American Journal of Environmental Engineering and Science
2018; 5(2): 34-38
http://www.aascit.org/journal/ajees
ISSN: 2381-1153 (Print); ISSN: 2381-1161 (Online)

American Association for Science and Technology

On the Statistical Methods for the Conclusion on the Climate Change

Vladimir Stepanovich Mukha

Department of Automated Data Processing Systems, Belarusian State University of Informatics and Radioelectronics, Minsk, Republic of Belarus

Email address

mukha@bsuir.by

Citation

Vladimir Stepanovich Mukha. On the Statistical Methods for the Conclusion on the Climate Change. *American Journal of Environmental Engineering and Science.* Vol. 5, No. 2, 2018, pp. 34-38.

Received: April 4, 2018; **Accepted**: April 28, 2018; **Published**: June 1, 2018

Abstract: Nowadays it is suggested, that the global warming is significant problem. There is a natural greenhouse effect which keeps the Earth warmer than it would otherwise be. This effect is enhanced by human activities and can lead to global warming and climate change. The Intergovernmental Panel on Climate Change (IPCC) publishes the reports of the state of knowledge on climate change in which there are also the instrumental evidences of the Earth warming. However, the scientific methods used for these conclusions are not mentioned. In this regard, there is a large number of "climate skeptics" who question both the fact of global warming and the role of the human in this process. In this work the importance of using instrumental scientific methods for the conclusions about climate change is emphasized. The analysis of the yearly mean value of the atmospheric temperature on the meteorological station Minsk (Belarus) over the last 20 years is considered. As a scientific method the statistical theory of regression analysis is used. At first the yearly mean values of the temperature on the base of the measurements of the temperature is calculated, then it is approximated by linear time-dependent regression function. Standard regression analysis procedures allow doing inference about significance of the linear dependence. In our case the linear stochastic dependence the yearly mean temperature on the time (empirical regression function) over the last 20 years has a slight tendency to increase. However, using the both statistical tests on the significant of the parameters and on the linearity of the regression function shows that the trend of the mean temperature is insignificant. The similar analysis shows that the linear approximation of the monthly mean temperature increases slightly for some months and decreases slightly for others, but these changes are insignificant.

Keywords: Regression Analysis, Hypothesis Testing, Greenhouse Effect, Global Warming, Climate Change, Atmospheric Temperature

1. Introduction

It is suggested, that the global warming is a real and significant problem. For the assessment of this problem the Intergovernmental Panel on Climate Change (IPCC) was established in 1988. The IPCC have published five assessments reports of the state of knowledge on climate change. The fifth report was published in 2013–2014.

In the first assessment report, the existence of the natural greenhouse effect was stated as a starting point. It is suggested, that this effect is enhanced by human activities and can lead to global warming and climate change: "We are certain of the following: there is a natural greenhouse effect which already keeps the Earth warmer than it would otherwise be; emissions resulting from human activities are substantially increasing the atmospheric concentrations of the greenhouse gases carbon dioxide, methane, chlorofluorocarbons (CFCs) and nitrous oxide. These increases will enhance the greenhouse effect, resulting on average in an additional warming of the Earth surface. The main greenhouse gas, water vapor, will increase in response to global warming and further enhance it" [1].

In the first assessment report there is the figure that shows current estimates of smoothed global-mean surface temperature over land and ocean since 1860 (figure 1). The conclusion is that "…we believe that a real warming of the globe of 0.3°C – 0.6°C has taken place over the last century, any bias due to urbanization is likely to be less than 0.05°C" [1].

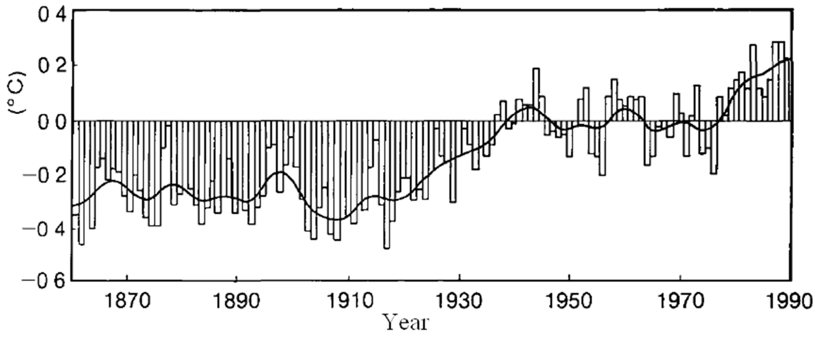

Figure 1. *Global-mean combined land-air and sea-surface temperatures, 1861–1989, relative to the average for 1951–80.*

One can see from the fifth assessment report [2], that the globally averaged combined land and ocean surface temperature data as calculated by a linear trend, show a warming of 0.85[0.65 to 1.06]°C, over the period 1880 to 2012 (figure 2).

