A Trend of the Five-Year Development Temperatures computed from Average Monthly Temperatures in the Territory of the Czech Republic

RENATA OSICKOVA, STANISLAV BARTON, VOJTECH KUMBAR
Department of Technology and Automobile Transport
Mendel University in Brno
Zemedelska 1, 613 00 Brno
CZECH REPUBLIC

xosicko7@node.mendelu.cz

Abstract: In this paper, its authors process and analyze values of average monthly temperatures recorded in 34 meteorological stations since January 2003 till December 2013 that are uniformly distributed in the territory of the Czech Republic. At first statistical relevance of each term of the used regression function is evaluated using the sum of squared residuals per degree of freedom. A second function containing only significant terms is tested on the edge of statistical reliability equal to 99%, 95% and 90% with respect to its individual terms. The Fisher-Snedecor function was used to determine really important terms of the evaluated function. The Final function enabled computation coefficients of 7 regression functions explaining recorded data in 7 time steps of 5 year intervals. These functions were used to determine border positions splitting areas of the Czech Republic into zones of local warming and cooling. Recorded data is plotted graphically and shows that this border oscillates over the whole territory of the Czech Republic.

Key-Words: global warming, mathematical modelling, regression function, linear correlation, space and time coincidence, temperature trends

Introduction

Problems of global warming represent a widely discussed theme that is the focus of interest of the major part of world population, for example [2, 9]. Many authors publish papers that accentuate the fact that the process of global warming is real and quite inevitable while some others write that this is a disputable phenomenon and that the global warming is a mere fiction, [3, 5].

In this paper we present a mathematical study of the development of diurnal temperatures in the territory of the Czech Republic within the period of the recent decade from January 2003 till December 2013. Using a Maple application, see [6], based on the method of least squares, we have developed a regression function $T(t,x,y,h)$, which explains the dependence of temperature on time, geographical position and height above the sea. To determine the significance of the regression function members that have been tested with a confidence interval of 90%, 95% and 99% have been used in the Fischer-Snedecor function (4). The resulting functions allow us to calculate the coefficients of the regression function, which is used at the end and are necessary for determining the motion boundary warming and cooling in the area of the Czech Republic. Calculations were made for seven five-year cycles, and gradually shifted by one year.

Material and Methods

Data concerning average monthly temperatures as recorded within the period of last ten years in 22 selected meteorological stations are normally available on the Internet, [10]. As far as further 12 stations are concerned, similarly data can be obtained from graphs that are available at the web page, [11].

The Czech Hydrometeorological Institute collects data about daily temperatures, as measured and recorded in a much higher number of meteorological stations already for a long time period. These data, however, can be obtained only on the base of payments and for that reason they are not available for wider public.

Nevertheless, data recorded in available 34 meteorological stations cover the territory of the Czech Republic adequately and in a satisfactory manner. The minimum airline distance between two stations is 12 km while the maximum does not exceed 54.7 km. Distances were recomputed from GPS coordinates onto *XY* coordinates with respect to [8, 12]. Data presented in this paper informs about an exact geographical location of the station, about its altitude and also about average monthly air temperatures. Temporary data is expressed as yearly fractions and the time t = 0 corresponds with the 1st January 2003. In case that partial data about the temperature is excluded or missing, the temperature is rewritten by -99 °C. Stations with incomplete data are highlighted in red, stations in Group 1, or in blue, Group 2. Data from Group 2 is reconstructed from graphs.

$$F(t,x,y,h) = c_1 x\cos(kt) + c_2 x\sin(kt) + c_3 x + c_4 xt + c_5 xt^2 + c_6 y\cos(kt) + c_7 y\sin(kt) + c_8 y + c_9 yt + c_{10} yt^2 + c_{11} h\cos(kt) + c_{12} h\sin(kt) + c_{13} h + c_{14} ht + c_{16} ht^2 + c_{17}\cos(kt) + c_{18}\sin(kt) + c_{19} t + c_{20} t^2, \quad (1)$$

The first step is to determine the statistical significance of each member of the regression function (1), where $k = 2\pi$, using the sum of squared residues on one degree of freedom – SQR_1 with a precision of 15 significant digits, (2).

$$SQR_1 = \frac{\sum_{i=1}^{N}\left(F(t_i,x_i,y_i,h_i) - T_i\right)^2}{N-p}, \quad (2)$$

where N = number of measurements, t = time, $[x,y,h]$ = spatial coordinates, p = number of function parameters F, F = seeking regression function, T_i = measured temperature. It turned out, not all members of the function (1) leads to a reduction in the sum squared residuals at one degree of freedom. Therefore, the resulting function (3) contains only members that meet this condition. In addition, each member in the list was sorted according to their importance.

