

# Local Warming\*

Robert J. Vanderbei†

**Abstract.** Using 55 years of daily average temperatures from a local weather station, I made a least-absolute-deviations (LAD) regression model that accounts for three effects: seasonal variations, the 11-year solar cycle, and a linear trend. The model was formulated as a linear programming problem and solved using widely available optimization software. The solution indicates that temperatures have gone up by about 2°F over the 55 years covered by the data. It also correctly identifies the known phase of the solar cycle; i.e., the date of the last solar minimum. It turns out that the maximum slope of the solar cycle sinusoid in the regression model is about the same size as the slope produced by the linear trend. The fact that the solar cycle was correctly extracted by the model is a strong indicator that effects of this size, in particular the slope of the linear trend, can be accurately determined from the 55 years of data analyzed. The main purpose for doing this analysis is to demonstrate that it is easy to find and analyze archived temperature data for oneself. In particular, this problem makes a good class project for upper-level undergraduate courses in optimization or in statistics. It is worth noting that a similar least-squares model failed to characterize the solar cycle correctly, and hence even though it too indicates that temperatures have been rising locally, one can be less confident in this result. The paper ends with a section presenting similar results from a few thousand sites distributed worldwide, some results from a modification of the model that includes both temperature and humidity, as well as a number of suggestions for future work and/or ideas for enhancements that could be used as classroom projects.

**Key words.** global warming, climate change, local warming, solar cycle, least absolute deviations, regression, least squares, median, linear programming

**AMS subject classifications.** 62J05, 62F40, 62G08, 62P12, 86A10, 90C05, 97M10

**DOI.** 10.1137/110827296

**1. Introduction.** Most research on climate change aims to produce a high-fidelity model of climate that spans centuries [13], if not millennia [11]. Since directly observed temperature data are not available over these time scales, such models are forced to resort to proxy climate indicators (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 less reliable than those obtained with the LAD model.

The purpose of this paper is not to attempt to improve on any of the global warming estimates that exist in the literature. Rather, the main purpose of this paper is to show that it is fairly easy to find and analyze archived temperature data with the hope that many others will make similar analyses for places that are of

\*Received by the editors March 14, 2011; accepted for publication (in revised form) November 14, 2011; published electronically August 7, 2012.

<http://www.siam.org/journals/sirev/54-3/82729.html>

†Department of Operations Research and Financial Engineering, Princeton University, Princeton, NJ 08544 (rvdb@princeton.edu).

interest to them. I also hope to inspire educators who teach courses in optimization and/or statistics to develop classroom projects based on the ideas presented here.

LAD 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. A secondary purpose of this paper is to demonstrate that, at least for the data considered here, LAD regression provides better results than the corresponding least-squares regression.

**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 [15], as can a list of the roughly 9000 weather stations [16]. McGuire Air Force Base, located not far from Princeton, New Jersey, is one of the archived weather stations. This particular weather station seemed good for a number of reasons: (i) it is about 50 miles from New York City and 30 miles from Philadelphia; (ii) it is in a rather undeveloped part of the state; (iii) it was established in 1937 and has been a major airbase since 1942; and finally (iv) it is only about 25 miles from the Atlantic Ocean and therefore its climate should be moderated somewhat by its proximity to an ocean (see Figure 2.1).

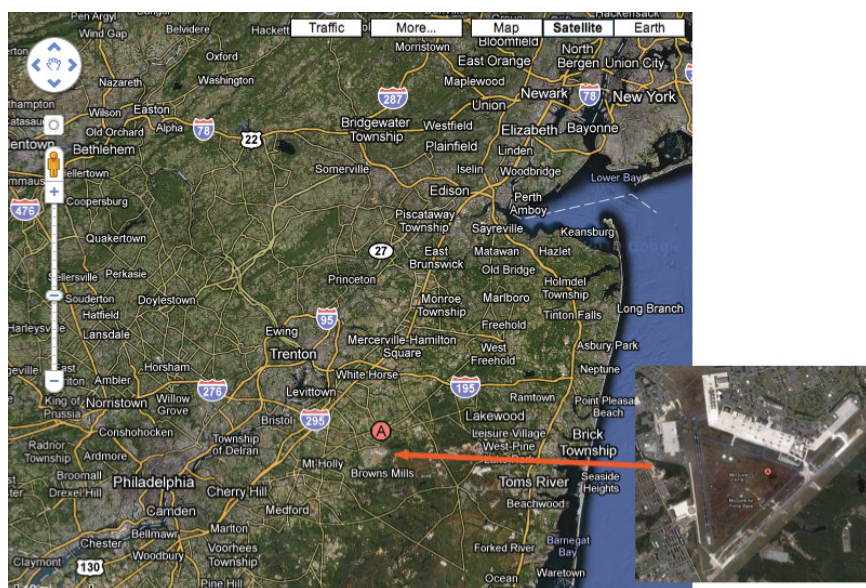


Fig. 2.1 McGuire Air Force Base. Imagery ©2012 TerraMetrics. Map data ©2012 Google.

