Review of my paper for non-scientists – Part 1

Should We Worry About the Earth's Calculated Warming at 0.7^{O}C Over Last the Last 100 Years When the Observed Daily Variations Over the Last 161 Years Can Be as High as 24^{O}C?

Darko Butina

The paper is unique and novel in its approach to man-made global warming in many respects: it is written by experimental scientists; it is published in journal that deals with data analysis and pattern recognition of data generated by a physical instrument; it treats the Earth atmosphere as a system where everything is local and nothing is global; and it is the first paper that looks for temperature patterns in the data that is generated by the instrument designed to and used by experimental scientists since early 1700s – calibrated thermometer. What is also unique is that every single graph and number that I have reported in the paper can be reproduced and validated by reader using data that is in public domain and analyse that data using simple excel worksheet. There are two main conclusions made in the paper:

1. That the global warming does not exists in thermometer data since it is impossible to declare one year either warmer or colder than any other year
2. That the Hockey Stick scenario does not exists in thermometer data and therefore it must be an artefact observed in a purely theoretical space of non-existing annual global temperatures

The paper is long 20 pages and analyses in great details a single weather station of daily data, dataset collected at Armagh Observatory (UK) between 1844 to 2004, one of the very few datasets in public domain that have not been either destroyed, corrupted or endlessly re-adjusted by the curators of the global thermometer data at East Anglia University or NASA.

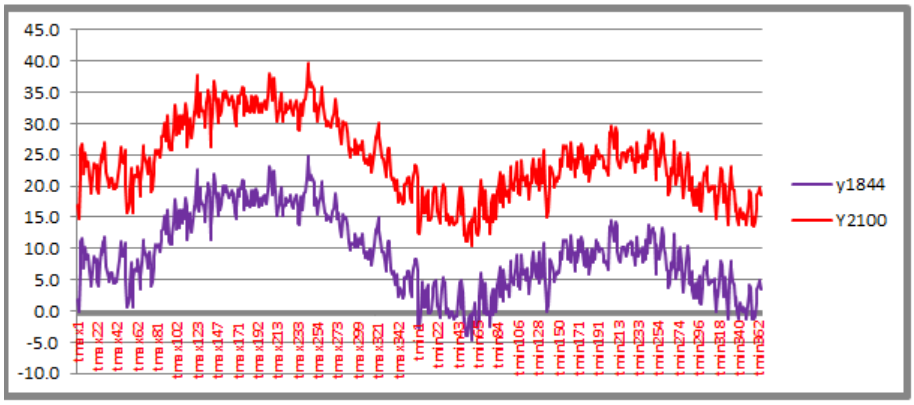
Before we start to analyse this paper, few points need to be made about experimental sciences for my paper to be properly understood. ALL our knowledge and understanding about the physical world around us comes from analysing and learning from data that has been generated by an experiment and measured or recorded by a physical instrument. Let me demonstrate this point by a very simple example of what happens when we record air temperature by some fixed-to-ground thermometer:

Temperature	Where (Grid Point)	Date	Time	Instrument	Symbol
15.1	X1,Y1,Z1	dd/mm/yyyy	ss/mm/hh	Thermometer	^{O}C

Thermometer reading of 15.1 has several links attached to it that cannot be broken: it is linked to a unique grid point, unique date and time stamp, unique instrument – thermometer and that thermometer to unique symbol (^{O}C). So if someone wants to analyse any temperature trends those trends have to come from thermometer readings; it follows that if thermometer to be used is calibrated using Celsius scale, no datapoint can be older than 1743, follow link to Anders Celsius. Since we know for a fact that the annual temperature ranges will depend on the location of that thermometer, and since mixing different datasets are not allowed in experimental sciences, it follows that if there are, say 6000 weather stations (or fixed thermometers) in existence and across the globe, the first step **before** raising an alarm would be to analyse and report temperature patterns for every single weather station. That was what I was expecting to see when I started to look into this man-made global warming hysteria three years ago, following the revelations of Climategate affair. But I could not find a single published paper that uses thermometer-based data. So we have situation that the alarm has been raised, the whole world alarmed, suicidal economic policies have been taken while totally

ignoring the data generated by the only instrument that has been invented to measure temperature – the thermometer. Instead, thousands of publications have been written looking for temperature trends in purely theoretical space that does not and cannot exist, the space of annual global temperatures. Two key papers have been published earlier, both arguing and explaining why global temperature as a single number does not exist, Essex et al., in 2007, using statistical arguments and written by recognized statisticians, while Kramm and Dlugi in 2011 have shown why the Earth's atmosphere cannot be treated as a homogeneous system but should be perceived as a network of local temperature systems, from astrophysics point of view.

