Paper 233-2008

Local Measures of Global Warming

Jun Lu, Texas Tech University, Lubbock, TX 79409-1042
Clyde. F. Martin, Texas Tech University, Lubbock, TX 79409-1042

ABSTRACT

Global warming is currently a focus of attention. This paper provides how to use SAS to create effective approaches for local measures of global warming. Based on analysis of a large data set, almost 100 years' daily temperatures in a local weather station, it gives the detail procedures for testing the equality of variance and mean, and finding the trend of time plot using "loess smoother" technique. The results suggest that global warming really exists in the local place, and they also provide evidences that may support global warming.

KEYWORDS

Global Warming, SAS, Test for Equality of Variance, Test for Equality of Mean, Test for Normality, Loess Smoother

1. INTRODUCTION

With the development of modern industry, many environmental problems arise, and global warming is currently a focus of attention. Based on the analysis of almost 100 years' local daily mean temperature records (Period: 1/1/1914-12/31/2006, Station: Lubbock International Airport, TX 79403, CoopID: 415411, and Data Source: http://weather-warehouse.com), this paper provides how to use SAS to make effective local measures of global warming. It also explores how to prove the existence of global warming and how to show that the temperature records are inconsistent with an unchanging climate.

2. HYPOTHESIS

Two parameters, mean and variance, are the key roles in evaluating that the temperature records are inconsistent with an unchanging climate. Thus, the equality of the variances and the means of the daily mean temperature TmeanF in each year are needed to be tested. The hypotheses are as follows:

a. hypothesis for test the equality of variance

H_0: $\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_n^2$

H_1: not all σ_i^2's are equal

b. hypothesis for test the equality of mean

H_0: $\mu_1 = \mu_2 = \ldots = \mu_n$

H_1: not all μ_i's are equal

3. TEST FOR THE NORMALITY OF TMEANF

Although several statistical tests for testing the equality of variance have been proposed, some of them are sensitive to normality assumption. When this assumption is not valid, they could not be used. So, we first test the normality of 93 years' local daily mean temperature TmeanF (degrees Fahrenheit. Use the function of "summary statistics" in SAS/Analysis, we get the summary statistics for TmeanF, see Appendix) which is the average of daily maximum temperature and daily minimum temperature. The function of "fit distributions" in SAS/Analysis was used, and the result was:

The UNIVARIATE Procedure
Fitted Distribution for TmeanF

Parameters for Normal Distribution

Parameter	Symbol	Estimate
Mean	Mu	60.23262
Std Dev	Sigma	16.40919

Goodness-of-Fit Tests for Normal Distribution

Test	---Statistic----		-----p Value-----	
Kolmogorov-Smirnov	D	0.080071	Pr > D	<0.010
Cramer-von Mises	W-Sq	53.494492	Pr > W-Sq	<0.005
Anderson-Darling	A-Sq	346.378151	Pr > A-Sq	<0.005

Quantiles for Normal Distribution

	------Quantile------	
Percent	Observed	Estimated
1.0	22.0000	22.0591
5.0	32.5000	33.2419
10.0	38.0000	39.2034
25.0	47.5000	49.1648
50.0	61.5000	60.2326
75.0	75.0000	71.3004
90.0	80.5000	81.2618
95.0	82.5000	87.2233
99.0	86.0000	98.4061

