The Influence of Station Numbers on Temperature Measurement

by Jonathan J. Drake

Reprinted from
ENERGY & ENVIRONMENT
VOLUME 21 No. 8 2010

SPPI
Science & Public Policy Institute
"science-based policy for a better world"

SPPI Reprint Series ♦ February 14, 2011

THE INFLUENCE OF STATION NUMBERS ON TEMPERATURE MEASUREMENT

by

Jonathan J. Drake (UK)

Reprinted from

ENERGY &
ENVIRONMENT

VOLUME 21 No. 8 2010

MULTI-SCIENCE PUBLISHING CO. LTD.
5 Wates Way, Brentwood, Essex CM15 9TB, United Kingdom

THE INFLUENCE OF STATION NUMBERS ON TEMPERATURE MEASUREMENT

Jonathan J. Drake
jdrake@ntlworld.com

1. INTRODUCTION

An intriguing article regarding the influence of the number of stations on the global temperature measurement has been written by Ross McKitrick and published on his website here: http://www.uoguelph.ca/~rmckitri/research/nvst.html. Looking at his graph, the eye is instantly drawn to the apparent correlation between the stations and temperature.

There is a huge step in the raw mean temperatures around 1990 with a synchronous change in the number of stations. The adjustments and gridding methods employed to create the accepted global temperature records, are reportedly satisfactory to deal with such data aberrations and on visual inspection there is no immediately obvious problem. However, that does not rule out the possibility that the gridded end product contains artefacts of the processing and/or character from the raw data that are not related to climate and that could potentially distort the overall picture.

This brief communication describes a method, utilising the dataset from Ross McKitrick and Joe D'Aleo, to calculate a historic temperature record through modelling the relationship between raw mean temperatures and the number of stations.

2. ANALYSIS

The dataset used here was obtained from Ross McKitrick (1). Unfortunately, it has a somewhat limited chronological range, only dating back to 1950. It is not known why it starts at that time, but there are adequate data points to make a reasonable assessment, at least for the purposes of this communication. Please note that the raw data is believed to be annual averages of unprocessed temperatures and thus will not necessarily reflect the adjusted historical records provided by GISS, CRU, etc.

The first graph, Figure 1, illustrates the issue by showing the raw mean temperatures and number of stations. Both series display a large jump in the data around 1990 but that is not all; there appears to be a correlation between two across the rest of the date span as well.

So, let's see what the correlation actually looks like by plotting temperature against number of stations in a scatter plot, Figure 2.

