WEIGHTED LEAST SQUARE METHODS AND CALIBRATION OF TEMPERATURE SENSORS

L. Kňazovická[1], *R. Strnad*[2]

[1]Department of Computing and Control Engineering, Institute of Chemical Technology Prague, Technická 5, 166 28 Prague 6, Czech Republic

[2]Czech Metrology Institute, Okružní 31, 638 00 Brno, Czech Republic

Abstract

In the calibration laboratories are used standard procedures for calculating of the calibration model coefficients based on well described standards (ČSN EN 60751, ITS-90, ČSN EN 60584, etc.). In practice, sensors are mostly calibrated in more points and redundant information is used as a validation of the model. This paper will present a comparison of the methods which are nowadays used for calculations of the calibration model coefficients for different type of the thermometers/standards. A weighted least square method will be used for the evaluation of the whole measured dataset and the investigation of the differences will be discussed.

1. Introduction

Thermocouples (TC) are used for measure of temperature in a lot of applications. A thermocouple is a device consisting of two different conductors (usually metal alloys) that produce a voltage proportional to a temperature difference between either ends of the pair of conductors [1]. Temperature range, in which can be TC used depends on composition of the material used for the conductors. Most common types of TC with their composition and temperature and voltage ranges are described in Table 1.

Table 1: TEMPERATURE RANGES AND COMPOSITION OF THE TC

Type of TC	Composition	Voltage [μV]	Temperature range [°C]
R	Pt - 13 % Rh/ Pt	-226 to 21103	-50 to 1768,1
N	Ni-Ct-Si/Ni-Si	-3990 to 47513	-270 to 1300
S	Pt – 10 % Rh/ Pt	-235 to 18694	-50 to 1768,1
B	Pt – 30 % Rh/ Pt – 6% Rh	0 to 13820	0 to 1820

Fact, that by using the TC is measured a voltage means that temperature is measured indirectly – and must be calculated from measured voltage by reference (1) and inverse (2) functions. These functions are defined in standard ČSN EN 60584-1 [2]. Form of these equations are equals for each type of TC, the only difference is in a constants a and d.

$$E_{ref} = \sum_{i=1}^{n} a_i (t_{90})^i \quad (\mu\text{V}) \qquad (1)$$

$$t = \sum_{i=1}^{n} d_i E^i \quad (°\text{C}) \qquad (2)$$

where t_{90} (°C) is temperature defined according to temperature scale ITS-90, E_{ref} is a reference voltage (μV) at temperature t_{90}, E is measured voltage (μV) and t represent measured temperature (°C) calculated from E.

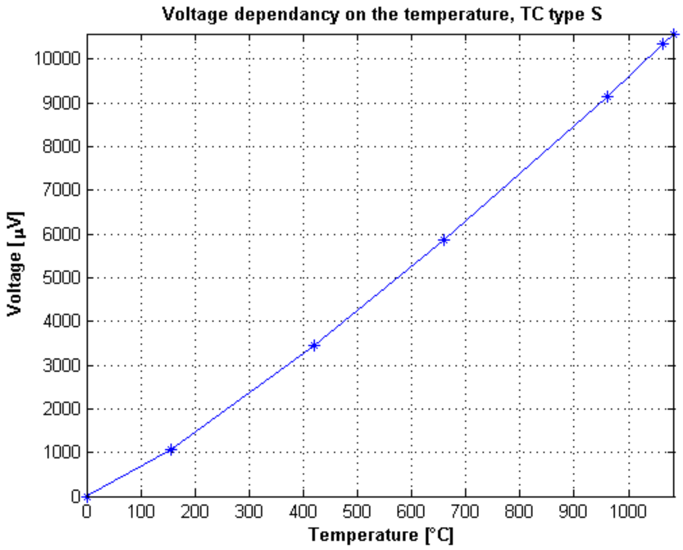

Figure 1: Voltage dependency on the temperature, TC type S

From described equations is calculated reference voltage (1) and real measured temperature (2). These values allow determining differences between measured and calculated values, which are necessary for further progress. Difference between measured and reference temperature is important for calculation of coefficients of deviation function (a, b, c). This function is represented by 3rd order polynomial function (3). Shapes of deviation function for most common TC`s are shown in the Fig. 2.