Table 1. Test for the Normality of TmeanF

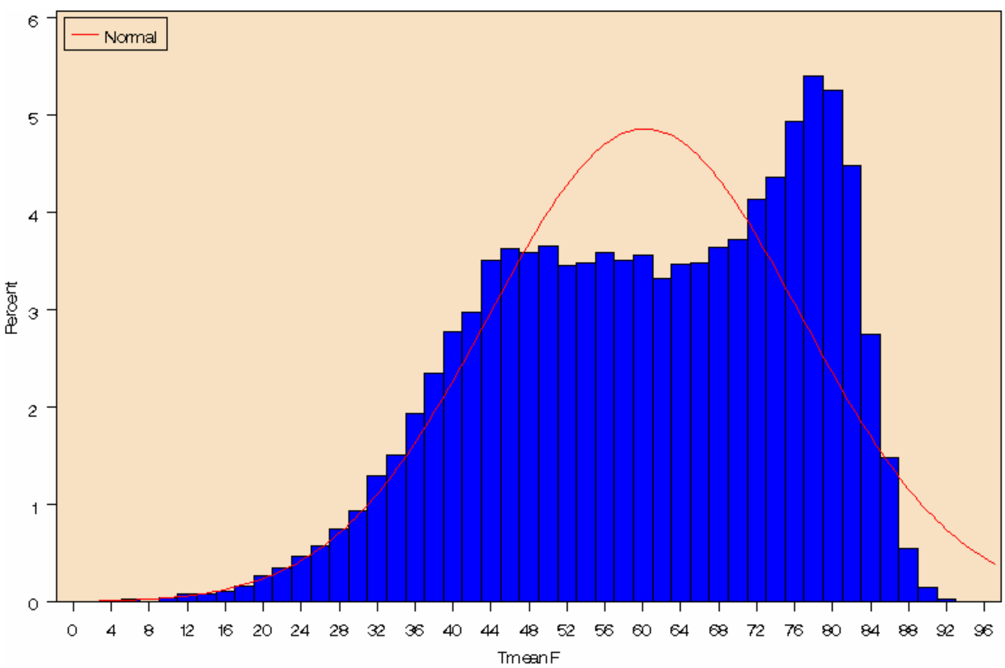

Figure 1. Histogram for Testing for the Normality of TmeanF

Since p-values of three tests, Kolmogorov-Smirnov, Cramer-von Mises, and Anserson-Darling, were all smaller than 0.01, the normality assumption was not valid. The Figure 1 also showed the shape of distribution was skewed to the left, which meant the number of lower TmeanF was relatively small, while the number of higher TmeanF was relatively large.

4. TEST FOR EQUALITY OF VARIANCE

The Brown and Forsythe's test is a very nice procedure that is robust to departure from normality. It uses the absolute deviation of the observations in each treatment from the treatment median. The function "tests for equal variance" in SAS/Analysis was used, and the result was:

The ANOVA Procedure

Brown and Forsythe's Test for Homogeneity of TmeanF Variance
ANOVA of Absolute Deviations from Group Medians

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Year	92	22134.0	240.6	3.15	<.0001
Error	33847	2586726	76.4241		

Table 2. Test for Equality of Variance

Since p-values of the Brown and Forsythe's test was smaller than 0.001, we rejected H_0: $\sigma_1^2 = \sigma_2^2 = \ldots = \sigma_{93}^2$, which meant not all the variances of the daily mean temperature TmeanF in each year were equal.

5. TEST FOR EQUALITY OF MEAN

Kruskal-Wallis nonparametric test that is not sensitive to the normality assumption is often used for comparing group means. Here, the function "Nonparametric One-way ANOVA → Wilcoxon (Kruskal-Wallis test)" in SAS/Analysis was used, and the result was:

Kruskal-Wallis Test

Chi-Square	190.6251
DF	92
Pr > Chi-Square	<.0001

Table 3. Test for Equality of Mean

Since the P-value for H = 190.6251 was smaller than 0.0001, we would reject the null hypothesis, and concluded that the population distribution of TmeanF in each year (1914-2006) was not all identical, and observations in some years were larger than others.