$$F_3(t,x,y,h) = 10.9319 - 10.5003\cos(kt) - 2.4811\sin(kt) - 0.0059h + 0.0012h\cos(kt) - 0.0004yt - 0.0022x - 0.0010x\cos(kt) - 0.0013x\sin(kt) - 0.0004h\sin(kt) - 0.0000ht^2 + 0.0015y\cos(kt) + 0.0001ht - 0.0022y + 0.0000xt^2, \quad (3)$$

The accuracy of calculation was then verified with a precision of 36 digits in force according to another algorithm, see [4]. All computations were done in the programme Maple, [6]. Both calculations differed only within the selected numerical precision, which means that their relative difference was of the order of 10^{13}%. The coefficient of linear correlation for each station for the entire period 2003-2013 ranges from 0.961 to 0.975, average correlation coefficient of linear spatial temperature distribution in the Czech Republic is 0.931. This means that the regression function can be considered as satisfactory.

Fischer-Snedecor function, see [13, 7], $F(z)$, (4), is used for testing and determining the significance of members of the regression function with more coefficients compared to simpler regression function. Level of the uncertainty is α.

$$\int_0^q F(z)\,dz = 1-\alpha, \text{ where } F(z) = \frac{\left(\frac{\kappa}{n}\right)^{\left(\frac{\kappa}{2}\right)} z^{\left(\frac{\kappa}{2}-1\right)}}{B\left(\frac{\kappa}{2},\frac{n}{2}\right)\left(1+\frac{z\,\kappa}{n}\right)^{\left(\frac{\kappa}{2}+\frac{n}{2}\right)}}, \quad \begin{array}{l}\kappa = \text{difference of count of parameters of functions}\\ n = \text{count of measurements.}\end{array} \quad (4)$$

This feature, with accuracy $1-\alpha$, tells us how to change the statistical significance of the function, if we add more function members. More complex functions is statistically significant if it satisfies the condition (5). Selection of the most suitable model is performed on the basis of a test which is based on the inequality (5), see tables XVIII 4a-4c in [1].

$$\frac{\frac{S_R(1)-S_R(2)}{p_2-p_1}}{\frac{S_R(2)}{n-p_2}} \geq q \quad \text{where} \quad \begin{array}{c} S_R(1) = \text{is the residual sum of squares of a simple model}\\ S_R(2) = \text{is the residual sum of squares of a complex model}\\ p_1 = \text{number of coefficients of a simple model}\\ p_2 = \text{number of coefficients of a complex model}\end{array}, \quad \kappa = p_2 - p_1. \quad (5)$$

Results and Discussion

Determining individual members of regression function

For each member of the regression function (1) containing 20 members is calculated the sum of squared residuals per 1 degree of freedom, further shorten as **SQR1,** see [14]. Selected is the value where **SQR1** is the lowest. This selected member is searched from the remaining 19 members, another member such that **SQR1** for two members was the lowest. It is being continued for each subsequent member by this manner until they are exhausted all the members or **SQR1** begins to rise. In this way, we find that members *2xt, y,* sin*(kt), yt*^{2}*, t, t*^{2} are meaningless. Graph showing **SQR1** as a function of the number of members is shown below, (see Fig. 1). From this graph it is clear that with the increasing number of sum of squared residuals on 1 degree of freedom to the 15th member decreases. Graph showing **SQR1** for two consecutive functions that differ from each other about member is for 5th to 15th member shown below, (see Fig. 2).

