Response to Gavin Schmidt on the Integrity of the Global Data Bases

By Joseph D'Aleo, CCM, AMS Fellow

Gavin Schmidt wrote a hit piece on Tuesday night's the Lou Dobbs show that featured a discussion of global cooling. He objected that he was the sole voice on the side on global warming. He said the piece was not fair and balanced. To Gavin and the other alarmists, it appears, a piece that is fair and balanced can make no mention of any other opinion except that carbon dioxide is causing global warming and action is needed now and will deliver gain and no pain, something the one sided media coverage has gotten them used to over the years.

Yet there is considerable evidence that man is not responsible anything but local climate change through urbanization (the world's population has grown from 1.5 to 6.5 billion people since 1900) and locally and regionally through land use changes like deforestation, forestation and irrigation. Consider for example United States temperatures cooled from the end of WWII in the 1940s to the late 1970s, warmed from 1979 to 1998 then leveled off and fell. That means 5 of the last 7 decades has seen cooling during the most rapid industrialization period, the post WWII boom. CO2 has increased the whole time. This does not give cause to believe CO2 is the primary climate driver as the alarmists want you to believe. We have shown you on this site many examples of how other factors like ENSO, the Pacific Decadal Oscillation, Atlantic Multidecadal Oscillation and cycles on the sun and volcanism are much better correlated with the observed temperatures than CO2 and are the real drivers.

In the Lou Dobbs' segment, Ines Ferre, the reporter who did an excellent job talking to many scientists on both sides of the issue reported: *"three independent research groups concluded that the average global temperature in 2008 was the ninth or tenth warmest since 1850, but also since the coldest since the turn of the 21st century.*

Lou also noted that according to NOAA 7 of the 8 warmest years on record have occurred since 2001. Lou asked me if I quibbled with that data. I replied:

"Yes, I do. In fact, if you look at the satellite data, which is the most reliable data, the best coverage of the globe, 2008 was the 14th coldest in 30 years. That doesn't jive with the tenth warmest in 159 years in the Hadley data set or 113 or 114 years in the NOAA data set. Those global data sets are contaminated by the fact that two-thirds of the globe's stations dropped out in 1990. Most of them rural and they performed no urban adjustment. And, Lou, you know, and the people in your studio know that if they live in the suburbs of New York City, it's a lot colder in rural areas than in the city. Now we have more urban effect in those numbers reflecting that show up in that enhanced or exaggerated warming in the global data set."

Gavin took umbrage at this slap at the data centers:

"D'Aleo is misdirecting through his teeth here. He knows that the satellite analyses have more variability over ENSO cycles than the surface records, he also knows that urban heat island effects are corrected for in the surface records, and he also knows that this doesn't effect ocean temperatures, and that the station dropping out doesn't affect the trends at all (you can do the same analysis with only stations that remained and it makes no difference). Pure disinformation."

Really Gavin? Lets take a look. Ironically just this week we featured a story highlighting that exact issue and the discrepancy that exists between the satellite data and the surface based station statistics. But let's start from the beginning.

Urban Heat Island

We summarized the Urban Heat Island (UHI) effect and the treatment of the UHI in assessing the climate trends by Hadley, NOAA and the IPCC in this pdf. We will summarize some of the points and then talk about GISS's (Goddard Institute for Space Studies) so called UHI adjustments.

There is no real dispute that weather data from cities, as collected by meteorological stations, is contaminated by urban heat island (UHI) bias, and that this has to be removed to identify climatic trends. Disputes center on whether corrections applied by the researchers on whom the IPCC relies for generating its climatic data are adequate for removing contamination.

And the UHI effect occurs not only for big cities but also for smaller towns that grow. For example, Oke (1973) and Hoyt (2002) have shown a town of 1000 could see a warming of 2C or 3F) especially in winter. Hinkel et al (2003) showed even the village of Barrow, Alaska with a population of 4600 has shown a warming of 2.2C (3.4F) in winter over surrounding rural areas.