$$diff_i = aE_i + bE_i^2 + cE_i^3. \tag{3}$$

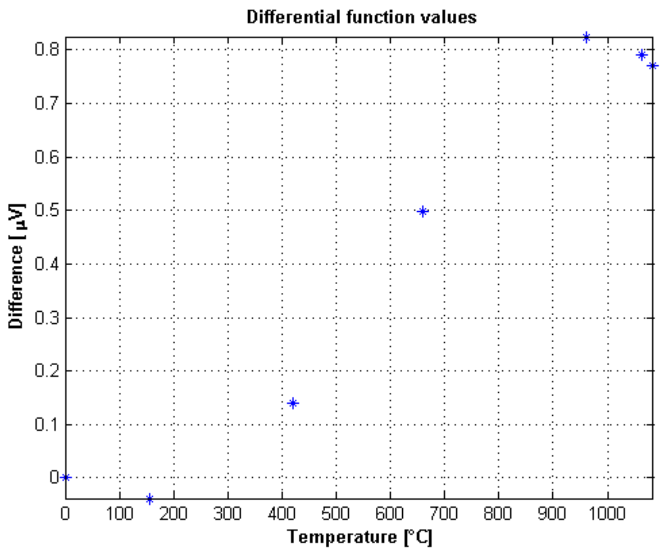

Figure 2: Differential function values for TC S, N, R and B

Values from this function allow evaluation of behavior of the thermocouple. It is a useful tool during the calibration of the sensor, because shapes of these functions are well-known, and some difference towards normal behave can indicate some problem during the calibration or some problem with TC itself.

Sensors are calibrated only in a few points of the measured range of the thermocouple (fixed points with exactly defined temperature – e.g. Triple Point of Water, Indium, Tin, Zinc, Silver, Copper), but they are used for measuring of the temperature in a whole range. Values of deviation function aren`t linear with the temperature, but it`s always a polynomial function. Determination of the accurate order of this polynomial is really important for correct use of the thermometer.

Important part of calibration is determination of the uncertainty. Uncertainty is a non-negative parameter, which characterize the dispersion of the values attributed to the measured quantity. Uncertainty is calculated for each of measured point separately, and consists of a lot of parts, e.g. characterization of the calibrated device and calibration procedures. For purposes of this paper values of the uncertainty will be used as a weight in Weighted Least Squares Method.

2. Weighted least square method

For finding of accurate orders of deviation function for various type of TC is used *Weighted Least Square Method* (WLS). This method is described in a lot of sources [3,4]. Usage of this method isn`t very complicated, but it provides a strong tool for function approximates. This method represents a modification of the *Method of Least Squares*, so this method will be described firstly.

The Method of Least Squares is a procedure to determine the best fit line to data, the proof uses simple calculus and linear algebra. For calculations are used data sets, which obtain independent variable *x* (in our case reference temperature t_{90}) and dependent variable *y* (values of the deviation function). Fitting curve *f(x)* has the deviation *e* from each data point

$$e_i = y_i - f(x_i). \tag{4}$$

According to the method of least squares, the best fisting curve minimize this deviation.

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} [y_i - f(x_i)]^2 = \min \tag{5}$$

One of the most common used curves in regression is polynomials with different orders. The Least-Squares line represent 1st order polynomial, when data are fitted by a straight line with equation

$$y = a_0 + a_1 x, \tag{6}$$

where $n \geq 2$ (*n* – number of measured points).

Polynomial of the 2nd order is represented by parabola with equation

$$y = a_0 + a_1 x + a_2 x^2 \tag{7}$$

where $n \geq 3$,

and polynomials of higher order are defined as follows:

$$y = a_0 + a_1 x + a_2 x^2 + \cdots + a_m x^m, \tag{8}$$

where $n \geq m + 1$.