Fig. 1 **SQR1** as a function of count of operands of F3

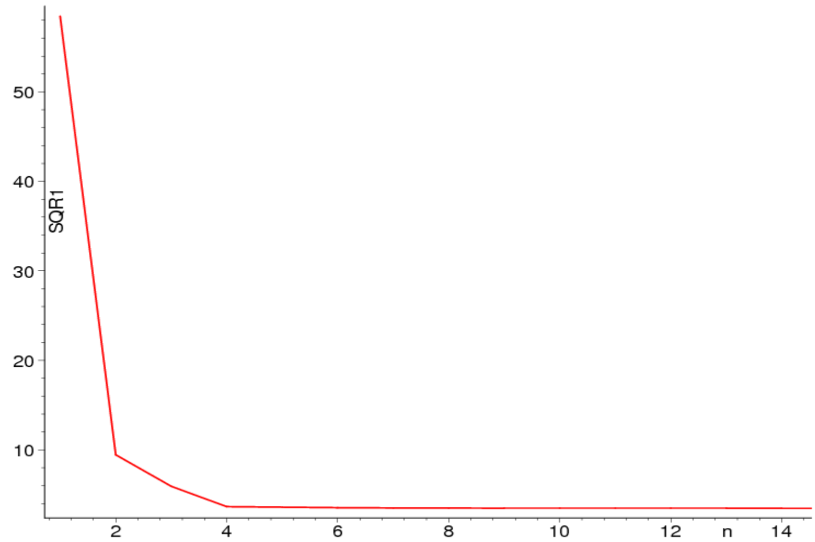
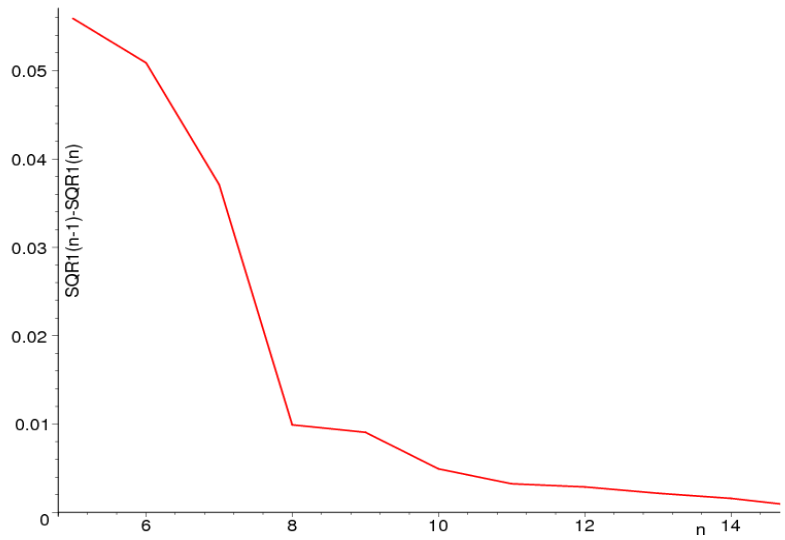
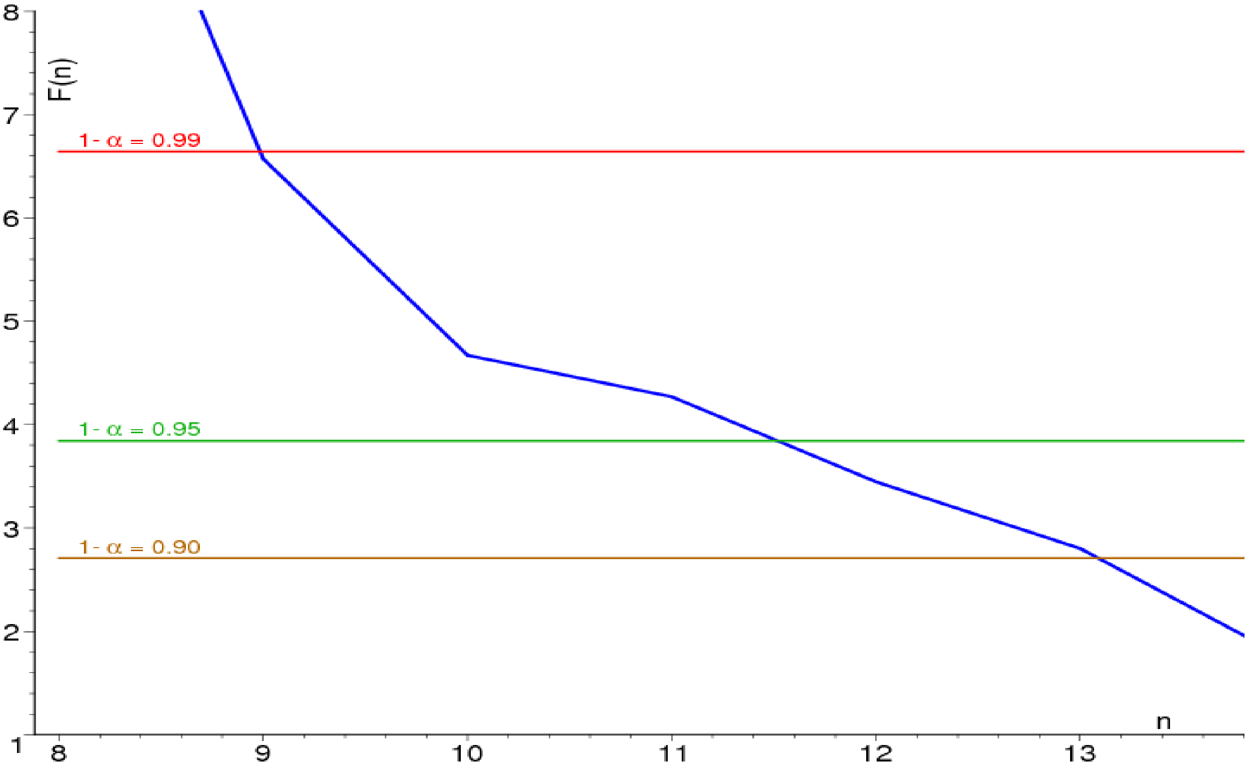