The starting point for my paper was based on facts that it is impossible to have arguments and ambiguity when it comes to thermometer. If you have two readings, only one outcome is possible: T2>T1 or T2<T1 or T2=T1. So if one wants, for some bizarre reason, to compare two annual patterns then one year can be unequivocally declared as warmer (IPCC claims that the warming is unequivocal – not me!) **only** if each daily reading of that year is larger than each corresponding daily reading of another year:



This artificially created graph above has real year in Tmax-Tmin space from the Armagh dataset while the year 'y2100' was result of adding 15C to each daily reading of y1844. My point here is that everyone seeing that graph would come up with an identical conclusion – y2100 is unambiguously warmer than y1844. So my perfectly valid question was – why would anyone went to trouble to invent something that does not exists while ignore obvious source of temperature data – the thermometer data itself? My 40 years of experience in experimental sciences offered a most obvious answer to that question – because nothing alarming could be found in thermometer-based data. There is a quite simple rule when it comes to interpretation of data – ***if more than a single conclusion could be made about any given dataset it means one of two things: either that dataset is of right kind but more data is needed to understand the data, or the data is of the wrong kind.*** Every single graph and number that is found in my paper can be independently reproduced and validated and therefore the thermometer data is the right-kind of data to use but we need more of it to fully understand temperature patterns observed on our planet. The opposite is true when we look at the calculated and not measured data called annual global temperatures.

When you dissect this very complex looking scientific problem of man-made global warming to its basic components, what you find is that ***the whole concept of global warming and climate change has nothing to do with science but everything to do with a very desperate attempt to connect temperatures with a few molecules of CO2 that have been generated by burning fossil fuels, while ignoring vast majority of CO2 molecules that have been generated by nature itself.*** It must follow that if those alarming trends could not be found in thermometer data, than that data must be removed

and new data created, type of data that cannot be either proved wrong or right and allow proponents of man-made global warming to generate any trend they need and to enable them to claim that they know everything about everything when it comes to our planet. But, the only problem with that approach is that you cannot cheat in experimental sciences and slowly but steadily, retired scientists like me, with bit of free time will start to look into this problem and use their respective expertise to critically evaluate the supposed science behind this man-made movement.

So even before I started to collect daily data that are available in public domain, I was almost 100% confident that I will not find any alarming trends in thermometer data. And I was proven right.

Let us now start with the experimental part of the paper, the part where all the details of the dataset and dataset itself are presented. The paper is 20 pages long and all conclusions are based on detailed analysis of the Armagh (UK) dataset that covers period between 1844 and 2004. Dataset can be downloaded from the Armagh Observatory website as two sets of files, Tmax and Tmin files:

http://climate.arm.ac.uk/calibrated/airtemp/tccmax1844-2004

http://climate.arm.ac.uk/calibrated/airtemp/tccmin1844-2004

Depending on the software that one wants to use to analyse data, it is important to format all datasets in the same way. Since all commercial software expect as default data to be read in row-wise manner, reformatted Armagh dataset was created as a matrix containing 161 rows (1 row for each year) and 730 columns (1 column for each day-night readings):

Table 1. Formatted annual and day fingerprints within X,Y coordinate system. Annual Fingerprints are all in X-space, 1x730 bits long, from X=1 to X=730 (Tmax + Tmin data), while Day Fingerprints are in Y space, 1x161 bits long, from Y=1 to Y=161 (number of years)

		X1	X2	...	X365	X366	X367	.....	X730
		Tmax1	Tmax2	.....	Tmax365	Tmin1	Tmin2	...	Tmin365
Y1	y2004								
Y2	y2003								
Y3	y2002								
Y4	y2001								
Y5	y2000								
Y6	y1999								
Y7	y1998								
Y8	y1997								
....	.....								
Y161	y1844								

BTW, all the graphs and tables from my paper are presented as JPG image and once I made my paper available free of charge on my own website you will be able to match all those graphs presented in this report to the original ones in the paper.

