Jacek MAJEWSKI
Oksana BOYKO

WEIGHTED LEAST-SQUARES POLYNOMIAL APPROXIMATION EMPLOYED TO RH SENSORS CALIBRATION POINTS

ABSTRACT *Weighted least-squares method was applied to derive polynomials which approximate the transfer functions of relative humidity (RH) sensors. The formulae for typical cases were discussed. The exemplary results were compared to the calibration equations obtained by the ordinary least-squares method and those presented in datasheets, and the conclusions were drawn.*

Keywords: *RH sensors calibration, weighted least-squares approximation*

1. INTRODUCTION

Relative humidity can strongly influence many technological processes, and the operation of electrical devices. In order to monitor and control the level of relative humidity, the sensors supplied with known transfer functions (an analytical description of the relationship between the sensor's input and output) are required [6,10]. The equation determining the input-output relationship is obtained as a result of a calibration procedure.

Jacek MAJEWSKI, Ph.D.
e-mail: j.majewski@pollub.pl

Lublin University of Technology

Oksana BOYKO, Ph.D.
e-mail: oxana_bojko@ukr.net

Lviv National Medical University

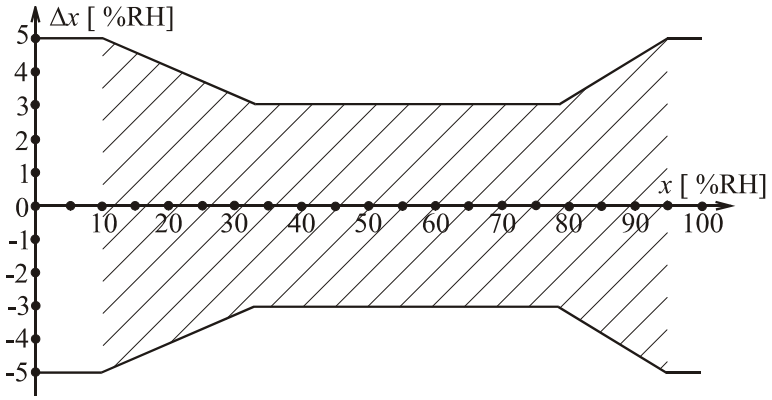

Fig. 1. Typical plot of error area of a RH sensor (the error distribution is quasi-symmetrical within the measurement range 10-95%RH)

The manufacturers of relative humidity (RH) sensors do not describe explicitly the methods employed to determine the calibration equations of the RH sensors offered to customers. It is usual to provide the maximum absolute error expressed as one figure in %RH (this is the unit for RH measuring) valid for the whole measurement range (e.g. 5%RH). Another way is to present a plot of the transfer function, and the maximum error value. For the higher grade sensors, the resulting calibration equation is only presented in the technical data sheets, sometimes accompanied with a tabulated set of calibration points without providing details about the calibration procedure. Some clues can be deduced form the datasheet plot of error area, which reveals that the error values at the endpoints of the measurement range are considerably higher than in the middle (typically about ⅓ – Figure 1).

If the user checks the manufacturer's equation formula by the ordinary least-squares (OLS) method [5] and calculates the transfer function himself, using the set of calibration points given by the manufacturer, sometimes small discrepancies between the datasheet formula and the expression obtained by the OLS method can be noticed. In some cases, these discrepancies can hardly be explained simply by the rounding of the equation coefficients or the numerical values of the calibration points – it seems that a method of approximation other than the OLS was used.

One of basic assumptions made in the OLS method states that variances (or: standard deviations) of the measurements of all the calibration points are equal (this property is called "homoscedasticity" [7, 9]). In the case of RH sensors, if a humidity generator is used, the calibration at high relative humidity is considerably less accurate than at medium and small RH values within the measurement range (Fig. 2). That implies that the weighted least-squares (WLS) method seems more adequate for RH sensors calibration points than OLS.