In these equations coefficients $a_0, \ldots, a_m$ are unknown, and they have to be determined. To obtain the least square error, unknown coefficients must yield zero first derivatives.

$$\begin{aligned}
\frac{\partial \mathrm{E}}{\partial a_0} &= 2\sum_{i=1}^{n}\left[y_i - \left(a_0 + a_1 x_i + a_2 x_i^2 + \cdots + a_m x_i^m\right)\right] = 0 \\
\frac{\partial \mathrm{E}}{\partial a_1} &= 2\sum_{i=1}^{n} x_i\left[y_i - \left(a_0 + a_1 x_i + a_2 x_i^2 + \cdots + a_m x_i^m\right)\right] = 0 \\
\frac{\partial \mathrm{E}}{\partial a_2} &= 2\sum_{i=1}^{n} x_i^2\left[y_i - \left(a_0 + a_1 x_i + a_2 x_i^2 + \cdots + a_m x_i^m\right)\right] = 0, \\
&\vdots \\
\frac{\partial \mathrm{E}}{\partial a_m} &= 2\sum_{i=1}^{n} x_i^m\left[y_i - \left(a_0 + a_1 x_i + a_2 x_i^2 + \cdots + a_m x_i^m\right)\right] = 0
\end{aligned} \tag{9}$$

These equations can be rewrite (after dividing by 2 and setting $\partial E/\partial a_0 = \partial E/\partial a_1, \ldots = 0$) as

$$\begin{aligned}
\sum_{i=1}^{n} y_i &= a_0 \sum_{i=1}^{n} 1 + a_1 \sum_{i=1}^{n} x_i + a_2 \sum_{i=1}^{n} x_i^2 + \cdots + a_m \sum_{i=1}^{n} x_i^m \\
\sum_{i=1}^{n} x_i y_i &= a_0 \sum_{i=1}^{n} x_i + a_1 \sum_{i=1}^{n} x_i^2 + a_2 \sum_{i=1}^{n} x_i^3 + \cdots + a_m \sum_{i=1}^{n} x_i^{m+1} \\
\sum_{i=1}^{n} x_i^2 y_i &= a_0 \sum_{i=1}^{n} x_i^2 + a_1 \sum_{i=1}^{n} x_i^3 + a_2 \sum_{i=1}^{n} x_i^4 + \cdots + a_m \sum_{i=1}^{n} x_i^{m+2} \qquad (10) \\
&\vdots \\
\sum_{i=1}^{n} x_i^m y_i &= a_0 \sum_{i=1}^{n} x_i^m + a_1 \sum_{i=1}^{n} x_i^{m+1} + a_2 \sum_{i=1}^{n} x_i^{m+2} + \cdots + a_m \sum_{i=1}^{n} x_i^{2m}
\end{aligned}$$

The unknown coefficients $a_0, \ldots, a_m$ are obtained by solving the above linear equations (10).

Weighted least squares method represents a modification of Method of Least Squares. Into the calculation enters another term – vector or matrix of weights (W_i). These values set to each pair of the data (x_i, y_i) some weight. Weighted least square method affects the points, which are used for the calculation of new regression function. The higher the value of weight, the greater is the influence of this point in the regression. The equation (5) than can be written as follows

$$\sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} W_i [y_i - f(x_i)]^2 = \min \qquad (11)$$

and other equations are modified in a same way.

A lot of possibilities exist for weights determination. Most common way is a calculation of the standard deviation (σ), and uses one of these forms:

$$W_i = \frac{1}{\sigma} \quad \text{or} \quad W_i = \frac{1}{\sigma^2} \qquad (12)$$

In presented case are weights determined otherwise – as weights are considered uncertainties of calibration

$$W_i = \frac{1}{u_i}. \qquad (13)$$

Method of Least Squares is also used for calculation of the RMSE (*Root Mean Square Error*).

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(y - f(x_i)\right)^2} \qquad (14)$$

From decreasing of the RMSE values appropriate order of the polynomial is determined.

3. Results

For thermocouples, which are mostly calibrated in CMI`s laboratories (TC types S, R, N, B) were found orders of the polynomials, which provides best fit of the measured data. Polynomials were calculated with using of both Least Squares Methods (weighted and non-weighted), and for the objective comparison and evaluation of these polynomials RMSE is calculated. Results are shown in Tables 2 and 3.