As a result, we now have annual temperature pattern, let us call it 'annual fingerprint', as a 730-bit fingerprint with the first 365 bits assigned to Tmax 1 to Tmax 365 (Jan1 to Dec 365 daytime readings) followed by 365 bits assigned to Tmin 1 to Tmin 365 (Jan1 to Dec 365 night-time readings). So, the annual fingerprint space can be seen as a 161 (years) x 730 (daily readings) matrix. Looking at the table above column-wise, we have 'day fingerprints', each of them 161-bits long representing the history of each day-night readings over period of 161 years. Once this table is created, we need to decide what to do with the missing values and with the extra day in February in leap years. We delete that extra day in February, but with great care not to get rest of the year out of the sync. There are two options when dealing with the missing datapoint – either replace it with some calculated one or remove the whole column. The danger of replacing missing value with some calculated one is that we are contaminating instrumental data with some theoretical data, and unless we really understand that data the safest way is to remove all columns that contain even a single missing datapoint. ***Once you remove all columns with missing data you end up with 649-bit annual fingerprints, 89% of the original data, i.e. loss of 11% of total information content that is contained in that dataset, but with knowledge that the starting set is not contaminated by any calculated data and all datapoints are generated by thermometer itself.***

Now we have our table in excel, table containing 161 years of data where each year is collection of 649 day-night readings and we can ask the data that 64 million worth question: Can we detect unambiguous warming trend over 161 years at Armagh (UK) in thermometer data? Good starting point is to take difference between the youngest (2004) and the oldest (1844) annual fingerprints and display it as a histogram:

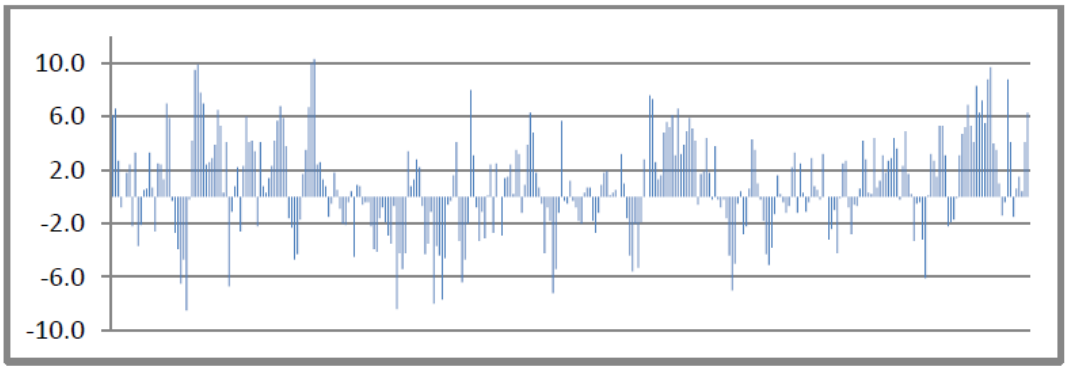


Figure 5. Difference between daily temperature readings of year 2004, the youngest year, and 1844, the oldest year. Note that the difference between two year's daily readings can switch the sign from "+" to "-"on almost daily bases.

Y-axis represents difference in temperatures between daily temperatures of years 2004 and 1844, while X-axis represents Tmax and Tmin readings.

Let me briefly digress here to make the following point – when you analyse instrumental data you have to know accuracy or error levels of the instrument that is used to generate the data. If we assume accuracy of thermometer used in 1800s at +/- 0.5C that means that for two readings to be declared as different, the difference between them should be larger than 1.0C. For example, if T2=10.0 and T1=10.8 we have to declare those two readings as same, i.e. T2=T1, since those two readings fall **within the error levels** of that instrument. If T2=10.0 and T1=20.0 then the difference is real since it is way **outside the error levels** of the instrument.