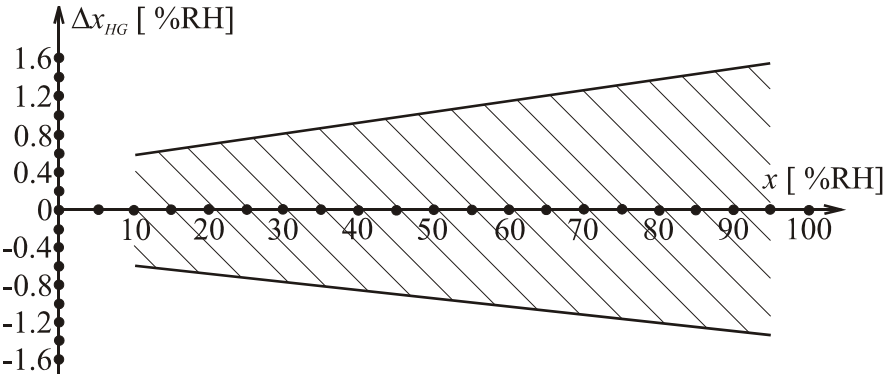

Fig. 2. Typical simplified plot of error area of an RH generator throughout its operating range (the error distribution changes monotonically)

2. THEORETICAL BACKGROUND

Let a set of n calibration points x_i, y_i is tabulated for $i = 1, ..., n$. If the calibration points are heteroscedastic, i.e. obtained with different standard deviations σ_i, $\sigma_1 \neq \sigma_2 \neq ... \neq \sigma_n$, the formulae of the WLS method are valid. In the case of linear approximation, the equations for the coefficients of the approximation line $y = A + Bx$ are calculated using weighted sums substituted into the formulae (1) and (2), where each weight is $w_i = (1 / \sigma_i)^2$. The formulae for A_w and B_w are written as (all summations are on the *i* indices; the summation limits are omitted) [8]:

$$A_w = \frac{\sum wx^2 \sum wy - \sum wx \sum wxy}{\sum w \sum wx^2 - \left(\sum wx\right)^2} \tag{1}$$

$$B_w = \frac{\sum w \sum wxy - \sum xw \sum wy}{\sum w \sum wx^2 - \left(\sum wx\right)^2} \,. \tag{2}$$

In practice, for RH sensors the polynomial approximation is applied. Usually, the polynomial of third degree: $y = A + Bx + Cx^2 + Dx^3$, is applied for sensors supplied with higher numbers of calibration points ($n > 10$). In general, the formulae for the coefficients A, B, C i D are rather complicated. However, if the x_i values are distributed symmetrically about the middle value of x and all the values of x_i can be shifted so that the sum of x's is equal to zero, much simpler formulae can be derived. In the case of RH sensors, when the weights must be taken into account, the condition to be met for simplified resolving of the equation set composed by the WLS method, is that the sum

of the products $w_i\, x_i$ is equal to zero. The full set of four equations for WLS method is expressed as:

$$\begin{cases} A_w \sum w + B_w \sum wx + C_w \sum wx^2 + D_w \sum wx^3 = \sum wy \\ A_w \sum wx + B_w \sum wx^2 + C_w \sum wx^3 + D_w \sum wx^4 = \sum wxy \\ A_w \sum wx^2 + B_w \sum wx^3 + C_w \sum wx^4 + D_w \sum wx^5 = \sum wx^2 y \\ A_w \sum wx^3 + B_w \sum wx^4 + C_w \sum wx^5 + D_w \sum wx^6 = \sum wx^3 y \end{cases} \tag{3}$$

If all terms containing summations of the products of odd powers of x_i multiplied by w_i are cancelled (because of the shift of all x_i by a value of a), the set (3) is largely simplified:

$$\begin{cases} \alpha_w \sum w + \gamma_w \sum w(x-a)^2 = \sum wy \\ \beta_w \sum w(x-a)^2 + \delta_w \sum w(x-a)^4 = \sum w(x-a)y \\ \alpha_w \sum w(x-a)^2 + \gamma_w \sum w(x-a)^4 = \sum w(x-a)^2 y \\ \beta_w \sum w(x-a)^4 + \delta_w \sum w(x-a)^6 = \sum w(x-a)^3 y \end{cases} \tag{4}$$

The set of four equations (4) is equivalent to two sets of two equations: one combined of the first and the third equations from (4), and the other combined of the second and the fourth equation. Although simplified, the formulas for the coefficients α_w, β_w and γ_w are still complicated for computing, and must be transformed back to the true coefficients A_w, B_w and C_w (D_w is equal to δ_w).