**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*,

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

```

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;

```

**Fig. 3.1** The model expressed in the AMPL modeling language.

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  $x_0, x_1, \dots, x_5$  are unknown regression coefficients. Our aim is to find the values of these parameters that minimize the sum of the absolute deviations:

$$\begin{aligned}
 \min_{x_0, \dots, x_5} \sum_{d \in D} & |x_0 + x_1 d + x_2 \cos(2\pi d/365.25) + x_3 \sin(2\pi d/365.25) \\
 & + x_4 \cos(2\pi d/(10.7 \times 365.25)) + x_5 \sin(2\pi d/(10.7 \times 365.25)) - T_d|.
 \end{aligned}
 \tag{3.1}$$

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. The result is a *linear programming* problem. One can check that the solution to the LP formulation is identical to the solution to the original model whenever the original problem is “convex.” In particular, they are the same whenever one minimizes a nonnegative weighted sum of absolute values of linear expressions subject to linear equality and inequality constraints [19, Chapter 12].

The linear programming problem, expressed in the AMPL modeling language [6], is shown in Figure 3.1. AMPL models, with their associated user-supplied data sets, can be solved online using the Network Enabled Optimization Server (NEOS) at Argonne National Labs [4].

An LAD model was chosen instead of a least-squares model because LAD regression, like the median statistic, is insensitive to “outliers” in the data. The least-squares variant is discussed in section 6.

**3.1. Confidence Intervals.** Let  $x_0^*, \dots, x_5^*$  denote the optimal solution to (3.1) and let  $\varepsilon_d$  denote the corresponding deviations:

$$\begin{aligned}\varepsilon_d = & x_0^* + x_1^*d + x_2^* \cos(2\pi d/365.25) + x_3^* \sin(2\pi d/365.25) \\ & + x_4^* \cos(2\pi d/(10.7 \times 365.25)) + x_5^* \sin(2\pi d/(10.7 \times 365.25)) - T_d.\end{aligned}$$

In the well-known *bootstrap* method [5], we assume that these  $\varepsilon_d$ 's form an empirical distribution associated with an unknown underlying error distribution. As such, we can sample from these errors and generate new data:

$$\begin{aligned}T'_d = & x_0^* + x_1^*d + x_2^* \cos(2\pi d/365.25) + x_3^* \sin(2\pi d/365.25) \\ & + x_4^* \cos(2\pi d/(10.7 \times 365.25)) + x_5^* \sin(2\pi d/(10.7 \times 365.25)) - \varepsilon'_d,\end{aligned}$$

where the  $\varepsilon'_d$  are drawn independently and with replacement from the set  $\{\varepsilon_d : d \in D\}$ . In this manner we can generate several alternate data sets that are statistically similar to the original one, and we can then recompute the parameters  $x_0, \dots, x_5$  using these replicated data sets to get multiple estimates for each parameter and thereby compute  $2\sigma$ -confidence intervals for each parameter (or any other derived parameter).

**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 together with their  $2\sigma$ -confidence intervals are

$$\begin{aligned}x_0 &= 52.6 \pm 0.27^\circ \text{F}, \\ x_1 &= 3.63 \pm 0.75^\circ \text{F/century}, \\ x_2 &= -20.4 \pm 0.16^\circ \text{F}, \\ x_3 &= -8.31 \pm 0.17^\circ \text{F}, \\ x_4 &= -0.197 \pm 0.137^\circ \text{F}, \\ x_5 &= 0.211 \pm 0.202^\circ \text{F}.\end{aligned}$$

