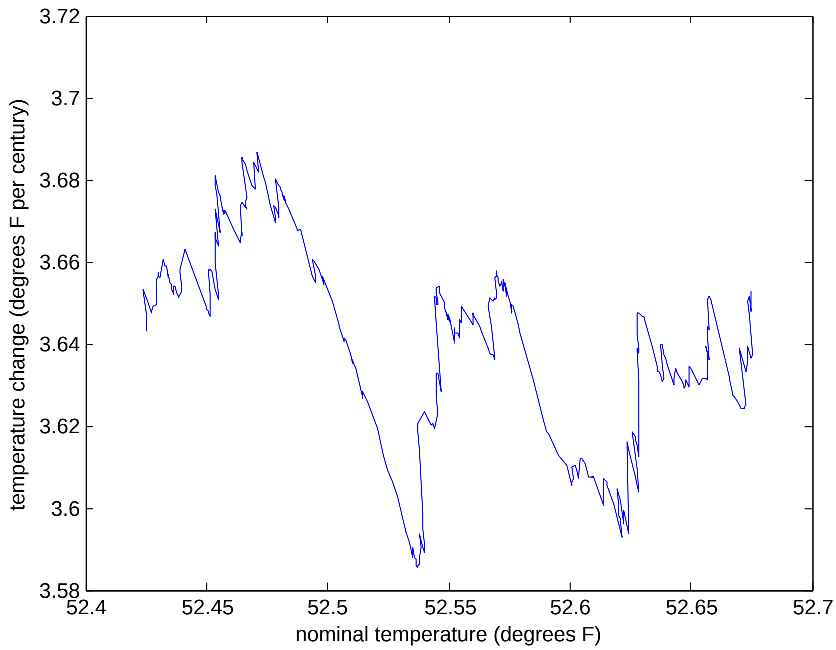

FIGURE 8. Plus/minus two-sigma confidence curve for the nominal temperature, x_0, and the rate of temperature change, x_1.

confidence interval for x_1 is $[3.588\ °\text{F},\ 3.687\ °\text{F}]/100$ yrs. In Figure 8, the projection of the curve onto the vertical axis gives this interval. Note that the confidence interval is much wider than what one would deduce from looking just at the values associated with $p_{\min}$ and $p_{\max}$.

7. ESTIMATING THE PERIOD OF THE SOLAR CYCLE

The length of the solar cycle is only approximately 10.7 years [18]. We can modify the model to predict, in addition to the parameters already being estimated, the period of this cycle. To do this, we introduce one new variable x_6 and change the solar-cycle sine and cosine terms to read:

$$+x_4 \cos(2x_6\pi d/(10.7 \times 365.25)) + x_5 \sin(2x_6\pi d/(10.7 \times 365.25)).$$

If the unknown parameter x_6 is *fixed at* 1, forcing the solar-cycle to have a period of exactly 10.7 years, then the problem reduces to the linear programming problem considered earlier. If, on the other hand, we allow x_6 to vary, then the problem is *nonlinear* and even *nonconvex* and therefore harder in principle. However, most nonlinear (local) optimization algorithms work well if the variables are given initial values fairly close to the correct optimal value. Such is the case with the problem at hand. The result is that the first six parameters remain virtually unchanged and $x_6 = 0.992$ which translates to a 10.8 year solar cycle, which is in close agreement with the nominal value of 10.7 years.