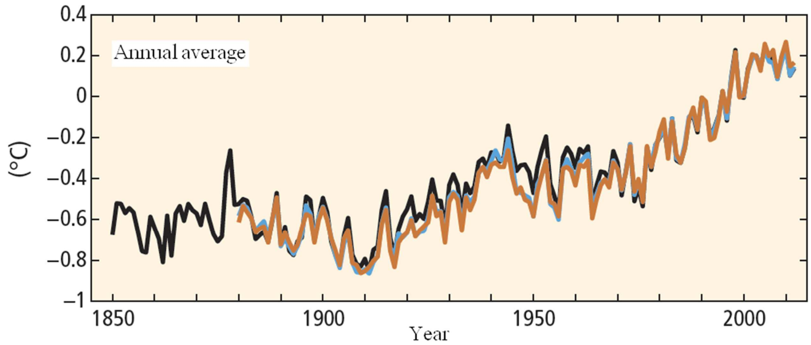

Figure 2. *Observed globally averaged combined land and ocean surface temperature anomaly 1850–2012.*

However, the scientific methods used for these conclusions are not mentioned. In this regard, there is a large number of "climate sceptics" who question both the fact of global warming and the role of the human in this process [3-5].

In this work the temperature changing at the local meteorological station over the last 20 years is investigated. As a scientific method the statistical theory of regression analysis is used.

2. The Content of the Research

The initial data for this research are the meteorological observations for separate months at the meteorological station 26850 Minsk (Belarus). The observations contain the quantitative features of the weather (atmospheric temperature, atmospheric pressure, direction and velocity of the wind, relative humidity of the air and others). They are measured at the meteorological stations with a periodicity of three hours beginning at 00:00 UTC/GMT (Greenwich Time).

At first the monthly mean values and yearly mean values of the temperature are calculated and then mean values are approximated by linear time-dependent regression functions.

The monthly mean value of the temperature for the fixed month of the fixed year was calculated by the formula

$$\overline{t}_{y,m} = \frac{1}{n_{y,m}} \sum_{j=1}^{n_{y,m}} t_{j,y,m} , \quad y = 1998...2017 , \quad m = \overline{1,12} , \qquad (1)$$

where $t_{j,y,m}$ is j-th measure of the temperature t in m-th month of the y-th year, $n_{y,m}$ is the number of the measures of the temperature t in m-th month of the y-th year, $\overline{t}_{y,m}$ is the mean value of the temperature t in m-th month of the y-th year. The received mean values of the temperature $\overline{t}_{y,m}$ for the January 1998–2017 are shown in the table 1. One can see, that the coldest January was in 2010 year of this period with the mean temperature minus 11.08°C. Note that the author has no data for 2002, 2004 and 2005 years.

Table 1. *The monthly mean values of the temperature* $\overline{t}_{y,m}$ *for the January 1998–2017 at the meteorological station Minsk.*

Year	1998	1999	2000	2001	2003	2006	2007	2008
Mean temp.°C	-1.351	-3.162	-3.937	-2.529	-5.632	-8.384	0.2133	-2.625

Year	2009	2010	2011	2012	2013	2014	2015	2016	2017
Mean temp.°C	-4.054	-11.08	-3.813	-5.023	-7.344	-7.443	-1.282	-7.381	-5.683

The monthly mean values of the temperature $\overline{t}_{y,m}$ received by the formula (1) were used to calculate the yearly mean values of the temperature $\overline{t}_y$ by the following formula:

$$\overline{t}_y = \left(\sum_{m=1}^{12} n_{y,m} \right)^{-1} \left(\sum_{m=1}^{12} n_{y,m} \overline{t}_{y,m} \right), \ y = 1998...2017 \quad (2)$$

The yearly mean values of the temperature $\overline{t}_y$ for 1998–2017 years calculated via the formula (2) are shown in the table 2 (there are no data for 2002, 2004 and 2005 years). One can see that 1998 was the coldest year in this period with the mean temperature 6.304°C.