Oke found that the warming can be directly tied to population. Oke (1973) found evidence that the UHI (in $_oC$) increases according to the formula

$$UHI = 0.73 \log_{10} POP$$

where *pop* denotes population. This means that a village with a population of 10 has a warm bias of 0.73C, a village with 100 has a warm bias of 1.46 C, a town with a population of 1000 people has a warm bias of 2.2 C, and a large city with a million people has a warm bias of 4.4 C (Oke, 1973).

Tom Karl et al 1988 employed a similar scheme for the first USHCN data base (probably the best data set available at any time). He noted that at that time the national climate network was predominantly rural or small towns population below 25,000 (as of 1980 census) and that a UHI effect was clearly evident.

He noted that the UHI warming was clearly greatest with respect to minimum temperatures with a slight cooling shown for maxima.

TABLE 7. The effect of urbanization on the annual temperature (°C) in the United States with respect to a station located in a rural environment (population < 2000). Each element is derived independently of other elements [(max + min)/2 = avg].

Population	Average	Maximum	Minimum	Diurnal temperature range
2000	0.06	−0.01	0.12	−0.13
5000	0.08	−0.02	0.16	−0.18
10 000	0.11	−0.02	0.22	−0.24
20 000	0.16	−0.03	0.32	−0.35
50 000	0.24	−0.05	0.48	−0.53
100 000	0.32	−0.07	0.63	−0.70
200 000	0.44	−0.09	0.87	−0.96
500 000	0.67	−0.14	1.33	−1.47
1 000 000	0.91	−0.20	1.81	−2.00
2 000 000	1.25	−0.27	2.48	−2.74
5 000 000	1.88	−0.40	3.73	−4.12
10 000 000	2.57	−0.55	5.10	−5.63

From: Karl, T.R., H.F. Diaz, and G. Kukla, 1988: Urbanization: its detection and effect in the United States climate record, J. Climate, 1, 1099-1123.

He did note that because of the bias towards smaller town or rural stations, the net contamination by UHI on the regional or national scale was still relatively small but that significant anomalies showed up in rapidly growing population centers (and thus should be addressed which USHCNv1 did).

In 2007, NCDC in its version 2 of USHCN inexplicably removed the Karl UHI adjustment and substituted a CHANGE POINT ALGORITHM that looks for sudden shifts (discontinuities). It is suited for finding site moves or local land use changes (putting a paved road or building next to sensors or shelters) but not the slow ramp up characteristic of a growing town or city. See the result of making this change on the global annual temperature in this graph:

USHCN V2-V1

Note how this change introduced a warming in the early 1900s cold period, a warming in the middle century cold period and significant boost in the late 1990s and first half of this decade. The net result is to make the recent warm cycle max more important relative to the early century max in the 1930s. The change can be seen clearly in this animation. (courtesy of climate skeptic). This serves to virtually eliminate the inconvenient finding that the 1930s was as warm or warmer than the recent warmth much as Mann's (Gavin's partner in crime at RC) hockey stick conveniently removed the hard to explain away Medieval Warm Period, shown clearly by Mcintyre and Mckitrick and then Wegman backed up by good detective work by the Idsos at CO2 Science here to be very real.

THE GLOBAL DATA BASES

Jones et al 1990 (Hadley) had concluded that UHI bias in gridded data could be capped at 0.05 deg C (not per decade, per century). Peterson et al 1999 agreed with the conclusions of Jones et al. (1990) and Easterling et al. (1997) that urban effects on 20th century globally and hemispherically averaged land air temperature time-series do not exceed about 0.05°C over the period 1900 to 1990. Peterson (2003) and Parker (2004) argue urban adjustment thus is not necessary.

NCDC for GHCN regarded urban areas as those with populations exceeding 10,000. Remember, Oke, the 2008 winner of the AMS Helmut Landsberg Award for his work on UHI showed warming for even 1,000 and smaller population towns.

All these ignore the findings of the more than half a dozen peer reviewed papers in this pdf that the lack of adequate UHI and local land use change adjustment could account for up to 50% of the warming since 1900 (the exaggeration I talked about with Lou Dobbs)

NASA GISS

Is NASA better? Steve McIntyre has taken an in depth look at the data adjustments made in the NASA GISS data set and the findings are summarized very well in the following pdf Correct the Correction by Ken Gregory..