For the calculation were used all measured points to estimate suitable approximation of higher orders. Maximum possible order of polynomial depends on number of measured points, so in presented cases maximum of the 6th order polynomial are established. Number of the measured points is different for each type of thermocouple because of theirs different temperature measurement ranges. Both methods are compared according to RMSE also in the Fig. 3. From this pictures can be easily seen, which order should be sufficient for data fitting. Best fitting is also visualized on the Fig. 4 for all used types of thermocouples. The data fitting by 4th order of the polynomial is appropriate for the thermocouples types S, R and B is, for the thermocouple type N is sufficient approximation with the polynomial of the 3rd order.

Table 2: RMSE values for different polynomials order (Least square)

	S	R	B	N
1st order	0,09396	0,15720	0,52840	0,49130
2nd order	0,10000	0,18040	0,22160	0,24970
3rd order	0,02318	0,04460	0,01307	**0,00106**
4th order	**0,00173**	**0,00000**	**0,00138**	0,00130
5th order	0,00052	-	0,00000	0,00131
6th order	0,00000	-	-	0,00000

Table 3: RMSE values for different polynomials order (Weighted Least square)

	S	R	B	N
1st order	0,03052	0,07133	0,22880	0,18420
2nd order	0,03411	0,08213	0,08486	0,09278
3rd order	0,00760	0,02038	0,00489	**0,00044**
4th order	**0,00173**	**0,00000**	**0,00055**	0,00053
5th order	0,00020	-	0,00000	0,00049
6th order	0,00000	-	-	0,00000

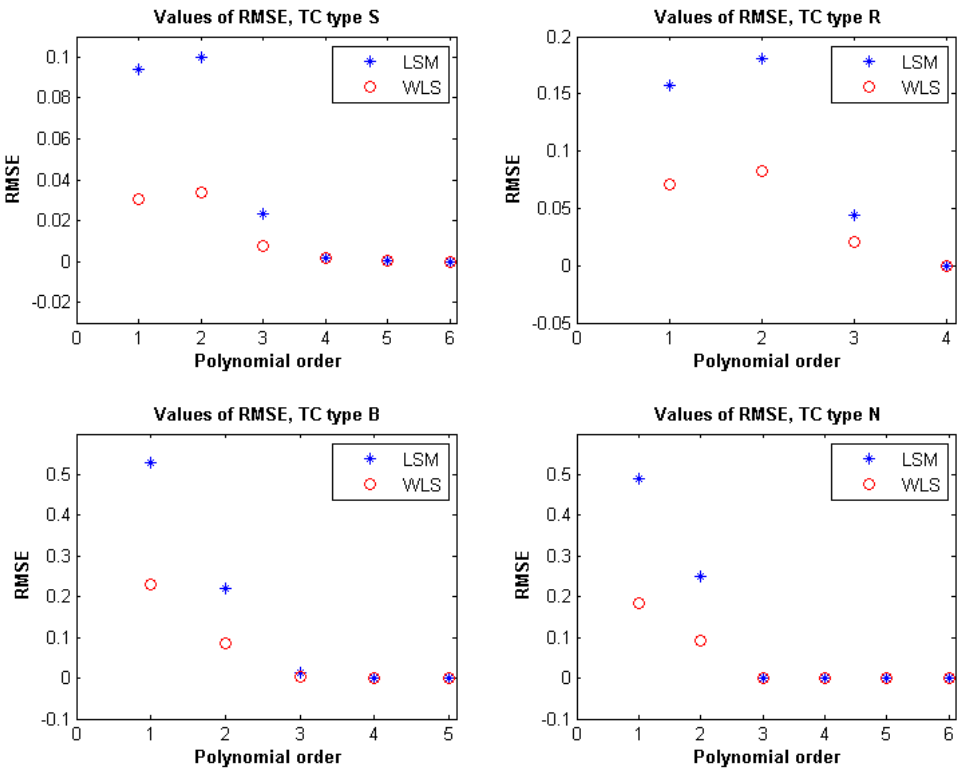

Figure 3: Approximation order determination according to RMSE and comparison of the Method of Least Squares (LSM) and Weighted Least Squares (WLS)