Table 2. *The yearly mean values of the temperature* $\overline{t}_y$ *for the 1998–2017 at the meteorological station Minsk.*

Year	1998	1999	2000	2001	2003	2006	2007	2008
Mean temp.°C	6.304	7.836	7.878	6.998	6.518	6.774	7.730	7.886

Year	2009	2010	2011	2012	2013	2014	2015	2016	2017
Mean temp.°C	6.988	6.925	7.577	6.826	7.479	7.792	8.732	7.782	7.650

The yearly mean values of the temperature $\overline{t}_y$ from the table 2 are demonstrated in the figure 3. One can see that they satisfy the conditions of the regression analysis. In this connection the estimations $\widehat{\theta}_1$, $\widehat{\theta}_2$ of the parameters θ_1, θ_2 of the linear regression function

$$z = \theta_1 + \theta_2 x = H^T(x)\overline{\theta}, \quad (3)$$

where $H^T(x) = (1, x)$, $\overline{\theta}^T = (\theta_1, \theta_2)$, $x = y - 1997$, $y = 1998...2017$, was calculated. The variable x takes the values 1…20. The received empirical regression function $\widehat{z} = \widehat{\theta}_1 + \widehat{\theta}_2 x = H^T(x)\widehat{\overline{\theta}}$ is demonstrated in the figure 3 (straight line). There is some increasing of this linear regression function. The question, however, is whether such increasing significant? The linear regression analysis allows answer this question by using two types of statistical tests: on the significance of the parameters and on the significance of the linear dependence of the regression function [6-8]. This analysis is performed below.

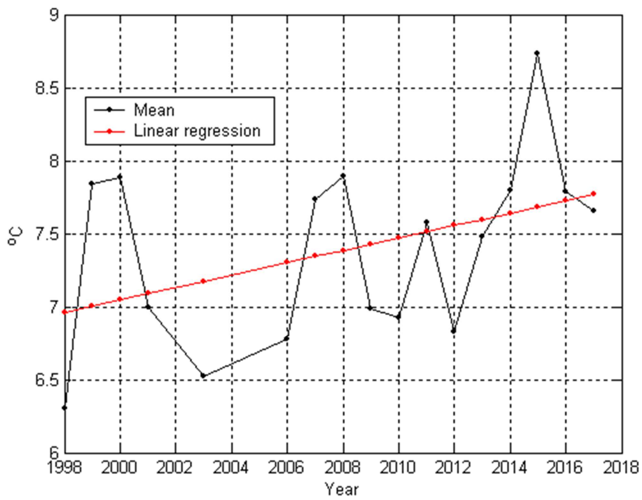

Figure 3. *The yearly mean temperature* $\overline{t}_y$ *for 1998–2017 years at the meteorological station Minsk.*

The vector estimate $\widehat{\overline{\theta}}^T = (\widehat{\theta}_1, \widehat{\theta}_2)$ of the vector parameter $\overline{\theta}^T = (\theta_1, \theta_2)$ in (3) is determined by the formula [3, 4].

$$\widehat{\overline{\theta}} = (F^T F)^{-1}(F^T Z), \quad (4)$$

where $F = (H(x_1), H(x_2), ..., H(x_n)) = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ x_1 & x_2 & \cdots & x_n \end{pmatrix}$, $Z^T = (\overline{t}_{y,1}, \overline{t}_{y,2}, ..., \overline{t}_{y,n})$.

The t-statistic

$$t_i = \frac{\widehat{\theta}_i}{\sqrt{\widehat{\sigma}_1^2 a^{i,i}}}, \ i = \overline{1,2}, \quad (5)$$

is used in the test on the significance of the parameter $\widehat{\theta}_i$ of the linear regression function (3). There $\widehat{\sigma}_1^2 = \frac{1}{n-k}(Z - F\widehat{\overline{\theta}})^T(Z - F\widehat{\overline{\theta}})$, $a^{i,i}$ is element of the matrix $A^{-1} = (a^{i,i}) = (F^T F)^{-1}$, n is the size of the sample, k is the number of the unknown parameters in the regression function (1). In our case $n = 17$, $k = 2$. The statistic t_i (5) has the Student distribution with $n-k$ degrees of freedom.