So, what is this simple graph (Figure 5) telling us? First thing to notice is that year 2004 cannot be declared either as warmer or colder than 1844 since every few days there is this switchover occurring making 2004 few days warmer than few days colder than 1844. Second thing to notice is that the size of those switchovers can be as large as 10C in one direction and 8C in another, i.e. 18C in total – way above the error levels of thermometer and therefore those switchovers are real. To make sure that those switchover patterns are not some artefacts unique to those two years, I wrote a special program (in C) to systematically compare every year to every other year (161 * 160 = 25760 comparisons) and on average each year is 50% of time warmer and 50% of time colder than any other year in the Armagh dataset. Let us zoom in into the above graph to see in more details frequency and magnitude of those switch-overs:

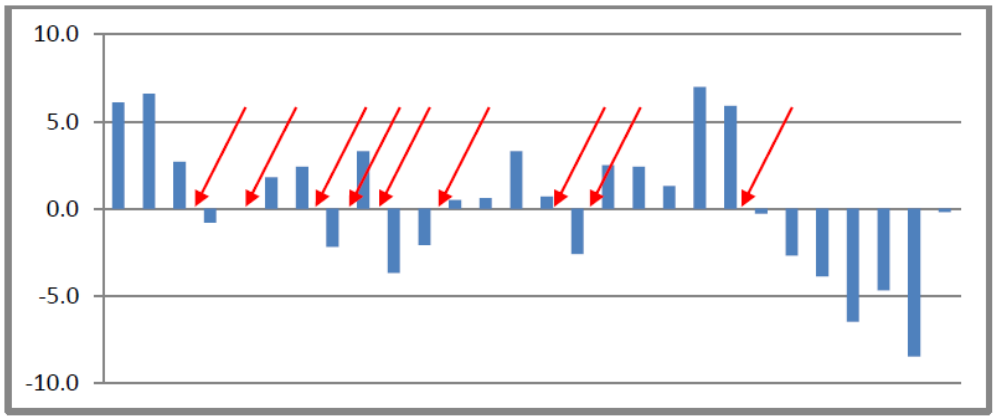
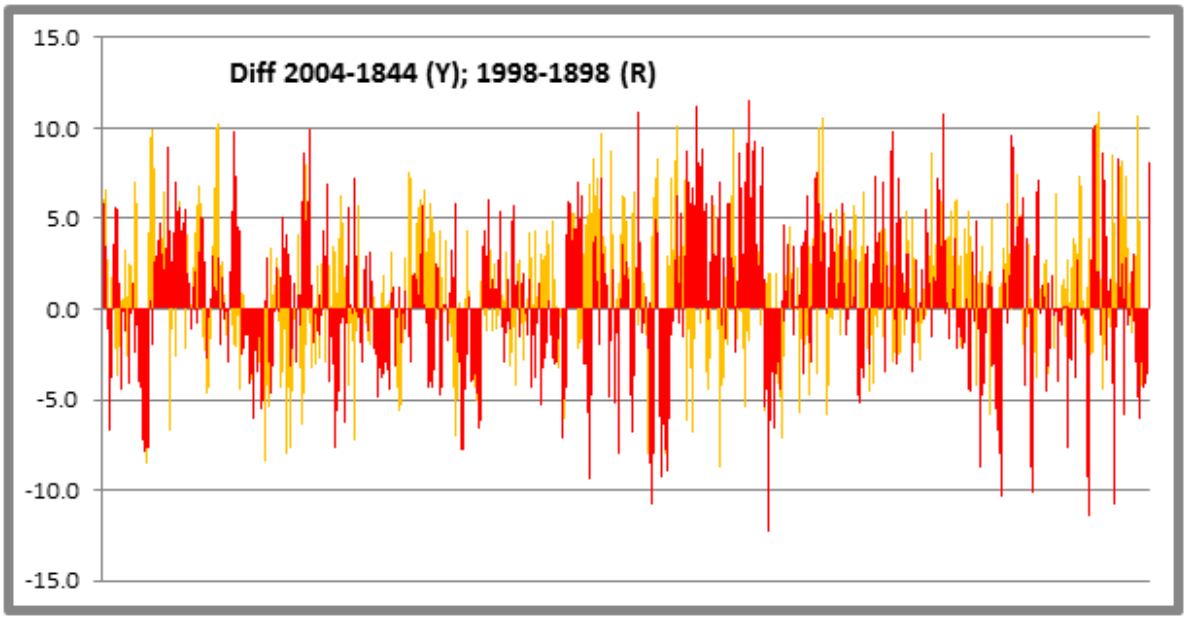