3. COMPUTATIONS RESULTS

For a few humidity sensors supplied by the manufacturers with both the sets of calibration points and the calibration equations, the computations by OLS and WLS methods were performed. The exemplary results for linear approximation are tabulated in Table 1.

In Table 1, N denotes the number of calibration points, and r is the correlation coefficient (for $r > 0.995$ the linearity assumption is taken as valid [11]). For two sensors (HS-1101 and HM-1520 – for both $r = 0.9999$) the maximum approximation error obtained by OLS was remarkable lower (ca. 2-5 times) than the value got from the datasheet equation. The difference is also 3 times larger than the maximum value of the possible error caused by the truncation of the

raw values of A and B coefficients of the datasheet equation when rounding them to four significant figures. The origin of this discrepancy between the datasheet equation and the equation obtained by OLS method need to be revealed. However, these differences are not higher than 1/5 of the overall sensor's accuracy and have no strong effect on the sensor's performance.

WLS computations of maximum error of linear approximation do not differ from the values obtained by OLS more than 0.4% RH and do not exceed 1.7% RH (except for the sensor HTF-3226 with the value 4.53% RH for rising weights distribution – but the datasheet of this sensor provides the endpoints accuracy of 10% RH). It should be stressed that all the maximum error values were attained at the endpoints of the input ranges of the sensors.

TABLE 1
Maximum absolute errors of linear approximation for RH sensors, obtained by different methods

Lp.	Sensor symbol	Input range		N	r	Calibration equation obtained:			
						from data sheet	by OLS	by WLS – weights distribution:	
		x_{min}	x_{max}					symmetric	rising
–		% RH		–	–	Maximum absolute error of linear approximation – in % RH			
1	HS-1101	10÷100		10	0.9999	0.42	0.21	0.29	0.30
2	HTF-3226	10÷95		18	0.9980	2.15	3.05	3.68	4.53
3	HTG-35Y3	10÷95		18	0.9997	1.40	1.39	1.72	1.13
4	HTM-2500	10÷95		18	0.9997	1.21	1.23	1.59	1.03
5	HM-1520	1÷20		20	0.9999	0.52	0.10	0.10	0.10

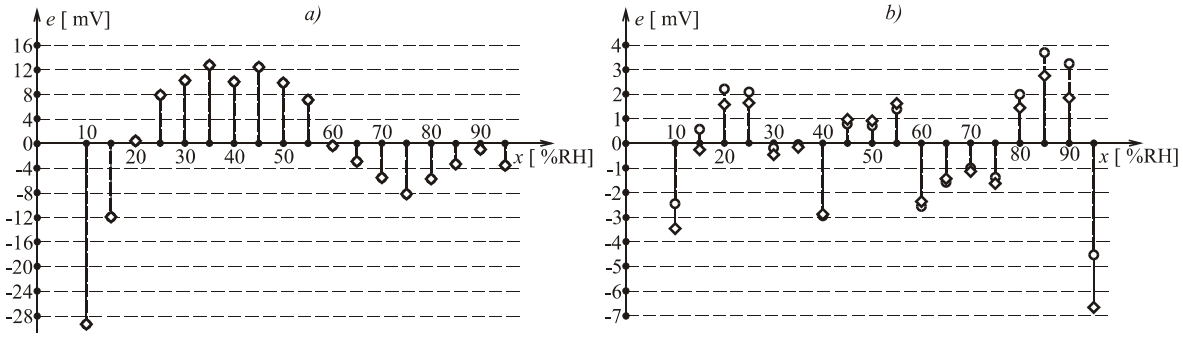

Fig. 3. The plots of residuals for the HTG-35Y3 sensor are shown (empty circles mark the residuals for OLS, and empty rhombi – for datasheet equation):
a) linear approximation (both marks sequences overlap closely);
b) nonlinear (cube) approximation (the differences at the endpoints can be noticed)