NASA's Goddard Institute of Space Studies (GISS) publishes a global temperature index.

The temperature record is contaminated by the effects of urban development and land use changes. NASA applies an "urbanization adjustment" to adjust the temperature histories to eliminate these effects. The resulting GISS temperature index is supposed to represent what the temperatures would have been in the absence of urbanization and land use changes. Most scientists assume that these adjustments are done correctly.

An audit by researcher Steve McIntyre reveals that NASA has made urban adjustments of temperature data in its GISS temperature record in the wrong direction, with almost as many urban areas adjusted to show more warming as less warming.

The audit shows that 74% of the USA stations are adjusted, but only 37% of the rest of the world stations are adjusted. There are almost as many negative adjustments as positive adjustments in the rest of world stations. The contiguous USA land area is only about 7% of the world surface area, so the other stations have a much larger effect on the global temperature index.
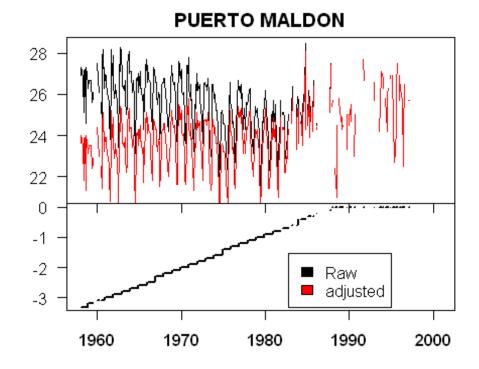
GISS uses two different methods of categorizing stations as rural or urban. Stations in the USA, southern Canada and northern Mexico are classified based on the amount of night time light measured by satellites from the station locations. Unlit stations are classified as rural stations.

(Note: Anthony Watts and Steve McIntyre have shown some errors in classification using night lights).

Outside of the USA, southern Canada and northern Mexico, GISS uses population data to define rural stations. Hansen et al 1999 provided the following definitions for "rural", "small" and "urban": "We use the definition of Peterson et al 1997 for these categories: that is, rural areas have a recent population of less than 10,000, small towns between 10,000 and 50,000 and urban areas more than 50,000. These populations refer to approximately 1980."

The GISS sites are defined to be "rural" if the town has a population of under 10,000. Unfortunately, the population data utilized by GISS to classify the stations is out of date. Stations at cities with populations greatly exceeding 10,000 are incorrectly classified as rural. For example, in Peru, there are 13 stations classified as rural. Of these, one station is located at a city with a population of 400,000. Five stations are at cities with populations between 50,000 and 135,000.

Steve McIntyre says here, "If the supposedly "rural" comparanda are actually "urban" or "small" within the Hansen definitions, then the GISS "adjustment" ends up being an almost completely meaningless adjustment of one set of urban

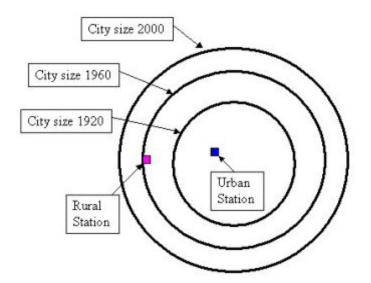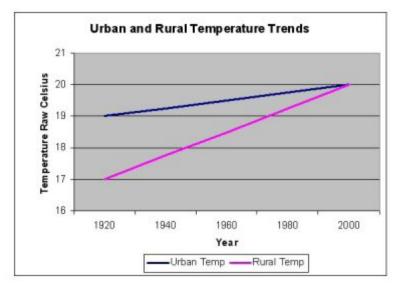values by another set of urban values. No wonder these adjustments seem so random."

Here is an example of an urban negative adjustment from Peru:

PUERTO MALDON

Note that the raw data shows no warming, but after applying the GISS urban adjustment, the adjusted data shows a significant warming trend. The adjustments are applied to reduce the past temperatures by up to 3 degrees Celsius. This is a very large adjustment when compared to the total warming of the twentieth century of 0.6 Celsius estimated by the IPCC.