Figure 6. Nine switch-overs in January between 2004 and 1884 ranging between -8C and +8C.

As you can see, in the month of January, there are 9 switchovers between 2004 and 1844 – first 3 days (Jan 1, 2 and 3) 2004 is warmer than 1844 up to 7C, than colder, than warmer fro couple of days, then colder for 1 day and so on.

What make things even more complex is that there is no obvious pattern in terms when the switchover occur and magnitude of it, as it can be seen when two different year pairs are plotted on the same graph:



So far, all I did was to plot the original data, without any adjustment and without making any prior assumptions. I did not start this exercise to either prove or disapprove existence of global warming, but to see what the actual data is telling us. And what the thermometer is telling us is that the sheer magnitude of those apparently random and chaotic switchovers are due to natural forces that we do not understand, yet, and the anti-scientific process in which all complexity of annual temperature patterns is removed, replaced by a single number and suddenly we 'see the light' cannot be used to acquire any knowledge. If we use a simple logic the following logical construct could be made: dataset that is based on thermometer readings contains 100% of information content when it comes to temperatures. If we reduce that 730-dimensional space into a single number, we reduce the information content of that dataset from 100% to 0% - i.e. there is no information there left to gain any knowledge. Let us do the following question/answer exercise to compare two datasets – one that has day-night thermometer readings for a single year and one where 1 number represents 1 year

Q. What is total range of temperatures observed in Armagh?

A. Lowest temperature observed was -15.1C on February 7 1895, the highest temperature +30.3C recorded on July 10 1895; Total range 45.4C

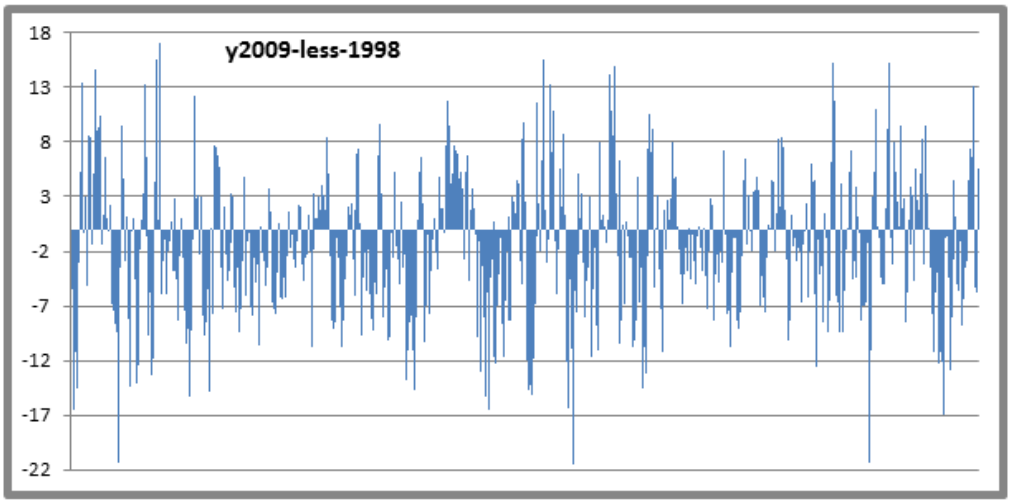
Q. What is the largest and the smallest natural fluctuations observed for individual days?