In Figure 3(a), the residuals for both the datasheet and OLS equations for the HTG-35Y3 sensor are plotted; the shape of the residual distribution exhibits a kind of regularity which indicates that a third-order approximation would do much better than the linear one.

Residual plots for other RH sensors would reveal similar regularities [3-4]. Perhaps that is the reason of the popularity of third-degree polynomials in the case of RH sensors (quadratic approximation is a rare case). Equidistant calibration points are not the best choice for nonlinear approximation but they are commonly accepted, although there are elaborated special optimisation methods for obtaining non-equidistant set of points for nonlinear calibration [1-2]. The exemplary results for nonlinear approximation are presented in Table 2.

TABLE 2
Maximum absolute error of nonlinear approximation for RH sensors, obtained by different methods

Lp.	Sensor symbol (in superscript the degree of the approximating polynomial is given)	Input range		N	Accuracy	Calibration equation obtained:			
						from data sheet	by OLS	by WLS – weights distribution:	
		x_{min}	x_{max}					symmetric	rising
–		% RH		–	% RH	Maximum absolute error of nonlinear approximation – in %RH			
1	HTF-3226(2)	10÷95		18	5 (10)	0.74	0.76	1.10	3.05
2	HTG-35Y3(3)	10÷95		18	3 (5)	0.32	0.23	0.28	0.28
3	HTM-2500(3)	10÷95		18	3 (5)	0.12	0.11	0.09	0.11
4	HTF-3000(3)	5÷100		20	3 (5)	0.20	0.13	0.14	0.21
5	HS-1101(3)	0÷100		21	3 (5)	0.10	0.09	0.09	0.08

In Table 2, the values in round brackets in the column “Accuracy” are the datasheet values of accuracy at the endpoints of the measurement range (cf. Fig. 1). For all the nonlinear approximations, the errors are considerably (4 to 10 times) smaller than for the linear case. None of the sensors exhibits remarkably higher errors for OLS than for errors calculated from datasheet equations; both rather tend to equality. The differences become almost negligible; they can be explained by rounding errors in the datasheet equation coefficients (A and B to four, and C and D to three significant figures).

For WLS computations, the errors obtained for the symmetric weights distribution are slightly smaller than for the rising one, and do not differ more than 0.1% RH from the errors obtained by OLS method for cube approximation. Only for the sensor HTF-3226 with second-order approximation, the difference is considerable – 0,4% RH and 2.3% RH – in favour of symmetric distribution.

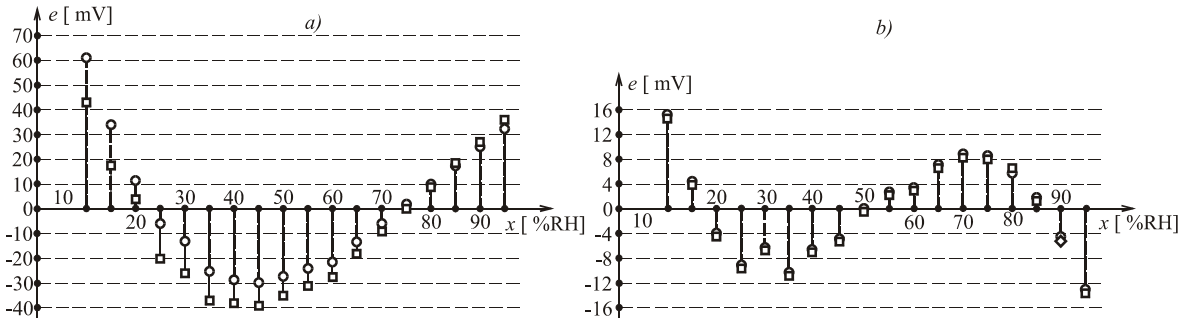

Fig. 4. The plots of residuals for the HTF-3226 sensor are shown (empty circles mark the residuals for OLS, and empty squares – for datasheet equation): a) linear approximation (both marks sequences tend towards each other increasingly); b) nonlinear (quadratic) approximation (the oscillating shape suggests that a third-degree approximation is feasible)