Fig. 2 **SQR1(n) - SQR1(n-1)** as a function of count of operands of F3



Test to determine whether a more complex model (multivariable) better than the simpler model is done with respect to the equation (5). The more complex function is omitted if $F > F_{1-\alpha}(p2-p1,\ n-p2)$.The count of operands of functions that correspond to the reliability of α = 90%, α = 95% and α = 99%, (see Fig. 3).

Fig. 3 Value of the q, (5), as a function of the count of operands of the function F3



Periodic components sin*(kt)* and cos*(kt)* are removed of these functions appropriate to the individual reliability. Then a derivation function by time is done, see (6).

$$f_{99}=10.93199-0.00586\,h-0.00038\,y\,t-0.00222\,x$$
$$f_{95}=10.93199-0.00586\,h-0.00038\,y\,t-0.00222\,x-0.00001\,h\,t^2 \qquad (6)$$
$$f_{90}=10.93199-0.00586\,h-0.00038\,y\,t-0.00222\,x-0.00001\,h\,t^2+0.00014\,h\,t\,.$$

These derivatives are set equal to zero, see (7). Furthermore, the coordinates *y* are calculated and the exact position of 3 boundary for a 5-year intervals are found.

$$ft_{99}=-0.00038\,y$$
$$ft_{95}=-0.00038\,y-0.00003\,h\,t \qquad (7)$$
$$ft_{90}=-0.00038\,y-0.00003\,h\,t+0.00014\,h\,.$$

The functions corresponding reliability α = 99% passes through the center of gravity of the Czech Republic. This fact shows the graph in Fig 4 – positions on the map of the Czech Republic and Fig. 5 showing average *y* position of the boundaries in 7 five years intervals.

From the calculations above it is clear that members of the 9 explain 99% of the data, the members of the 13 explain 95% of the data and the members of the 14 explain 90% of the data. The boundary, 95% and 99% reliability are plotted on the graph, (see Fig. 4).

In our case, we work with 95% reliability, for which the corresponding function contains 14 members. This reliability is designed for more complex models. The boundaries, corresponding to 99%, is stable and passes through the center of gravity of the Czech Republic. The functions corresponding 99% reliability, contains 9 members and it is independent on time.

Fig. 4 Average cooling/warming border position of 5 year intervals in the Czech Republic for 1-α = 95%