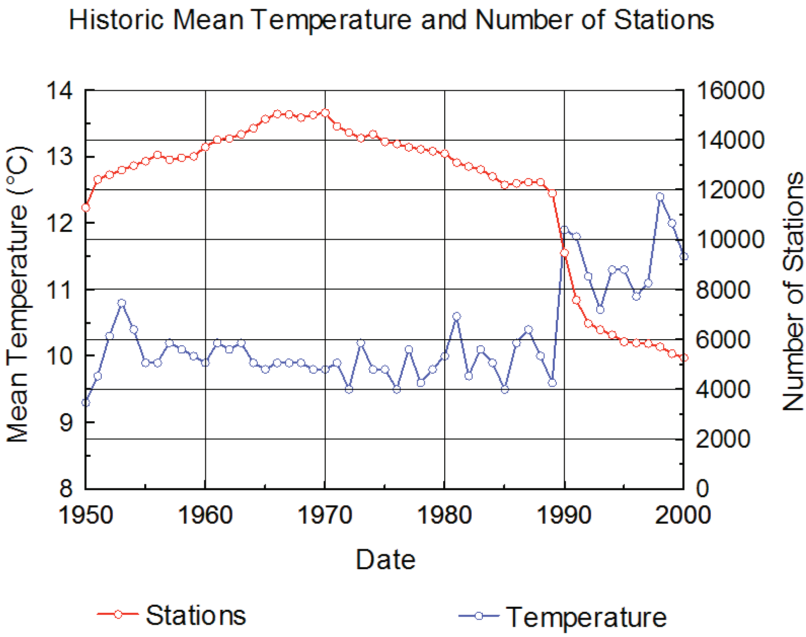

Figure 1

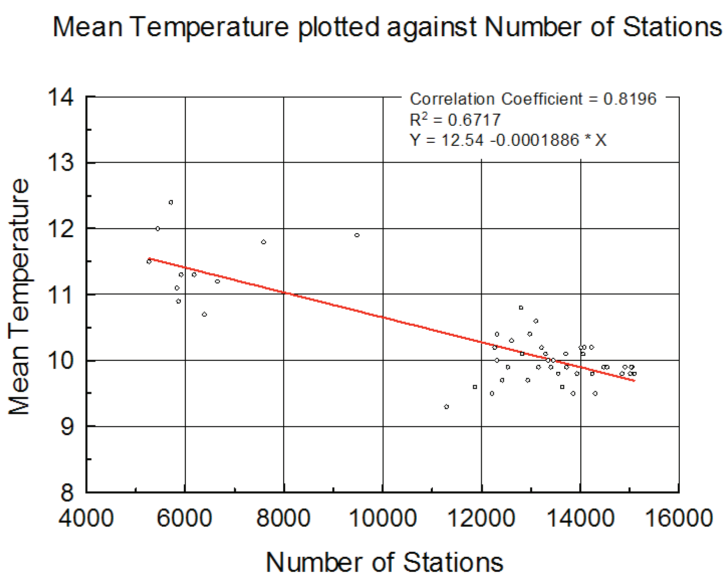

Figure 2

A linear fit has been applied yielding $R^2 = 0.67$ which suggests a reasonably strong dependence. Having determined the relationship, this can be used as a model. The station numbers can be fed back into the model as a check. The result is given below (Figure 3).

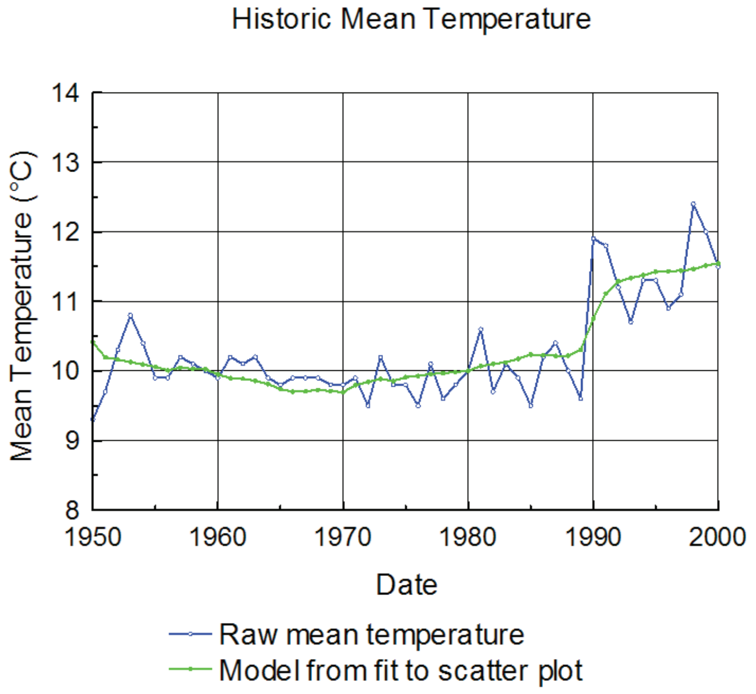

Figure 3

As can be seen, the model (green line) is a reasonable approximation to the real data.

Now we assume that the number of stations should not affect the measured temperature. Using this it is possible to exploit the model to predict what may have happened had the number of stations remained constant. The average number of stations was calculated to be 12,000. Plugging in the numbers produces the temperature history corrected for the influence of station numbers, as shown in Figure 4.

Clearly, 1990 is now the hottest year with 1998 second and 1953 third. The 1990 step has gone and if it wasn't for the peaks of 1990 and 1998 then there would probably be a slight downward temperature trend. Due to the large change in the number of stations around 1990 it is conceivable that some of the peak at that time could be the result of inter-sample temporal mismatch between the datasets. However, that should not affect the overall trend unduly.