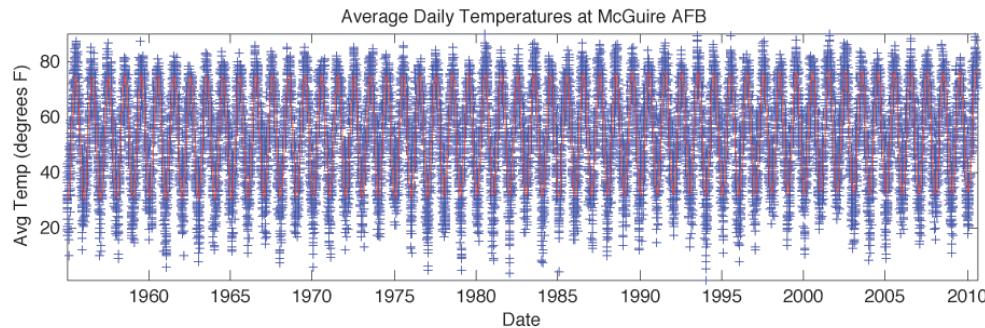
**4.1. Linear Trend.** From  $x_0$ , we see that the nominal temperature at McGuire AFB was  $52.56 \pm 0.27^\circ \text{F}$  (on January 1, 1955).

We also see, from  $x_1$ , that there is a positive trend of  $0.000099 \pm 0.000020^\circ \text{F/day}$ , which scales to  $3.63 \pm 0.75^\circ \text{F per century}$ . This result is consistent with results from global climate change models, which predict a per century warming rate of between  $2.0^\circ \text{C}$  and  $2.4^\circ \text{C}$  (see [12, 9]).

**4.2. Amplitude 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:

$$\sqrt{x_2^2 + x_3^2} = 22.02 \pm 0.15^\circ \text{F}.$$

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, 74.58 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.



**Fig. 4.1** Plot showing actual data and regression curve. Blue: Average daily temperatures at McGuire AFB from 1955 to 2010. Red: Output from LAD regression model.

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:

$$\sqrt{x_4^2 + x_5^2} = 0.2887 \pm 0.156^\circ \text{F}.$$

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

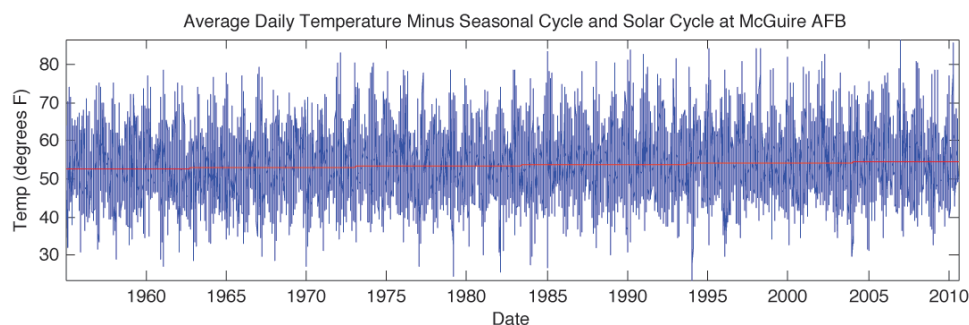
**4.3. Phase of the Sinusoidal Fluctuations.** Close inspection of the output shows that January 22 is nominally the coldest day in the winter and July 24 is the hottest day of summer. The plus/minus  $2\sigma$ -error on these estimates is less than half a day. It is perhaps worth noting that the coldest 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 10.7 year solar cycle. In this case, the plus/minus  $2\sigma$ -error on this estimate is about 400 days. It is well known that the solar cycle had its last minimum in 2007 [17]. 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  $0.29^\circ \text{F}$ , is strong support for fidelity of the LAD model.

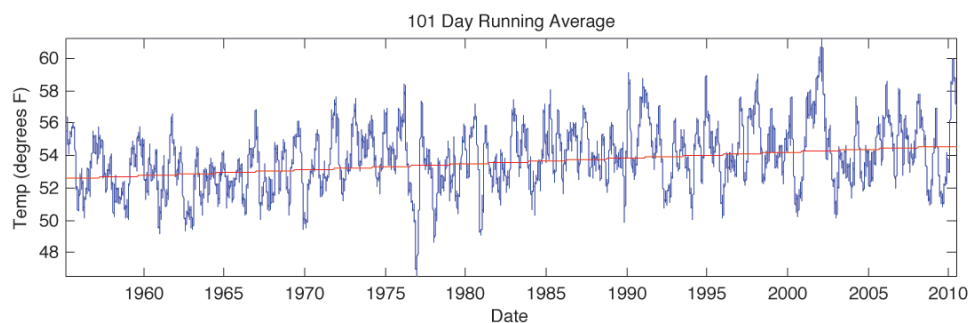
**4.4. Visualizing the Results.** Figure 4.1 shows a plot of all 20,309 data points. Overlaid on these data points is the solution of the LAD regression model. Daily and seasonal fluctuations completely dominate other effects. It is impossible to “see” any linear warming trend or the solar cycle.