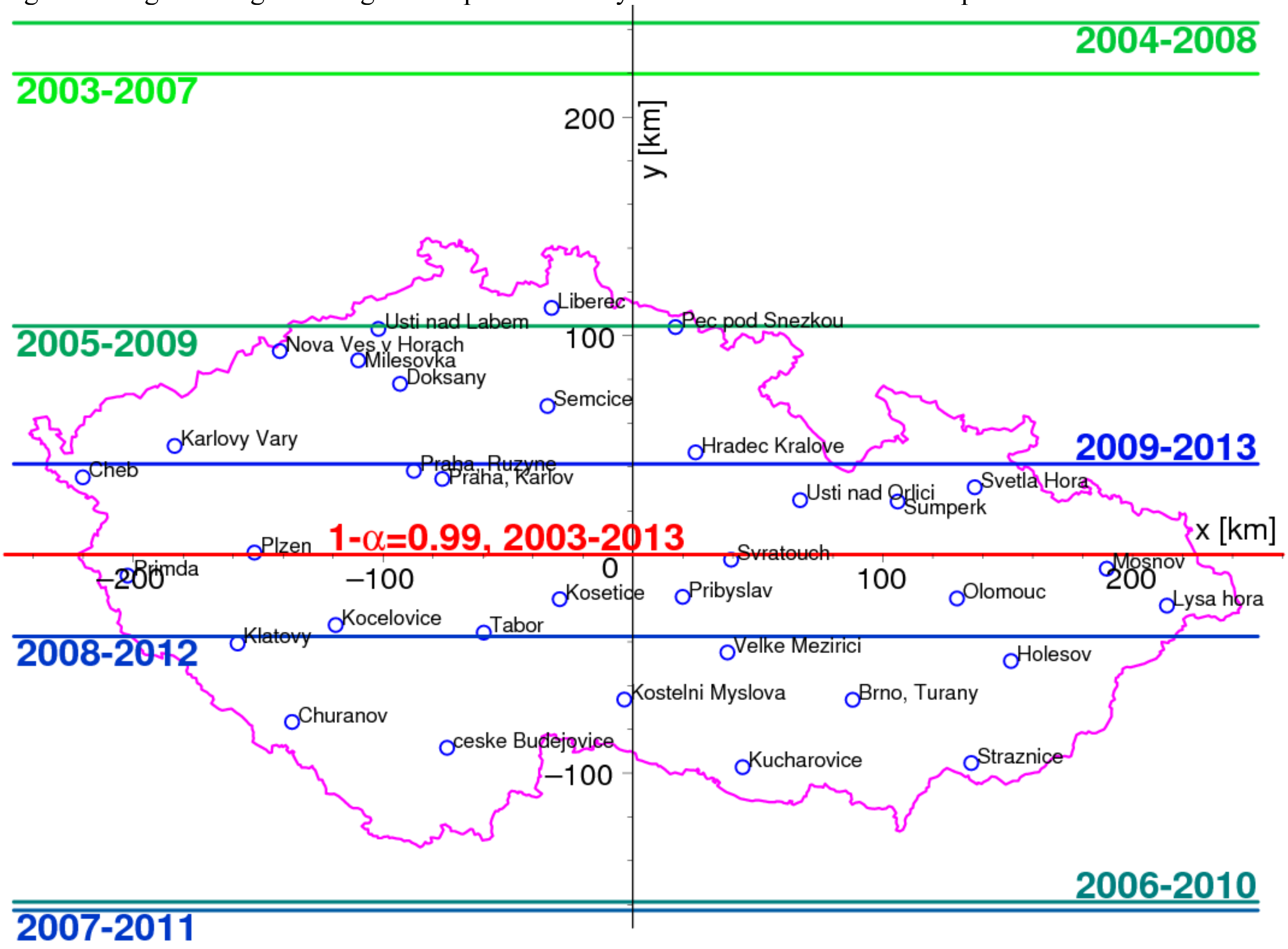
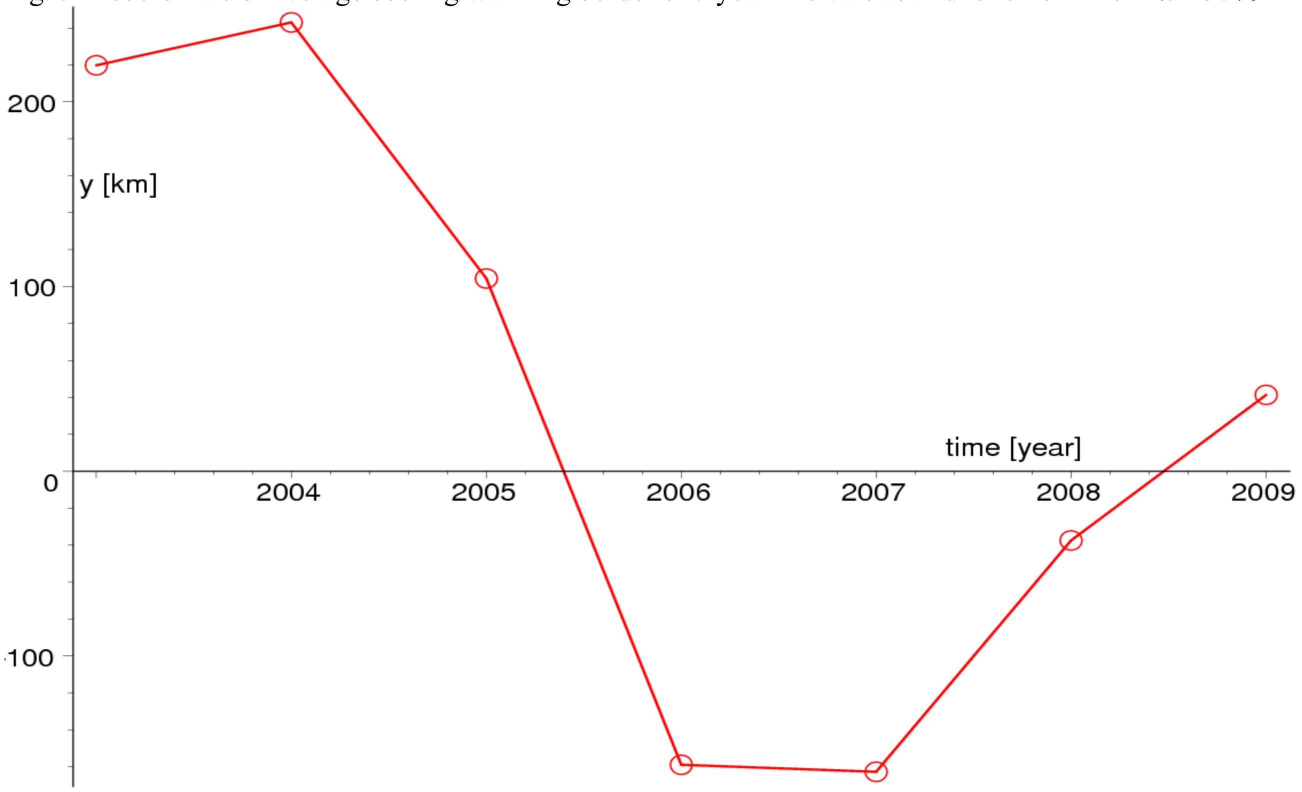