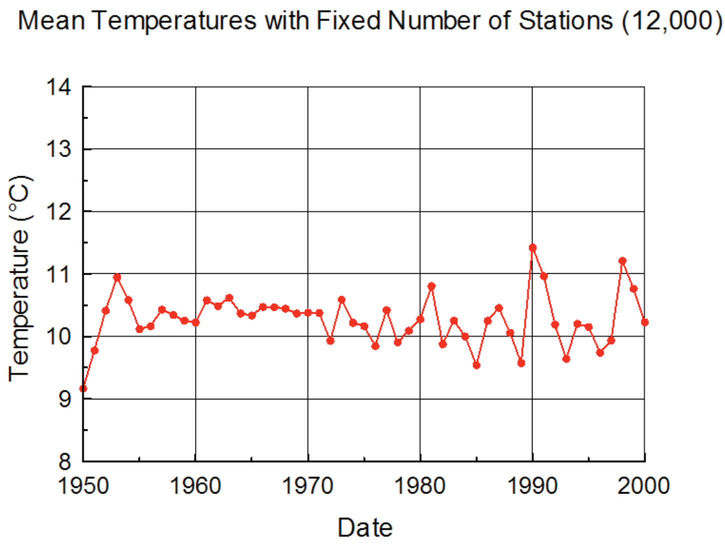

Figure 4

A comparison between the original and corrected records is given in Figure 5.

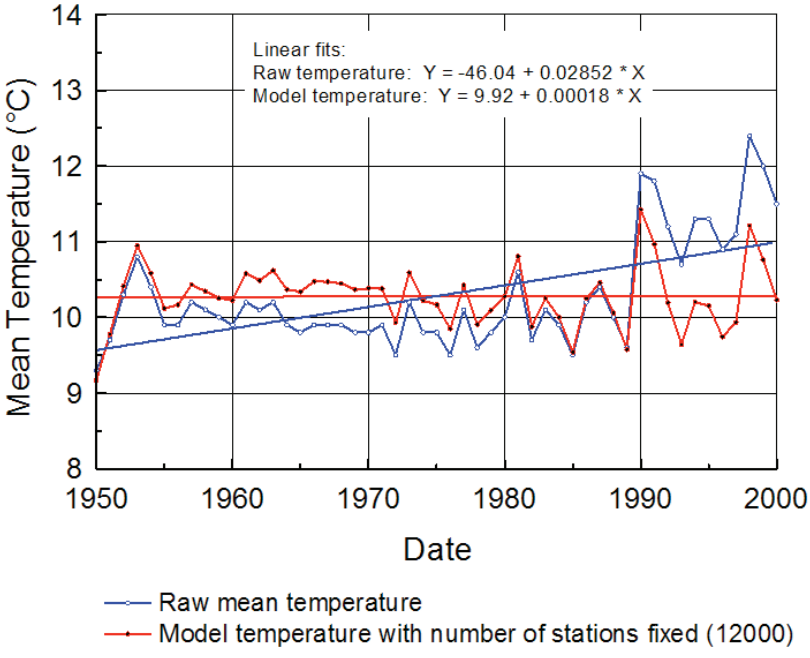

Figure 5

Having applied linear trend fits, it can be seen that the trend gradient has dropped from about +0.29 °C per decade to +0.002 °C per decade for the 50 year period.

Using this simple process it appears from the corrected data that the last decade of the twentieth century is on average no longer unusually hot although the well known peak years retain relatively high values.

Of course, there are probably better ways of correcting for this effect if the disturbances can be quantified and accounted for adequately. However, as an approximate correction, in the absence of any other information, it is probably reasonable.

3. ASSESSMENT

Here comparisons are made between satellite records and the surface station temperature records. It needs to be remembered that the satellite data is only available for a relatively short period of time and the measurements are not in the absolutely strictest sense, directly comparable. Thus the following is only a guide to the validity of such a correction. The raw and corrected data (fixed number of stations) with associated linear fits are presented in Figure 6 & Figure 7 respectively:

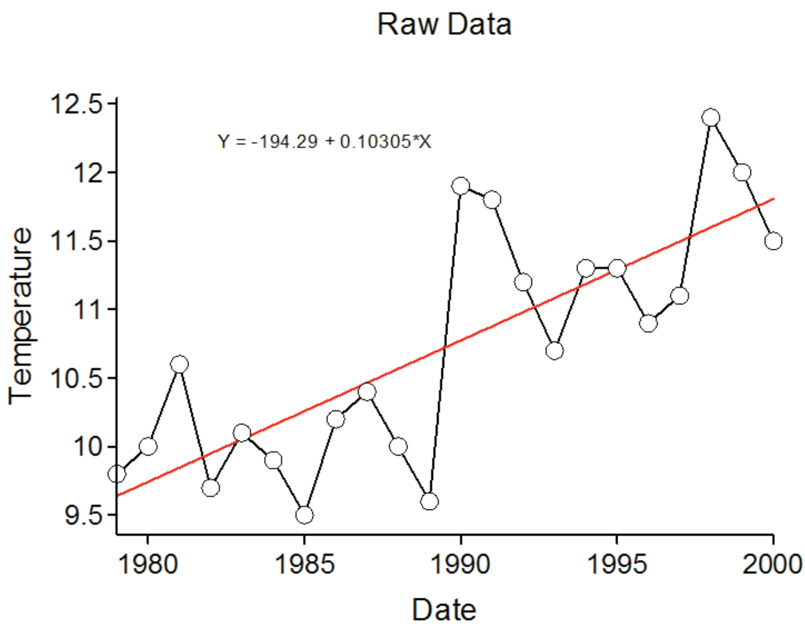

Figure 6

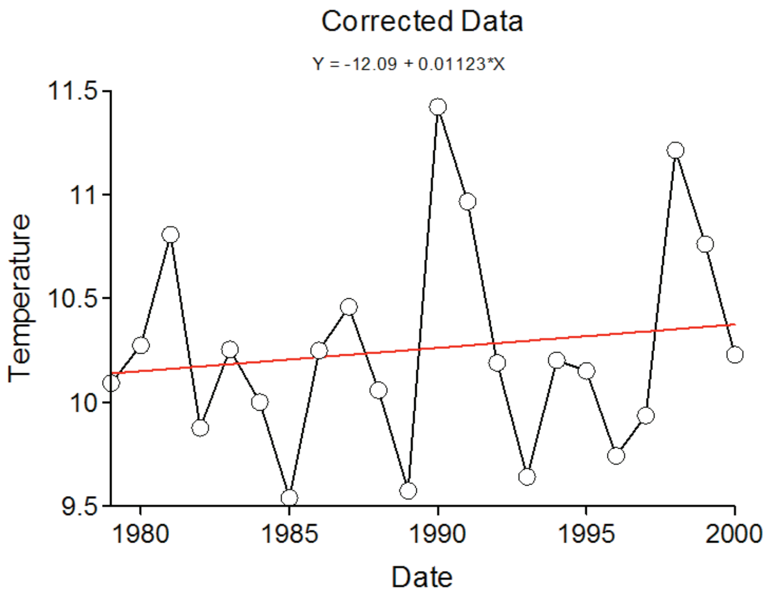

Figure 7

RSS (2) and UAH (3) satellite data are used for comparison. The overlapping date range in these datasets is 1979 to 2000, and thus only that period can be assessed. Plots of the data, along with their associated linear trends are given in Figure 8 & Figure 9.