The f-statistic

$$f = \frac{R_1^2 - R_0^2}{k-1} \Big/ \frac{R_0^2}{n-k} \quad (6)$$

is used in the test on the significance of the linear dependence of the regression function. There $R_1^2 = \sum_{i=1}^{n} (\overline{t}_{y,i} - \overline{z})^2$, $\overline{z} = \frac{1}{n}\sum_{i=1}^{n} \overline{t}_{y,i}$, $R_0^2 = \sum_{i=1}^{n} (\overline{t}_{y,i} - H^T(x_i)\widehat{\overline{\theta}})^2$.

The f-statistic (6) has the Fisher F-distribution with $k-1$, $n-k$ degrees of freedom. It is the known analysis of variance method (ANOVA).

As a result, the following estimations of the parameters of the linear regression function for the yearly mean temperature at the meteorological station Minsk via the formula (4) was calculated: $\hat{\theta}_1 = 6.9217$, $\hat{\theta}_2 = 0.0421$.

The hypothesis on the significance of the parameters of the linear regression function $\{H_0 : \theta_i = 0;\ H_1 : \theta_i \neq 0\}$, $i = 1,2$, was tested by the two-side test of significance $P(|t_i| > t_{\alpha/2}) = \alpha$. The significance limit $t_{\alpha/2}$ on the typical significance level $\alpha = 0,05$ at $n-k=15$ degrees of freedom has value $t_{\alpha/2} = 2.1314$. The empirical values of the statistics t_i calculated via the formula (5) are equal: $t_{1\text{emp.}} = 23.0656$, $t_{2\text{emp.}} = 1.7795$. As $t_{1\text{emp.}} > t_{\alpha/2}$, the hypothesis $H_0 : \theta_1 = 0$ is rejected, i.e. the parameter θ_1 is significant. As $t_{2\text{ emp.}} < t_{\alpha/2}$, the hypothesis $H_0 : \theta_2 = 0$ is received, i.e. the parameter θ_2 is insignificant.

The hypothesis on the significance of the linear dependence of the regression function in our case has a view $\{H_0 : \theta_2 = 0;\ H_1 : \theta_2 \neq 0\}$. It was tested by the right-side test of significance $P(f > f_\alpha) = \alpha$. The significance limit f_α on the typical significance level $\alpha = 0,05$ at $k-1=1$, $n-k=15$ degrees of freedom has value $f_\alpha = 4.5431$. The empirical value of the statistic f calculated via the formula (6) is equal $f_{\text{emp.}} = 3.1667$. As $f_{\text{эмп.}} < f_\alpha$, the hypothesis $H_0 : \theta_2 = 0$ is received.

As a result, two test show that there is no significant linear stochastic dependence the yearly mean temperature on the time for 1998–2017.

The same analysis is performed also for the data from the table 1 about the monthly mean values of the temperature $\bar{t}_{y,m}$ for the January 1998–2017 at the meteorological station Minsk. These data are showed in the figure 4 as an irregular curve. The empirical linear regression function $\hat{z} = \hat{\theta}_1 + \hat{\theta}_2 x = H^T(x)\hat{\theta}$ for these data is showed in the figure 4 as the straight line which decreases. The estimations of the parameters are $\hat{\theta}_1 = -2.7482$, $\hat{\theta}_2 = -0.1779$. The changing of the temperature is determined by the parameter θ_2. The significance of this parameter was checked by the t-statistic (5), i.e. the hypothesis $\{H_0 : \theta_2 = 0;\ H_1 : \theta_2 \neq 0\}$ was checked by the two-side test of significance $P(|t_2| > t_{\alpha/2}) = \alpha$. The empirical values of the statistic t_2 calculated by the formula (5) is equal $t_{2\text{emp.}} = -1.5583$. The significance limit $t_{\alpha/2}$ on the significance level $\alpha = 0,05$ at $n-k=15$ degrees of freedom is $t_{\alpha/2} = 2.1314$. As $|t_{2\text{ emp.}}| < t_{\alpha/2}$, the hypothesis $H_0 : \theta_2 = 0$ is received, i.e. the parameter θ_2 is insignificant. The hypothesis $H_0 : \theta_2 = 0$ on the significance of the linear dependence of the regression function was checked also by the ANOVA, i.e. by the f-statistic (6) and the right-side test of significance $P(f > f_\alpha) = \alpha$. It was calculated: $f_{\text{emp.}} = 2.4282$, $f_\alpha = 4.5431$, $\alpha = 0,05$, $k-1=1$, $n-k=15$, $f_{\text{эмп.}} < f_\alpha$, so the hypothesis $H_0 : \theta_2 = 0$ is received. As a result, the both tests show that there is no significant linear stochastic dependence the mean temperature on the time for the January 1998–2017 at the meteorological station Minsk.