The data shows that the stations classified as rural are almost as likely to have as much a warming trend as urban stations. Why would almost half of the urban stations have lower warming trends than the nearby rural stations? It is very unlikely that heat sources near urban stations were gradually removed.

A population increase of 500 in a town of 2000 people would have a much larger effect on temperature measurements than the same increase in a city of 500,000 people. A city with a growing population generally increases its area. A temperature station inside the city would be little affected by the expansion of the suburbs. However, a temperature station located just outside a city would be greatly affected by the city expanding around the station. This effect is shown in the following diagram.

Urban and Rural Temperature Trends

A hypothetical urban station is shown located in a city and a rural station is located outside the city in the year 1920. By 1960, the city has grown out to the rural station. The city growth has little effect on the urban station, but a much larger affect on the rural station. By 2000, the rural station is completely surrounded by the city, so it has the same temperature as the urban station.

As indicated in the graph, the unadjusted rural temperature trend is much greater than the urban station trend. ***According to the GISS urban adjustment procedure, the urban station trend is increased to match the rural station trend by reducing the past temperatures.***

A proper urban correction algorithm would reduce the warming trends of both stations to make an adjusted temperature record represent what would have happened if nobody lived near the stations.

Ross McKitrick and Patrick Michaels published a paper in December 2007 that shows a strong correlation between urbanization indicators and the "urban adjusted" temperatures, indicating that the adjustments are inadequate. The conclusion is: Fully correcting the surface temperature data for "non-climatic effects reduce the estimated 1980-2002 global average temperature trend over land by about half."

Dutch meteorologists, Jos de Laat and Ahilleas Maurellis, showed (2006) that climate models predict there should be no correlation between the spatial pattern of warming in climate data and the spatial pattern of industrial development. But they found this correlation exists, and is statistically significant. They also concluded it adds a large upward bias to the measured global warming trend.
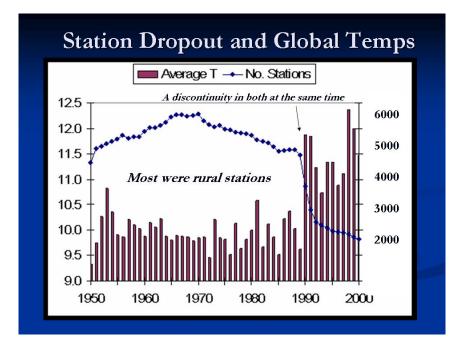
These studies convincingly show that the urban corrections fail to correct for the effects of urbanization, but do not indicate why the corrections fail. The audit of GISS urban adjustments by Steve McIntyre, shows why the corrections failed. Governments around the world intend to spend billions of dollars based on the belief that the temperature indexes are properly corrected for the effects of urbanization

NASA GISS should be commended for attempting to correct for urbanization, but their efforts may be best described as a Chinese fire drill with lots of movement but no resulting impact.
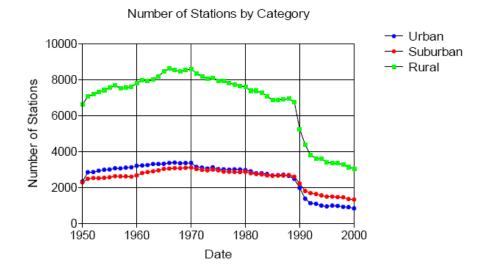
STATION DROPOUT AND OTHER INTEGRITY ISSUES

Another important issue is station dropout. Over 2/3rds of the world's stations, many rural areas in the FSU stopped reporting around 1990. Climatologist David Legates at the University of Delaware are prepared this animation. See the lights go out in 1990.
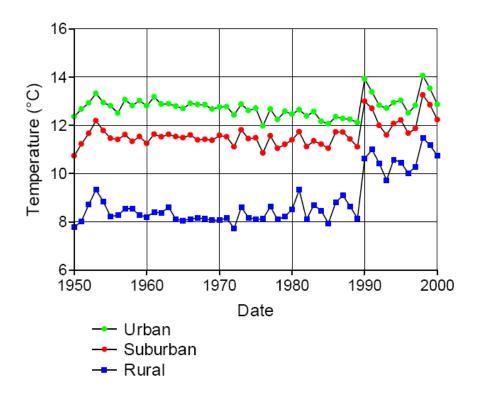
The animation above shows that Siberia is one area with the biggest change