4. CONCLUSIONS

From the tabulated results it can be seen that – for some RH sensors – between the datasheet linear calibration equation and the OLS equations small discrepancies emerge. These discrepancies could be explained by the use of the WLS approach for deriving the datasheet equations but only if a complementary information about the calibration procedure would be supplied by RH sensors manufacturers. As the discrepancies are small for linear approximation, they become negligible for nonlinear (third-degree) polynomial approximation, in comparison with the overall sensor accuracy. That implies that in the case of RH sensors nonlinear calibration equation, the WLS method, although theoretically supported, can be substituted by the OLS method without much influence on the approximation error, and in consequence on the sensor's accuracy.

LITERATURE

1. François N., Govaerts B., Boulanger B.: Optimal designs for inverse prediction of univariate nonlinear calibration models, Chemometrics and Intelligent Laboratory Systems, 2004, Vol. 74.
2. Hajiyev Ch.: Determination of optimum measurement points via A-optimality criterion for the calibration of measurement apparatus, Measurement, 2010, Vol. 43.
3. Kang-Hung H., Chiachung C.: The effect of calibration equations on the uncertainty of UV-Vis spectrophotometric measurements, Measurement, 2010, Vol. 43.

4. Kessel R., Kacker R.N., Sommer K.-D.: Uncertainty budgeting for range calibration, Measurement, 2012, Vol. 45.
5. Liao J.J.Z.: An insight into linear calibration: univariate case, Statistics &Probability Letters, 2002, Vol. 56.
6. Pi-Guey S., Ren-Jang W.: Uncertainty of humidity sensors testing by means of divided-flow generator, Measurement, 2004, Vol. 36.
7. Piotrowski J., Kostyrko K.: Wzorcowanie aparatury pomiarowej, Wydawnictwo Naukowe PWN, Warszawa, 2000.
8. Taylor J.P.: An Introduction to Error Analysis. Second ed. (revised), in Polish, PWN, Warszawa, 1999.
9. Tellinghuisen J.: Weighted least squares in calibration: The problem with using ”quality coefficients” to select weighting formulas, Journal of Chromatography B, 2008, Vol. 872.
10. Tzehung L., Chiachung C.: Uncertainty evaluation of humidity sensors calibrated by saturated salt solutions, Measurement, 2007, Vol. 40.
11. Van Loco J. et al.: Linearity of calibration curves: use and misuse of the correlation coefficient, Accreditation Quality Assurance, 2002, Vol. 7.

Manuscript submitted 27.07.2012

ZASTOSOWANIE APROKSYMACJI WIELOMIANOWEJ METODĄ WAŻONYCH NAJMNIEJSZYCH KWADRATÓW DO PUNKTÓW KALIBRACYJNYCH SENSORÓW WILGOTNOŚCI WZGLĘDNEJ

Jacek MAJEWSKI, Oksana BOYKO

STRESZCZENIE *Metodę ważonych najmniejszych kwadratów zastosowano do wielomianowej aproksymacji funkcji przetwarzania sensorów wilgotności względnej (RH). Przedyskutowano wybrane zależności odpowiadające typowym przypadkom aproksymacji. Wyniki otrzymane dla wybranych sensorów RH porównano z równaniami kalibracyjnymi otrzymanymi zwykłą metodą najmniejszych kwadratów oraz podanymi przez wytwórców, i podano wnioski.*

Słowa kluczowe: *kalibracja sensorów wilgotności względnej, aproksymacja metodą ważonych najmniejszych kwadratów*