6. TREND OVER TIME

In order to find a clear trend of the variances and means over time, smoothing techniques are used to reduce irregularities (random fluctuations) in time series data and provide a clearer view of the true underlying behavior of the series. One common smoothing technique is called "Lowess," which stands for LOcally WEighted Scatter plot Smoothing. A newer formula based version of "Lowess" is "Loess" which is short for Local Regression. The "Loess" procedure in SAS provides an automatic smoothing parameter selection which minimizes a bias corrected Akaike Information Criterion and balances the residual sum of squares against the smoothness of the fit. The smoothing parameter controls the shape of fitted "Loess" curve. It is the proportion of the data which influence the smooth at each local neighborhood. Larger values of the smoothing parameter give more smoothness. The SAS program of "Loess" Smoother for time plot of variance was:

```
data TmeanVar;
    input  Year TmeanVar @@;
    format Year d4.0;
    format DepVar d5.2;
datalines;
/*input data set*/;

symbol1 color=black value=dot ;
proc gplot data=TmeanVar;
        title1 'Scatter Plot of Tmean Variance Data';
        plot TmeanVar*Year;

proc loess data=TmeanVar;
        model TmeanVar=Year/details(OutputStatistics);

proc loess data=TmeanVar;
        model TmeanVar=Year;
        ods output OutputStatistics=Results;

symbol1 color=black value=dot;
symbol2 color=black interpol=join value=none;

%let opts=vaxis=axis1 hm=3 vm=3 overlay;

axis1 label=(angle=90 rotate=0);

proc gplot data=Results;
        title1 'TmeanVar Data with Default LOESS Fit';
        plot  DepVar*Year Pred*Year/ &opts;
run;
```

Table 4. SAS Program of Loess Smoother for Time Plot of Variance

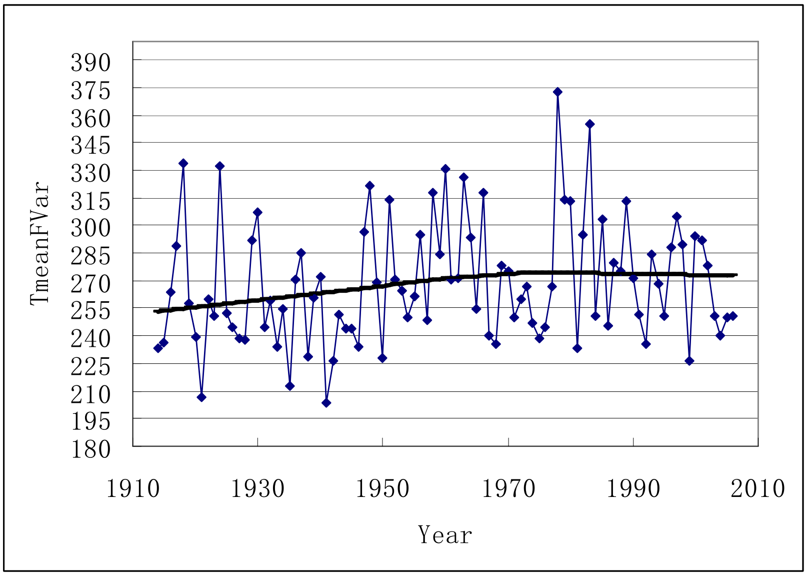

Figure 2: Time Plot of Variances with Fitted “Loess” Curve

(Default optimal smoothing parameter: 0.91935)

From Figure 2, we could find the overall trend of the variances of TmeanF in each year over time was increasing, and it went from a level of about 253 in 1914 to about 273 in 2006.

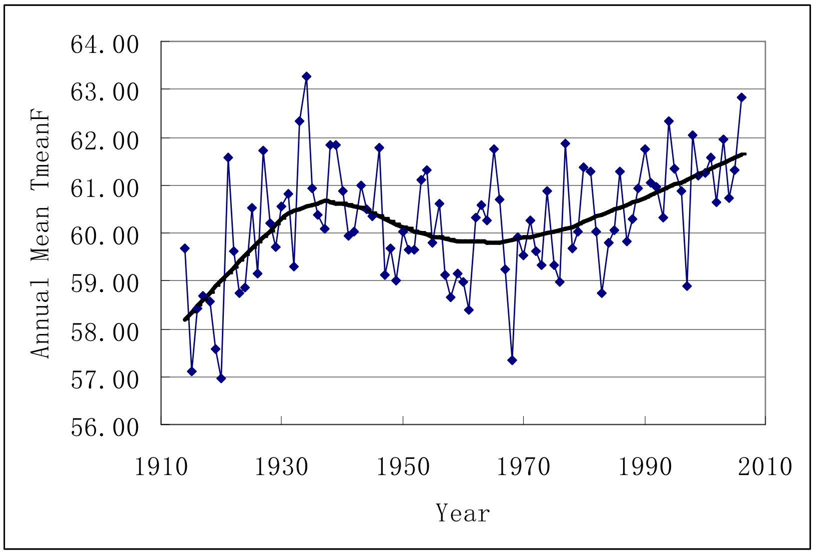

Figure 3: Time Plot of Means with Fitted "Loess" Curve

(Default optimal smoothing parameter: 0.39247)