Station Dropout and Global Temps

In the chart above you see the number of global sites drops off coincident with the sudden rise in mean of all remaining stations. Jonathan Drake did this analysis below of the broken down by rural, suburban and urban categories. It clearly shows the big dropoff of rural stations.

Average temperatures for the station categories jumped when the stations dropped out, suggesting that it was mainly due to colder stations were no longer in the record.
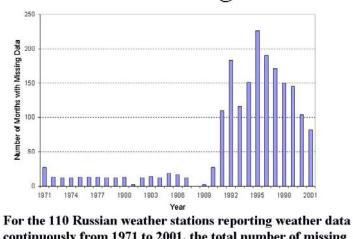
Station Temperatures by Category

The global data bases all compile data into grid squares globally and calculate the temperatures for the box using data from stations (already adjusted) within that box or using the closest stations weighted by distance in nearby boxes. Thus a grid square, which at one time had rural stations now gone, will find temperatures increasingly determined by urban areas in that box or distant boxes. This is why the global data bases suggest the greatest warming has been in Siberia where the greatest dropout occurred.

Also a factor is that in the Soviet era, the cities and towns temperatures reportedly determined allocations of funds and fuel, so it is believed that the cold was exaggerated in the past which introduced a seeming warming when more honest measurements began to be made.
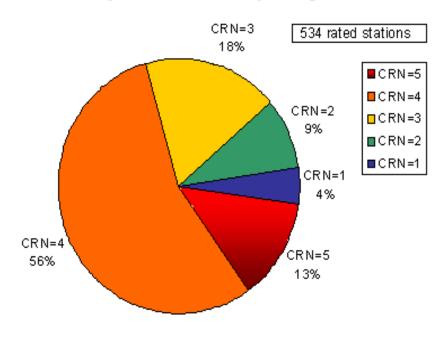
In addition to insufficient or no UHI adjustment, and station dropout, there has been a tenfold increase in missing months of data in places like the Soviet Union.

Number of Missing Months

For the 110 Russian weather stations reporting weather data continuously from 1971 to 2001, the total number of missing monthly observations each year (McKitrick and Michaels)

There has been the use of new instruments with known warm biases that go uncorrected, and of course bad siting of instruments. As of 2008, when nearly half way through the 1200 USHCN stations, Anthony Watts and his team of volunteers found only 4% met the governments own standards and 69% were poor or very poorly sited (with errors of greater than 2C).

Surveyed CRN Site Quality Rating

NOAA and NASA will tell you their adjustments take care of that. Can you trust that?

OCEAN TEMPERATURES

Finally Gavin mentions oceans as if NASA had them covered. The world is 70% ocean. Hadley only trusts their own merchant ship data mainly travelling select northern hemisphere routes. They have virtually no ocean data from the southern hemisphere's oceans (80% of that hemisphere).

Notable too is the gradual change from buckets to ship intakes for measuring ocean temperatures. Different sampling levels will make results slightly different. How to adjust for this introduced difference and obtain comparable data sets has yet to be resolved adequately, especially since the transition occurred over many decades.
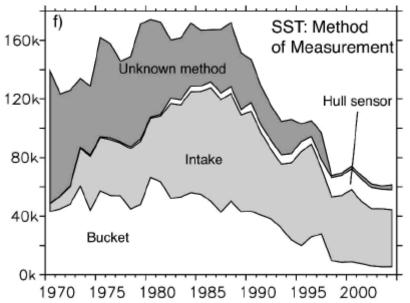