Fig. 5 Y coordinate of average cooling/warming border of 5 year intervals as a function of time 1-α = 95%



Conclusion

In area of the Czech Republic there is a border between areas where warming and cooling are occurring. If we use a simpler function, then a position of this boundary has coordinates y = 0 (the border passes through the center of gravity of the Czech Republic) and is not dependent on time. This feature is able to explain 99% of the measured data.

From the properties of the more complex function, it follows that the position of the boundary is moved over the whole area of the Czech Republic, (see Fig. 4 and Fig. 5).

Acknowledgement

The authors want to express their thanks to Mrs. Pýchová from the Czech Hydrometeorological Institute for information about publicly accessible meteorological data.

The research has been supported by the project TP 4/2014 "Analysis of degradation processes of modern materials used in agricultural technology" financed by IGA AF MENDELU.

References:

[1] Anděl J, Matematická statistika. SNTL, Praha, 1985, pgs. 327-329
[2] Barros V, Globální změna klimatu. Praha, 2006, pp. 165.
[3] Dymnikov VP, Filatov AN, *Mathematics of Climate Modeling*. Birkhäuser, Boston-Basilej-Berlin, 1997, pp. 264.
[4] Gander W, Hřebíček J, *Solving problems in Scientific Computing Using Maple and MATLAB*. Springer-Verlag Berlin Heidelberg 1993, 1995. Germany, pp. 315.
[5] Klaus V, *Modrá, nikoli zelená planeta.*, Dokořán. Praha, 2009, pp. 164.
[6] MAPLE 11, *User Manual*, 1st edition, Maplesoft, 2007, pp. 396.
[7] Meloun M, Militký J, Kompendium statistického zpracování dat. Academia, Praha, 2006, pg. 677, ISBN 80-200-1396-2
[8] Meyer TH, *Introduction to Geometrical and Physical Geodesy*, 3rd edition, ESRI Press, 2010, pp. 246.
[9] Potter TD, Colman BR, *Handbook of weather, climate, and water: dynamics, climate, physical meteorology, weather systems, and measurements*. Hoboken, NJ, USA, 2005, pp. 973.
[10] WEB1, http://www.czso.cz/csu/2012edicniplan.nsf/kapitola/0001-12-r_2012-0200
[11] WEB2, http://portal.chmi.cz/portal/dt?action=content&provider=JSPTabContainer&menu=JSPTabContainer/P4_Historicka_data/P4_1_Pocasi/P4_1_9_Mesicni_data&nc=1&portal_lang=cs#PP_Mesicni_data
[12] WEB3, http://www.fd.cvut.cz/department/k611/PEDAGOG/files/webskriptum/kartografie/kartografie.html
[13] WEB4, https://ucnk.ff.cuni.cz/bonito/nahoda.php
[14] WEB5, http://dsp.vscht.cz/konference_matlab/matlab03/balda.pdf