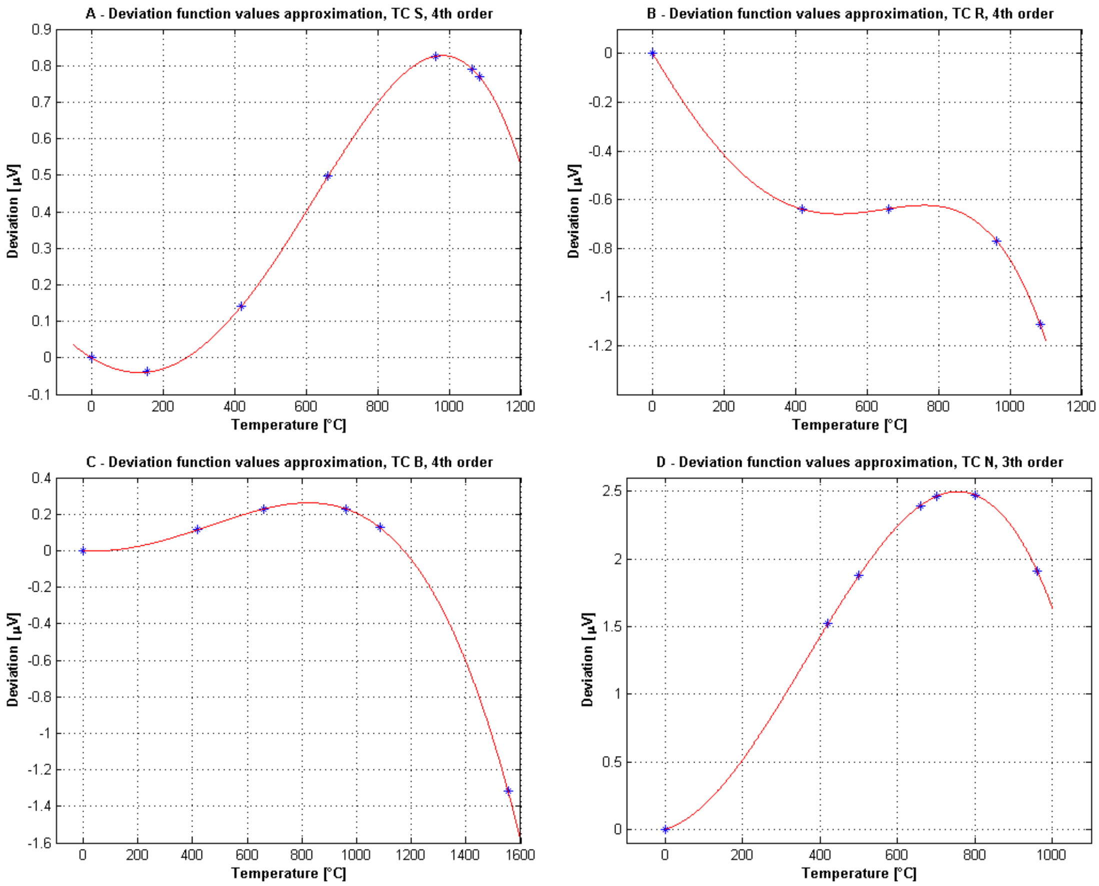

Figure 4: Best approximation of the thermocouples

4. Conclusion

Method of Least Squares and Weighted Least Squares are used for determination of appropriate order of polynomial, which are used for fitting data obtained at calibration of thermocouples. Following the results can be said that WLS is more suitable method for fitting the data. WLS takes into account measurement uncertainties and fit the data with lower errors than Method of Least Squares.

5. References

[1] Nicholas, J.V; White, D.R. *Traceable Temperatures*, 1994, ISBN 0471938033

[2] ČSN EN 60584-1, *Termoelektrické články, Část 1: Referenční tabulky*, December 1997

[3] http://www.efunda.com/math/leastsquares/leastsquares.cfm, October 2011

[4] C.R. Rao, H. Toutenburg, A. Fieger, C. Heumann, T. Nittner and S. Scheid. *Linear Models: Least Squares and Alternatives*. Springer Series in Statistics, 1999, ISBN 0387988483

Acknowledge

This report has been made by support of research study MSM 6046137306 and financial support from specific university research MSMT No 21/2011.

Lenka Kňazovická
Department of Computing and Control Engineering, ICT Prague, Lenka.Knazovicka@vscht.cz

Radek Strnad
Thermometry Department, Czech Metrology Institute, rstrnad@cmi.cz