Chart from Kent (Kent, E. C., S. D. Woodruff, and D. I. Berry. 2007. Metadata from WMO Publication No. 47 and an Assessment of Voluntary Observing Ship Observation Heights in ICOADS. Journal of Atmospheric and Oceanic Technology 24, no. 2: 214-234.)
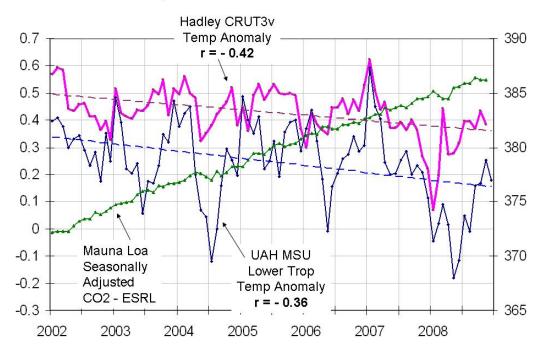
We have reanalysis data based on reconstructions from ships and buoys (subject to some of the same adjustment issues) and satellites which see only ocean surface skin temperatures but are hampered by cloud cover.

See much more on these data integrity issues here.

Please note, I thought it was too overwhelming and impossible to bring all these issues up, given the time constraints with a panel of 3 given 4 minutes to answer questions with the interviewer Lou Dobbs. So I stopped with the station dropout and UHI. The next in item would have been bad siting if I had 10 seconds more.

WARMING TREND AND GREENHOUSE CONNECTION INTACT

One last point. Gavin in his own statement makes the claims the warming trend remains intact and so by extension the greenhouse warming theory. Here is a plot of the temperatures from Hadley and UAH MSU since 2002 relative to CO2. Note the strong downtrend and negative correlations of temperature with the CO2 (seasonally adjusted from NOAA ESRL). Correlation is most negative with the Hadley data, the long term trend of which was shown on the Lou Dobbs show.

Hadley CRUT3v and UAH MSU vs CO2

And here courtesy of SPPI, a blend of four data sites (satellite and surface) since 2001 versus the IPCC model forecasts.

Eight straight years' global temperature downtrend: The authoritative SPPI composite index of global mean surface temperature anomalies, taking the mean of two surface and two satellite datasets, shows a pronounced downtrend for eight full years. Not one of the climate models relied upon by the IPCC had predicted this downturn. The pink region shows the IPCC's projected rates of temperature increase: the thick red straight line shows the least-squares linear regression on the composite temperature anomalies.

IF I WERE IN CHARGE

I would spend some of the money 'earmarked' for climate change research or carbon control and invest in doing a bottom up reanalysis of the global data. It would include a UHI scheme like Oke or Karl 1988 that adjusts the raw data, already corrected for time of observation and site moves. I would assemble a staff recreating the metadata of station siting and population for the world's stations and assemble a team to determine the best approach or blend for an ocean reanalysis (similar to what is described here). A few hundred thousand or even a million dollar investment in trying to get this right could save us spending a trillion dollars more on a non-issue at a time when real issues are already threatening our economic future.

REFERENCES

Hansen, J. et al., 1999: "GISS analysis of surface temperature change", Journal of Geophysical Research, Volume 104.

Hansen, J. et al , 2001: "A closer look at United States and global surface temperature change", Journal of Geophysical Research, Volume 106.

Karl, T.R., H.F. Diaz, and G. Kukla, 1988: Urbanization: Its Detection and Effect in the United States Climate Record. *J. Climate*, **1**, 1099–1123.

McKitrick, R., Michaels, P., 2007: "Quantifying the influence of anthropogenic surface processes and inhomogeneities on gridded global climate data", December 2007, Journal of Geophysical Research, Volume 112.
http://www.uoguelph.ca/%7Ermckitri/research/jgr07/M&M.JGRDec07.pdf

Laaat de, A.T.J., Maurellis, A.N., 2006: "Further evidence for influence of surface processes on lower tropospheric and surface temperature trends", International J. of Climatology, 26, p. 897-913, 2006. http://www.knmi.nl/~laatdej/2006joc1292.pdf

Additional references can be found [here.](#)