From Figure 3, we could find the overall trend of the annual means of TmeanF over time was increasing, and it creased from a level of about 58.2 (degrees Fahrenheit) in 1914 to about 61.6 (degrees Fahrenheit) in 2006.

7. CONCLUSIONS

We analyzed the daily mean temperatures (degrees Fahrenheit) measured at the station of Lubbock International Airport, TX 79403, from 1/1/1914 to 12/31/2006. Based on the above analysis, we found the data set was not normal distribution. It was skewed to the left which meant the number of higher daily mean temperature was relatively large. From the test of equality of variance and the test of equality of mean, we rejected the null hypothesis of equal variance and equal mean. Furthermore, the fitted "loess" curve showed the trend of the variance of the daily mean temperature over year was increasing. From 1914 to 2006, it increased about 20 levels. The fitted "loess" curve of the mean of the daily mean temperature over year also showed the hypothesis of equal mean was not valid, and the overall trend was also increasing. There was about a 3 ^{o}F increase during these 93 years. These results suggested global warming really existed in the local place, and they also provided evidence that may supports the existence of global warming.

APPENDIX

YEAR	Mean	Std. Dev.	Std. Error	Variance	N	N Miss	Min	Max	Median
.	60.23	16.41	0.09	269.26	33940	28	0.50	96.00	62.00
1914	59.68	15.28	0.81	233.53	358	7	23.50	85.00	60.75
1915	57.10	15.37	0.81	236.15	359	6	16.00	81.50	59.00
1916	58.42	16.25	0.85	263.94	366	0	15.50	84.50	60.00
1917	58.68	17.00	0.89	289.12	365	0	15.00	85.00	58.00
1918	58.58	18.27	0.96	333.76	365	0	4.00	89.00	60.50
1919	57.59	16.04	0.84	257.32	365	0	6.50	85.00	58.50
1920	56.97	15.47	0.81	239.32	366	0	12.50	83.50	58.50
1921	61.57	14.37	0.75	206.53	365	0	24.50	83.50	61.50
1922	59.61	16.12	0.84	259.70	365	0	11.50	86.50	61.00
1923	58.74	15.83	0.83	250.48	365	0	17.00	84.50	58.00
1924	58.87	18.22	0.95	331.91	366	0	9.50	90.00	61.00
1925	60.52	15.89	0.83	252.40	365	0	18.00	90.50	62.00
1926	59.15	15.63	0.82	244.42	365	0	12.50	85.50	59.00
1927	61.73	15.44	0.81	238.54	365	0	19.50	85.00	62.50
1928	60.20	15.43	0.81	238.17	366	0	13.50	91.50	60.50
1929	59.70	17.08	0.89	291.57	365	0	12.50	85.50	62.00
1930	60.55	17.52	0.92	306.80	365	0	6.00	88.50	63.00
1931	60.83	15.64	0.82	244.76	365	0	25.50	86.00	60.00
1932	59.29	16.10	0.84	259.29	366	0	17.00	86.00	61.75
1933	62.35	15.30	0.80	234.08	365	0	0.50	89.50	63.50
1934	63.28	15.96	0.84	254.64	365	0	24.50	89.50	64.00
1935	60.92	14.58	0.76	212.53	365	0	17.00	88.00	62.50
1936	60.39	16.44	0.86	270.25	366	0	17.00	90.00	61.50
1937	60.09	16.89	0.88	285.14	365	0	10.50	86.50	61.50
1938	61.85	15.14	0.79	229.08	365	0	21.00	84.50	65.00