Figure 4.2 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 4.3 is derived from Figure 4.2 by applying a 101-day rolling average to each data point. The rolling average reduced the extreme fluctuations to about one-tenth their original amplitude, thus making the linear warming trend quite apparent. In New Jersey we have *local warming*.



**Fig. 4.2** As before but with sinusoidal seasonal variation removed and sinusoidal solar-cycle variation removed as well.



**Fig. 4.3** Smoothed seasonally subtracted plot. To smooth out high-frequency fluctuations, we use 101-day rolling averages of the data.

**5. Estimating the Period of the Solar Cycle.** The length of the solar cycle is only approximately 10.7 years [21]. 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 potentially harder to solve. However, nonlinear (local) optimization algorithms produce provably locally optimal solutions “near” the initially provided values for the variables. Hence, for problems in which rough estimates of the optimal values are known, one can expect such algorithms to produce the desired solution. 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 10.78-year solar cycle, in close agreement with the nominal value of 10.7 years. One might argue that it is unfair to initialize  $x_6$  so that the starting point is so close to the known correct solution. Table 5.1 shows the period of the solar cycle obtained using various initial guesses.

**Table 5.1** The derived period, in years, of the solar cycle computed using different initial guesses for variable  $x_6$  in the nonlinear versions of both the LAD and least-squares optimization models. In the table values of  $x_6$  are specified by giving the period of the sinusoid measured in years (recall that  $x_6 = 1$  corresponds to a 10.7-year solar cycle). Note that, for the LAD model, four of the eleven cases considered produced an answer close to the known answer, whereas none of them is very close using the least-squares model.

Initial period (in years)	Locally optimal period (in years)	
	LAD	Least Squares
6	6.64	14.65
7	10.78	6.53
8	10.78	6.53
9	5.67	6.53
10	8.12	8.75
10.7	10.78	14.65
11	-14.61	14.65
12	4.19	14.65
13	-14.61	5.74
14	10.78	14.65
15	-10.78	$\infty$

**6. 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 (*nonlinear*) *least-squares model*. The problem is no longer representable as a linear programming problem, but its objective function is still convex and the problem is still easy to solve using nonlinear optimization software. Fixing  $x_6$  to one (i.e., fixing the solar cycle to 10.7 years), the problem becomes a linear least-squares regression model that can also be solved using any number of statistical packages, such as R [20]. The solution, however, is *sensitive* to outliers, which may or may not be present. Here's the output associated with the nonlinear least-squares formulation:

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

In this case, the rate of local warming is  $4.37^\circ\text{F}$  per century. This number lies near the upper limit of the confidence interval given before. As Table 5.1 shows, the model also produces a *wrong answer* for the period of the solar cycle for each of the eleven values I used to initialize  $x_6$ , including  $x_6 = 1$ , which corresponds to a solar cycle of 10.7 years. It is well known that the period of the solar cycle has been about 10.7 years for the past few centuries.

**7. Humidity.** It is well known that higher temperatures imply more evaporation of water into the atmosphere; that is, increased temperature implies increased humidity. It turns out that the NOAA data sets contain not only temperature data but

also dew point data. *Dew point* is the temperature below which water in the atmosphere condenses out to form clouds/fog. It is a quantitative measure of the amount of water in the atmosphere—the more water vapor, the higher the temperature must be for the air to hold that water vapor. It is common in summer months for there to be fog in the morning. The temperature and the dew point match when it is foggy. One says that the *relative humidity* is 100% at such times. In the summer months at midlatitudes, dew points are often up in the 60°F to 70°F range. In the winter, obviously, the dew point is much lower—winter air is significantly dryer than summer air. Anyway, while weathermen typically report humidity in terms of relative humidity, dew point is a more direct measure of the amount of water vapor in the atmosphere.

Given that NOAA tabulates this measure in the same data sets as used above, it turns out to be almost trivial to modify the scripts to perform a dew point regression. What we discover is that, in 1955 at McGuire AFB, the nominal dew point (given by  $x_0$ ) was 41.4°F and that on the dampest summer days it was expected to be 22.6°F higher and on the driest winter days it was expected to be that much lower. Furthermore, the dew point is going up at a rate of 5.4°F per century—a rate even greater than the rate that temperatures are rising, which means that relative humidity is also increasing.