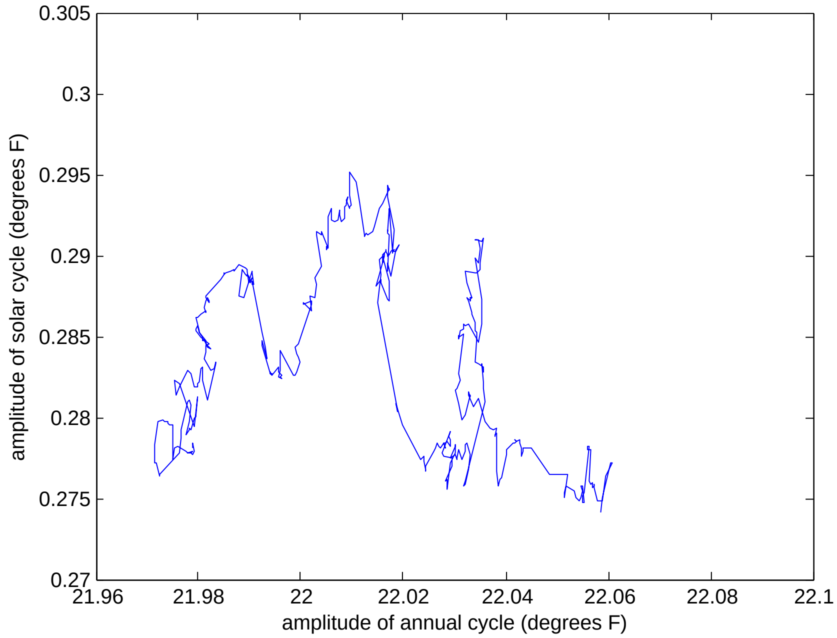

FIGURE 9. Plus/minus two-sigma confidence curve for the amplitude of the seasonal cycle, $\sqrt{x_2^2 + x_3^2}$, and the amplitude of the solar cycle, $\sqrt{x_4^2 + x_5^2}$.

8. LEAST SQUARES SOLUTION (MEAN INSTEAD OF MEDIAN)

Suppose we change the objective to a sum of squares of deviations:

```
minimize sumdev: sum {d in DATES} dev[d]^2;
```

The resulting model is a *least squares model.* The objective function is now convex and quadratic and the problem is still easy to solve. The solution, however, is *sensitive* to outliers. Here's the output:

$$
\begin{aligned}
x_0 &= \quad 52.6\ ^\circ\text{F} \\
x_1 &= \quad 1.2 \times 10^{-4}\ ^\circ\text{F/day} \\
x_2 &= -20.3\ ^\circ\text{F} \\
x_3 &= \quad -7.97\ ^\circ\text{F} \\
x_4 &= \quad 0.275\ ^\circ\text{F} \\
x_5 &= \quad 0.454\ ^\circ\text{F} \\
x_6 &= \quad 0.730
\end{aligned}
$$

In this case, the rate of local warming is 4.37 °F per century. This number is significantly higher than the corresponding value from the LAD model—it lies outside the confidence interval we computed. The model also produces a *wrong answer* for the period of the solar cycle: $10.7/0.730 = 14.7$ years. It is very well known that the solar cycle is close to 10.7 years.

9. Final Remarks

It is remarkable that the solar cycle can be seen and a warming trend can be extracted from just one weather station's 55-year dataset.

We made one assumption in our model that is wrong. We assumed that the temperature fluctuations from day to day are independent. In reality they are correlated over short time intervals—the correlation length is probably about a week or so. This error makes our confidence intervals too tight. If we know, or can guess, the correlation length m, then we can determine a smaller effective number of points

$$n_{\text{eff}} = n/m$$

and use this in our computation of $p_{\min}$ and $p_{\max}$. For example, if we assume that the correlation length is 9 days, then the new confidence interval for x_1 is $[3.541\ ^\circ\text{F},\ 3.802\ ^\circ\text{F}]/100$ yrs.

The original model can be improved in a few key ways. First of all, the assumption that the seasonal variation is sinusoidal is only an approximation and it completely falls apart for any data collection site in the tropics where, in principle, each year has two dates at which the Sun passes directly overhead and two dates in between when the Sun is furthest (to the north/south) from passing overhead. Also, the linear trend could be a modeled as a function of global (or local) population density. Over 55 years, such a function is probably fairly, but certainly not exactly, linear. Hence, the model we have considered perhaps could be improved. We leave these improvements for future endeavors. Finally, the model could be applied at thousands of local weather stations around the world to produce a global "warming map".