A. Day that has the most variability is May 4 (Tmax125) with total observed range of 23.8C, while day with least amount of variability is October 29 (Tmax302) with the observed range of 'only' 9.9C

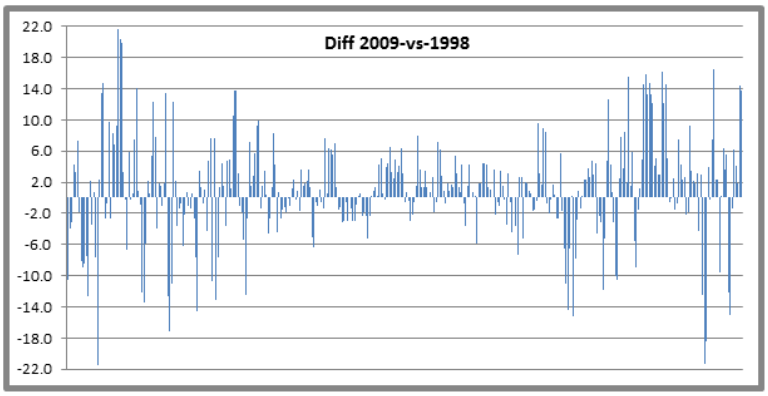
In contrast, each year is presented as a single number in the annual temperature space, number obtain by averaging all daily data to a single number, and there are not too many questions that you can ask about that single number. Actually there are none – not a single question could be asked the number that has no physical meaning! For example, if two years have identical annual average we don't know why they are the same, and if they have two different annual averages, we don't know why they are different. If we do the same exercise in daily data, we know exactly which days are moving in a same direction and which days are moving in opposite direction.

Let us now ask the most obvious question – are those patterns, or rather lack of patterns, observed in Armagh unique to UK, i.e. are they local or do they reflect some global patterns? Scientific logic would suggest that the same random/chaotic switchover patterns observed in Armagh should be observed across the globe with the only difference being the size and magnitude of those switchovers, i.e. local variations. To test that I took two annual temperature samples from two weather stations on two different continents, one in Canada and one in Australia:

Waterloo (Canada):



Melbourne (Australia):



Please note that switch-overs are the norm in UK, Canada and Australia.

Let me make a very clear statement here – the choice of Waterloo and Melbourne weather stations was driven by the ease to find weather stations with relatively easy-to-download formats and I did not get involved in method of cherry picking weather stations that fit patterns found in Armagh, as it is normal practice in man-made sciences. To prove that last point and to challenge readers to start looking into the real measured data and stop looking into non-existing and calculated data like annual global temperatures, I will offer a modest financial reward of £100.00 (UK) from my pension, to first person who finds a single example of year pair where one year has every single daily thermometer readings larger than another year. **Any weather station** that is not on permanent ice or sand (I don't know what to expect in those cases) and **any gap** between two years. Obviously, the winner will have to give the link to the original data and contact me at darkobutina@l4patterns.com to claim the award.

The way I see it, I am here in win-win situation. If nobody can find weather station that shows unambiguous warming trend, and if we keep record of all those analysed weather stations I saved the money but gain large number of additional information that should finally kill any notion of the man-made global warming hypothesis. In strictly scientific terms and using null hypothesis that either all weather stations count or none does, I have already proven that those patterns are real and observed on three different continents, and therefore prove that the global warming trend does NOT exist in thermometer data. On other hand, if someone does find clear and unambiguous warming trend in thermometer data, that work will again make the same point – all temperature patterns are local and

ONLY way to declare that trends are global is if ALL individual weather stations are showing the same trends.

This concludes this Part One report in which I explained how the first conclusion “That the global warming does not exists in thermometer data since it is impossible to declare one year either warmer or colder than any other year” in my paper has been reached.

The second conclusion in my paper which explains why the Hockey Stick scenario does not exist in thermometer data will be reported in a separate report. In Part Two report I will introduce two different bits of software, my own clustering algorithm and k Nearest Neighbours algorithm, or kNN, both used in sciences like pattern recognition, datamining and machine learning and apply them to annual temperature patterns observed in Armagh. The overall conclusions will obviously be the same as we have reached so far, but I will demonstrate how the observed differences between different annual patterns can be quantified and how we can use those computational tools to detect ‘extreme’ or unusual annual temperature patterns, like annual patterns of 1947 which is the most unique not only in Armagh but also in the rest of UK.