1939	61.84	16.15	0.85	260.81	365	0	19.00	89.50	64.00
1940	60.87	16.50	0.86	272.32	366	0	13.50	88.00	63.25
1941	59.94	14.27	0.75	203.54	365	0	31.00	83.50	61.00
1942	60.02	15.05	0.79	226.41	365	0	15.50	85.50	61.50
1943	60.98	15.85	0.83	251.23	365	0	14.50	89.00	62.00
1944	60.51	15.63	0.82	244.23	366	0	23.00	88.00	61.00
1945	60.34	15.62	0.82	243.85	365	0	23.50	86.50	60.00
1946	61.77	15.30	0.80	234.20	363	2	19.50	88.00	64.00
1947	59.11	17.22	0.92	296.44	352	13	9.00	88.00	60.75
1948	59.68	17.93	0.94	321.66	366	0	12.50	88.00	61.25
1949	59.00	16.41	0.86	269.24	365	0	13.50	85.50	60.50
1950	60.02	15.10	0.79	227.92	365	0	17.00	86.50	63.00
1951	59.66	17.73	0.93	314.32	365	0	7.00	88.50	59.00
1952	59.65	16.45	0.86	270.74	366	0	24.00	87.50	59.25
1953	61.12	16.26	0.85	264.37	365	0	19.00	89.50	60.00
1954	61.31	15.80	0.83	249.75	365	0	19.50	86.00	61.50
1955	59.80	16.16	0.85	261.09	365	0	22.00	87.50	61.50
1956	60.62	17.16	0.90	294.64	366	0	13.50	86.00	62.75
1957	59.13	15.77	0.83	248.85	365	0	17.00	91.00	59.00
1958	58.67	17.83	0.93	317.82	365	0	18.00	90.50	59.50
1959	59.15	16.86	0.88	284.27	365	0	12.00	87.00	59.00
1960	58.99	18.18	0.95	330.48	366	0	9.50	88.00	62.00
1961	58.38	16.44	0.86	270.24	365	0	16.50	83.00	60.00
1962	60.32	16.47	0.86	271.29	365	0	6.00	86.00	62.00
1963	60.58	18.05	0.94	325.93	365	0	1.00	86.50	65.50
1964	60.25	17.13	0.90	293.55	366	0	17.50	86.50	61.50
1965	61.76	15.95	0.84	254.56	365	0	24.00	86.00	63.00
1966	60.71	17.82	0.93	317.53	365	0	12.00	92.00	63.00
1967	59.25	15.49	0.81	239.94	365	0	21.00	86.50	62.00
1968	57.33	15.34	0.80	235.26	366	0	14.00	85.50	57.75
1969	59.91	16.68	0.87	278.25	365	0	22.00	87.00	60.00
1970	59.54	16.59	0.87	275.14	365	0	20.50	88.50	59.00
1971	60.27	15.80	0.83	249.74	365	0	11.50	88.50	60.50
1972	59.62	16.12	0.84	259.98	366	0	14.50	87.00	62.75
1973	59.34	16.34	0.86	267.04	365	0	11.00	86.50	60.50
1974	60.88	15.72	0.82	246.98	365	0	18.00	87.50	62.50
1975	59.33	15.44	0.81	238.54	365	0	17.50	85.50	61.00
1976	58.99	15.65	0.82	244.83	366	0	14.50	85.00	60.50
1977	61.88	16.34	0.86	266.86	365	0	10.00	88.50	63.00
1978	59.69	19.31	1.01	372.79	365	0	11.00	90.50	63.00
1979	60.03	17.71	0.93	313.74	365	0	5.50	88.50	62.50
1980	61.37	17.71	0.93	313.59	366	0	20.50	92.00	60.50
1981	61.28	15.27	0.80	233.05	365	0	21.00	90.00	62.00
1982	60.03	17.17	0.90	294.86	365	0	9.00	88.50	61.50
1983	58.75	18.84	0.99	354.98	365	0	4.00	90.50	61.00
1984	59.79	15.83	0.83	250.57	366	0	10.50	86.00	59.25
1985	60.05	17.42	0.91	303.49	365	0	6.00	88.00	62.00
1986	61.29	15.67	0.82	245.48	365	0	13.50	88.50	62.00
1987	59.83	16.72	0.88	279.60	365	0	16.00	90.00	62.50
1988	60.28	16.60	0.87	275.44	366	0	17.50	86.00	62.00
1989	60.94	17.70	0.93	313.22	365	0	6.50	89.50	65.00
1990	61.75	16.48	0.86	271.53	365	0	9.00	94.00	63.00
1991	61.04	15.87	0.83	251.79	365	0	22.00	86.00	63.00
1992	60.96	15.36	0.80	235.90	366	0	22.00	89.00	64.00
1993	60.33	16.86	0.88	284.17	365	0	19.50	89.50	61.00
1994	62.33	16.39	0.86	268.59	365	0	21.50	96.00	63.00
1995	61.35	15.83	0.83	250.66	365	0	21.00	92.00	61.00
1996	60.87	16.98	0.89	288.45	366	0	14.50	90.00	63.50
1997	58.88	17.46	0.91	305.02	365	0	9.50	86.00	58.50
1998	62.04	17.01	0.89	289.47	365	0	14.50	91.50	62.50
1999	61.19	15.04	0.79	226.15	365	0	26.00	88.50	60.00
2000	61.27	17.15	0.90	293.97	366	0	19.00	90.50	62.00
2001	61.57	17.08	0.89	291.89	365	0	22.50	89.50	62.50
2002	60.64	16.69	0.87	278.49	365	0	23.00	87.50	61.50
2003	61.97	15.85	0.83	251.08	365	0	18.00	90.50	64.50
2004	60.72	15.49	0.81	239.80	366	0	14.00	87.00	62.25
2005	61.32	15.80	0.83	249.77	365	0	16.50	87.00	61.00
2006	62.84	15.84	0.83	251.04	365	0	22.00	89.00	64.00