9.1. Getting the Data.

Since the NOAA data is archived in one year batches, I wrote a UNIX shell script to grab the 55 annual data files for McGuire and then assemble the relevant pieces of data into a single file. Here is the shell script:

http://www.princeton.edu/∼rvdb/ampl/nlmodels/LocalWarming/McGuireAFB/data/getData.sh

The resulting pair of data files that I used as input to my local climate model are posted at:

http://www.princeton.edu/∼rvdb/ampl/nlmodels/LocalWarming/McGuireAFB/data/McGuireAFB.dat

and

http://www.princeton.edu/∼rvdb/ampl/nlmodels/LocalWarming/McGuireAFB/data/Dates.dat

References

[1] P. Bloomfield and W.L. Steiger. *Least Absolute Deviations: Theory, Applications, and Algorithms*. Birkhäuser, Boston, 1983.

[2] A. Bovik, T. Huang, and D. Munson Jr. A generalization of median filtering using linear combinations of order statistics. *IEEE Proceedings on Acoustics, Speech and Signal Processing*, 31(6):1342–1350, 1983.

[3] M.J. Campbell and M.J Gardner. Statistics in medicine: Calculating confidence intervals for some nonparametric analyses. *British J. of Medicine*, 296:1454, 1988.

[4] J. Czyzyk, M.P. Mesnier, and J.J. More. The NEOS server. *Computational Science and Engineering*, 5(3):68–75, 1998.

[5] B. Efron. Bootstrap Methods: Another Look at the Jackknife. *Ann. Stat.*, 7(1):1–26, 1979.

[6] R. Fourer, D.M. Gay, and B.W. Kernighan. *AMPL: A Modeling Language for Mathematical Programming*. Scientific Press, 1993.

[7] J.R. Gott, M.S. Vegeley, S. Podariu, and B. Ratra. Median statistics, h_0, and the accelerating universe. *Astrophysical Journal*, 549(1):1–17, 2001.

[8] G.C. Hegerl, H. von Storch, K. Hasselmann, B.D. Santer, U. Cubasch, and P.D. Jones. Detecting greenhouse-gas-induced climate change with an optimal fingerprint method. *Journal of Climate*, 9(10):2281–2306, 1996.

[9] J.T. Houghton, L.G. Meira Filho, D.J. Griggs, and K. Maskell. An introduction to simple climate models used in the IPCC second assessment report. *IPCC Technical Paper II*, 1997.

[10] P.J. Huber and E.M. Ronchetti. *Robust Statistics*. Wiley, 2nd edition, 2009.

[11] P.D. Jones, K.R. Briffa, T.P. Barnett, and S.F.B. Tett. High-resolution palaeoclimatic records for the last millennium: interpretation, integration and comparison with general circulation model control-run temperatures. *The Holocene*, 8(4):455–471, 1998.

[12] T.R. Karl, R.W. Knight, and B. Baker. The record breaking global temperatures of 1997 and 1998: Evidence for an increase in the rate of global warming? *Geophysical Research Letters*, 27(5):719–722, 2000.

[13] M.E. Mann, R.S. Bradley, and M.K. Hughes. Global-scale temperature patterns and climate forcing over the past six centuries. *Nature*, 392:779–787, 1998.

[14] NOAA. Climate data format and download instruction, 2011. ftp://ftp.ncdc.noaa.gov/pub/data/gsod/readme.txt.

[15] NOAA. List of weather stations, 2011. ftp://ftp.ncdc.noaa.gov/pub/data/gsod/ish-history.txt.

[16] N. Scafetta and R.C. Willson. ACRIM-gap and TSI trend issue resolved using a surface magnetic flux TSI proxy model. *Geophysical Research Letters*, 36, 1997.

[17] R.J. Vanderbei. *Linear Programming: Foundations and Extensions*. Springer, 3rd edition, 2007.
[18] R.C. Willson and H.S. Hudson. The sun's luminosity over a complete solar cycle. *Nature*, 351:42–44, 1991.

Acknowledgments. The author thanks Kurt Anstreicher, Jianqing Fan, J. Richard Gott, Matthew Saltzman, and Henry Wolkowicz for helpful discussions about this work.

DEPARTMENT OF OPERATIONS RESEARCH AND FINANCIAL ENGINEERING, PRINCETON UNIVERSITY, PRINCETON, NJ 08544.

E-mail address: `rvdb@princeton.edu`