In the U.S., the National Weather Service reports a number called the *heat index* on hot summer days. This index combines temperature and humidity into a single number on a temperature scale. However, this measure is fairly subjective as it is based on relative humidity and aims to approximate how the temperature “feels” to a person. In Canada, a similar measure called *humidex* is used (see [14, 18]). An advantage of the humidex over heat index is that humidex is based on temperature and dew point rather than temperature and relative humidity. In fact, the formula for humidex  $H$  (in degrees Fahrenheit) is rather simple:

$$H = T + 6.11e^{5417.7530(\frac{1}{273.16} - \frac{1}{D})} - 10,$$

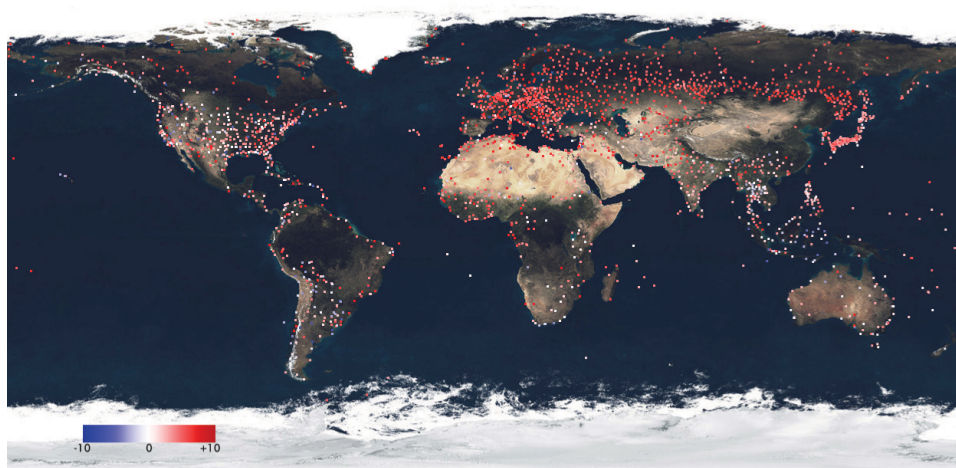
where  $T$  is temperature in degrees Fahrenheit and  $D$  is dew point in degrees Kelvin (of course, it is easy to convert degrees Fahrenheit to degrees Kelvin).

Again, the fact that NOAA tabulates dew point together with temperature data makes it easy to produce values for the humidex. What we discover is that, in 1955 at McGuire AFB, the nominal humidex (given by  $x_0$ ) was 52.9°F and that on the hottest/dampest summer days it was expected to be 30.8°F higher and on the coldest/driest winter days it was expected to be that much lower. Furthermore, the humidex is going up at a rate of 6.5°F per century—a rate even greater than the rate that temperatures are rising.

**8. Going Global.** For fun, I applied the model to thousands of local weather stations around the world to produce a global “local warming map.” Specifically, data files were downloaded from every NOAA weather station for which collection commenced prior to January 1, 1955, and is currently ongoing. There may be, and it turns out that there usually are, gaps in the data—sometimes just a day here or there is missing but often there are multiyear gaps. Hence, the download script stipulated that the location must have collected at least 3650 days of data (i.e., 10 years worth). The resulting map is shown in Figure 8.1.

I should mention that no attempt was made to filter out bad data. Also, seasonal variations are not sinusoidal in the tropics, and so the simple model presented earlier is not very good at tropical latitudes.





**Fig. 8.1** A map of 2399 local warming results distributed around the globe. In units of degrees Fahrenheit per century, the mean of these values is 4.18, the median value is 4.53, and the standard deviation is 2.94. Dividing by the square root of the sample size, a confidence interval for the mean is  $[4.12, 4.24]$ . Map used with permission.

One can clearly see from Figure 8.1 that warming is more pronounced in the Northern Hemisphere. Such an observation is consistent with the hypothesis that warming over the last century is largely anthropogenic.

**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 data set.

In producing confidence intervals, one of our assumptions is wrong. In the bootstrap method, 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 scale the estimated standard deviation by the square root of  $m$ . For example, if we assume that the correlation length is 9 days, then the new confidence interval for  $x_1$  is  $[1.38^\circ \text{F}, 5.88^\circ \text{F}]$  per century.