Table 5. Summary statistics for TmeanF

(Period: 1/1/1914 - 12/31/2006, Station: Lubbock International Airport, TX 79403, CoopID: 415411)

REFERENCE

1. Agresti, A. and C. A. Franklin. 2006. *Statistics: the Art and Science of Learning from Data*. Prentice Hall.
2. Cleveland, W. S. 1979. Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the American Statistical Association* 74: 829-836. Available WWW: http://www.jstor.org/view/01621459/di985942/98p1302g/0.
3. Cleveland, W. S. 1981. "LOESS": A Program for Smoothing Scatterplots by Robust Locally Weighted Regression. *The American Statistician* 35(54). Available WWW:

http://www.jstor.org/view/00031305/di020566/02p0132d/0.

4. Cleveland, W. S., S. J. Devlin, and E. Grosse. 1988. Regression by Local Fitting. *Journal of Econometrics* 37: 87-114.
5. Cleveland, W. S. and E. Grosse. 1991. Computational Methods for Local Regression. *Statistics and Computing* 1: 47-62.
6. Cleveland, W. S., E. Grosse and W. M. Shyu. 1992. *Local Regression Models*. Chapman and Hall.
7. Cohen, R.A.1998. An Introduction to PROC LOESS for Local Regression. *Proceedings of the Twenty-fourth Annual SAS Users Group International Conference*. Available WWW: http://www.ats.ucla.edu/STAT/sas/library/loesssugi.pdf.
8. Coles, S. 2001. *An Introduction to Statistical Modeling of Extreme Values*. London; New York: Springer.
9. Johnson, R. A. and G. K. Bhattacharyya. 2005. *Statistics: Principles and Methods*. Hoboken, NJ: John Wiley & Sons.
10. Mansouri, H. 2007. Handouts of SAS/Base, SAS/Analysis, SAS/Enterprise Guide Short Courses. Technology Support, Texas Tech University, Lubbock, TX.
11. Montgomery, D. C. 2001. *Design and Analysis of Experiments*. New York: John Wiley.
12. SAS Institute Inc. SAS 9.1 Help and Documentation.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author(s):

Name: Jun Lu

Enterprise/Address: Department of Mathematics and Statistics, Texas Tech University, Lubbock, TX 79409-1042

E-mail: june.lu@ttu.edu