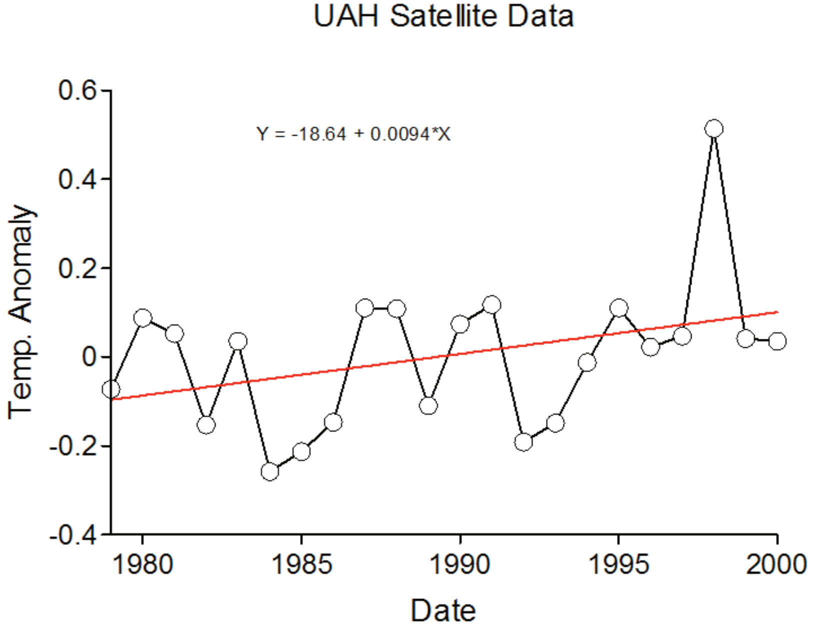

Figure 8

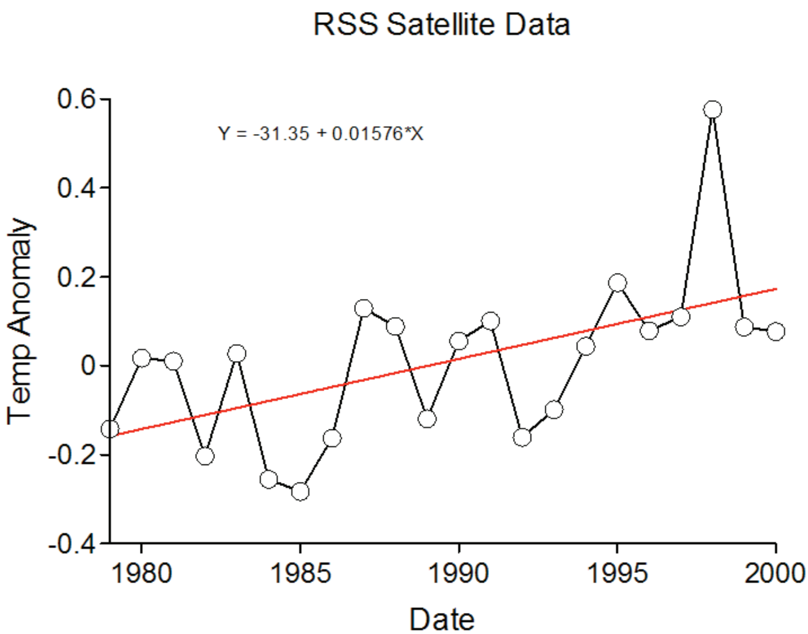

Figure 9

Clearly, the surface station record contains significantly greater variability than the satellite versions. This may be the result of spatial effects, possibly due to ocean thermal damping; however, explaining it is beyond the scope of this communication.

Least squares linear fits from the above plots are summarised below:

Surface Stations (Raw)	1.030°C per decade
Surface Stations (Corrected)	0.112°C per decade
RSS Satellite	0.158°C per decade
UAH Satellite	0.094°C per decade

Correcting the surface station record with the method outlined reduces the temperature trend over this period by almost an order of magnitude. Moreover, the trend of the corrected data is between that of the two satellite measurements, giving confidence that the technique appears to produce reasonable results.

Could this be coincidental?

4. CONCLUSION

There appears to be a signal in the raw historic temperature record that is related to the number of weather stations. Using this simple technique to remove the perturbing influence reveals a significantly reduced temperature trend that is notably different from the historic records presented by CRU, GISS, etc. The corrected data has a linear trend that is in good agreement with satellite data in the overlapping period. Whether or not this is a coincidence requires further investigation.

REFERENCES

1. Ross McKitrick's compiled data: http://www.uoguelph.ca/~rmckitri/research/stations.v.T.xls
2. RSS (Remote Sensing Systems, Santa Rosa) data (RSS Data Version 3.1): ftp://ftp.ssmi.com/msu/monthly_time_series/rss_monthly_msu_amsu_channel_tlt_anomalies_land_and_ocean_v03_1.txt
3. UAH (University of Alabama, Huntsville) data for Lower Troposphere: http://vortex.nsstc.uah.edu/public/msu/t2lt/tltglhmam_5.2

Cover photo of the Ceahlau Toaca weather station in Romania after a winter storm uploaded by ionutz at wunderground.com.