It should be noted that in the works [9, 10] other statistical approaches were used.

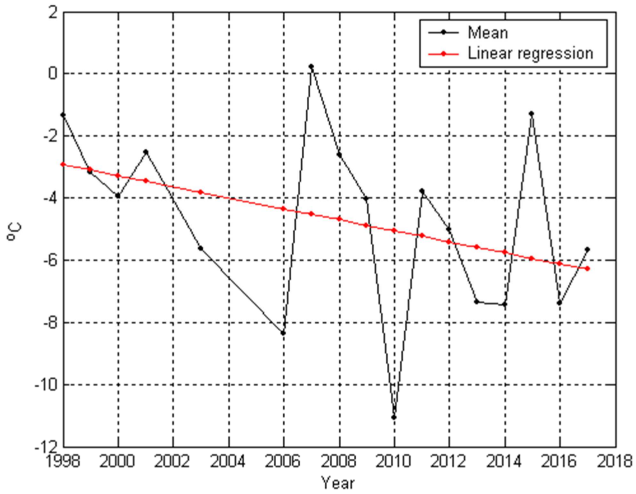

Figure 4. *The monthly mean values of the temperature* $\bar{t}_{y,m}$ *for the January 1998–2017 at the meteorological station Minsk.*

3. Conclusion

In this work the analysis of the yearly mean value of the atmospheric temperature on the meteorological station Minsk over the last 20 years was performed. As an investigation method the statistical theory of regression analysis was chosen. The linear stochastic dependence the yearly mean temperature on the time (empirical regression function) over the last 20 years has a slight tendency to increase. However, using the both statistical tests on the significant of the parameters and on the linearity of the regression function shows that the parameter θ_2 of the regression function (3) which defines the trend of the mean temperature is insignificant. The similar analysis for the average monthly mean temperature for all months of the year was performed (the analysis for January was outlined above on the base of the data in the table 1). The investigations show that the linear approximation of the monthly mean temperature slightly decreases in the January, February, April and October and in other months of the year slightly increases. However, the parameter θ_2 of the regression function (3) was confirmed as insignificant by both tests for all the months. As a result, it can be stated, that there have been no significant linear changing in the yearly and monthly mean temperature at the meteorological station Minsk from 1998 to 2017.

Thus, the importance of using of the instrumental scientific methods for the conclusions about climate change was emphasized.

References

[1] IPCC First Assessment Report 1990 (FAR). Available at http://www.ipcc.ch.

[2] IPCC Fifth Assessment Report. Available at http://www.ipcc.ch.

[3] http://www.climate-skeptic.com.

[4] https://en.wikipedia.org/wiki/Climate change denial.

[5] Björnberg, Karin Edvardsson et al. Climate and environmental science denial: A review of the scientific literature published in 1990-2015. Journal of Cleaner Production. 2017. Volume 167. Pp. 229-241.

[6] I. N. Vuchkov, L. Boyadjeva, E. Solakov. Applied linear regression analysis. Moscow, Finance and Statistics, 1987. 238 p. In Russian.

[7] V. S. Mukha. Statistical methods of data processing: tutorial. Minsk, BSU, 2009. 183 p. In Russian.

[8] C. S. Rao, Linear statistical inference and its applications. Wiley, 2ed, 1973. 648 p.

[9] Milivoj b. Gavrilov, Slobodan B. Marković, Natalija Janc, Milena Nikolić et al. Assessing average annual air temperature trends using the Mann–Kendall test in Kosovo. Acta geographica Slovenica. Volume 58. Issue 1. Pp. 7-25.

[10] McKitrick, Ross R., Stephen McIntyre and Chad Herman. Panel and multivariate methods for tests of trend equivalence in climate data sets. Atmospheric Science Letters. 2010. Volume 11. Issue 4. Pp. 241-331.