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—it falls apart for data collection sites 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 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 basic models we have considered can be improved in various ways. Such improvements could inspire many interesting student projects.

**9.1. Getting the Data.** Since the NOAA data are 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>

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

#### REFERENCES

- [1] P. BLOOMFIELD AND W. STEIGER, *Least Absolute Deviations: Theory, Applications, and Algorithms*, Birkhäuser, Boston, 1983.
- [2] A. BOVIK, T. HUANG, AND D. M. JR., *A generalization of median filtering using linear combinations of order statistics*, IEEE Trans. Acoustics Speech Signal Process., 31 (1983), pp. 1342–1350.
- [3] M. CAMPBELL AND M. GARDNER, *Statistics in medicine: Calculating confidence intervals for some nonparametric analyses*, British J. Medicine, 296 (1988), pp. 1454–1456.
- [4] J. CZYZYK, M. MESNIER, AND J. MORE, *The NEOS server*, Comput. Sci. Engrg., 5 (1998), pp. 68–75.
- [5] B. EFRON, *Bootstrap methods: Another look at the jackknife*, Ann. Statist., 7 (1979), pp. 1–26.
- [6] R. FOURER, D. GAY, AND B. KERNIGHAN, *AMPL: A Modeling Language for Mathematical Programming*, Duxbury Press, Belmont, CA, 1993.
- [7] J. GOTT, M. VEGELEY, S. PODARIU, AND B. RATRA, *Median statistics,  $h_0$ , and the accelerating universe*, Astrophys. J., 549 (2001), pp. 1–17.
- [8] G. HEGERL, H. VON STORCH, K. HASSELMANN, B. SANTER, U. CUBASCH, AND P. JONES, *Detecting greenhouse-gas-induced climate change with an optimal fingerprint method*, J. Climate, 9 (1996), pp. 2281–2306.
- [9] J. HOUGHTON, L. M. FILHO, D. GRIGGS, AND K. MASKELL, *An Introduction to Simple Climate Models Used in the IPCC Second Assessment Report*, IPCC Technical Paper II, Intergovernmental Panel on Climate Change, Geneva, 1997.
- [10] P. HUBER AND E. RONCHETTI, *Robust Statistics*, 2nd ed., Wiley, Hoboken, NJ, 2009.
- [11] P. JONES, K. BRIFFA, T. BARNETT, AND S. TETT, *High-resolution palaeoclimatic records for the last millennium: Interpretation, integration and comparison with general circulation model control-run temperatures*, The Holocene, 8 (1998), pp. 455–471.
- [12] T. KARL, R. KNIGHT, AND B. BAKER, *The record breaking global temperatures of 1997 and 1998: Evidence for an increase in the rate of global warming?*, Geophys. Res. Lett., 27 (2000), pp. 719–722.
- [13] M. MANN, R. BRADLEY, AND M. HUGHES, *Global-scale temperature patterns and climate forcing over the past six centuries*, Nature, 392 (1998), pp. 779–787.
- [14] J. MASTERTON AND F. RICHARDSON, *A Method of Quantifying Human Discomfort Due to Excessive Heat and Humidity*, AES, Environment Canada, Downsview, ON, Canada, 1979.
- [15] NOAA, *Climate Data Format and Download Instruction*, <ftp://ftp.ncdc.noaa.gov/pub/data/gsod/readme.txt>.
- [16] NOAA, *List of Weather Stations*, <ftp://ftp.ncdc.noaa.gov/pub/data/gsod/ish-history.txt>.
- [17] N. SCAFETTA AND R. WILLSON, *ACRIM-gap and TSI trend issue resolved using a surface magnetic flux TSI proxy model*, Geophys. Res. Lett., 36 (1997).
- [18] K. SMOYER-TOMIC AND D. G. C. RAINHAM, *Beating the Heat: Development and Evaluation of a Canadian Hot Weather Health-Response Plan*, 2001; available online from <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1240506/pdf/ehp0109-001241.pdf>.
- [19] R. VANDERBEI, *Linear Programming: Foundations and Extensions*, 3rd ed., Springer, New York, 2007.
- [20] W. VENABLES AND D. SMITH, *An Introduction to R*, Network Theory Ltd., Godalming, UK, 2004.
- [21] R. WILLSON AND H. HUDSON, *The sun's luminosity over a complete solar cycle*, Nature, 351 (1991), pp. 42–44.