Curve Fitting of Sensors’ Characteristics

Boyanka Marinova Nikolova, Georgi Todorov Nikolov and Milen Hrabarov Todorov

Abstract - More complex technical systems and higher levels of integration of electronic circuits lead to new requirements on the data treatment of modern signal processing systems, especially for sensing systems. This paper presents curve fitting methodology for analogue sensor’s characteristics described by datasheet lookup tables. This is achieved by MATLAB’s Curve Fitting Toolbox, which provides a library of linear, nonlinear, and nonparametric fitting models. A method for implemented polynomial equations as calibration template in transducer electronic data sheets is suggested in the end of presentation.

Keywords – Curve fitting, Goodness of fit statistics, MATLAB, Intelligent sensors, TEDS

I. INTRODUCTION

At the core of any data acquisition system is interpretation of a voltage signal based on information about the analog sensor that makes the measurement intelligible. Typically, these are standard curves and equations specific to the type of transducer. Sensor calibration, however, takes this process one step further by considering transducers on an individual basis. The sensor output voltage is mapped to a physical measurement based on metrics obtained from a specific sensor calibration. Although many sensors are linear over the limited range, these sensors exhibit a slight but progressively more nonlinear characteristic as the measurement range widens. Consequently, over an extended span, curve fitting is necessary if the system is to achieve a high level of precision.

With the facility of computation now available through digital computers and microprocessors, the problem of estimation of transducer’s transfer characteristics is being increasingly tackled using software techniques. However, for inherent nonlinear sensors, a software solution depends upon the proper approach through mathematical modeling of the response curve [5, 9].

The purpose of this paper is to assist engineers and scientists to implement the newly released Curve Fitting Toolbox in order to achieve more precise results. The presented material facilitates users to create calibration equations from basic calibration data and use these equations to make accurate measurements.

II. PARAMETRIC CURVE FITTING

Parametric fitting involves finding coefficients (parameters) for one or more models that fit to data [1, 2]. The model is a function of the independent variable and one or more coefficients. The appropriate mathematical model of sensors characteristics is can be obtained by Curve Fitting Toolbox™ software. This toolbox is a collection of graphical user interfaces and M-file functions for curve and surface fitting that operate in the MATLAB® technical computing environment. The toolbox includes a library of various parametric models, summarized in table 1. [3, 4]

Table 1. Curve Fitting Toolbox Library Models

<table>
<thead>
<tr>
<th>Model</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exponentials</td>
<td>( y = a e^{bx} ) and ( y = a e^{bx} + c e^{dx} ), where ( a, b, c, ) and ( d ) are model’s parameters</td>
</tr>
<tr>
<td>Fourier Series</td>
<td>models a intercept term in the data and is associated with the ( i = 0 ) cosine term, ( \omega ) is the fundamental frequency of the signal, ( n ) is the number of terms (harmonics) in the series, and ( 1 \leq n \leq 8 )</td>
</tr>
<tr>
<td>Gaussian</td>
<td>( y = \sum_{i=1}^{n} a_i e^{-\frac{(x-h_i)^2}{c_i^2}} ), where ( a_i ) is the amplitude, ( h_i ) is the centroid (location), ( c_i ) is related to the peak width, ( n ) is the number of peaks to fit, and ( 1 \leq n \leq 8 )</td>
</tr>
<tr>
<td>Polynomials</td>
<td>( y = \sum_{i=1}^{n} p_i x^{n-i} ), where ( p_i ) are model’s parameters, ( n+1 ) is the order of the polynomial, ( n ) is the degree of the polynomial, and ( 1 \leq n \leq 8 )</td>
</tr>
<tr>
<td>Power Series</td>
<td>( y = a x^b ) and ( y = a + b x^c ), where ( a, b, ) and ( c ) are model’s parameters</td>
</tr>
<tr>
<td>Rationals</td>
<td>( y = \sum_{i=1}^{n} p_i x^{n+i} ), where ( q_i, ) and ( p_i ) are model’s parameters, ( n ) and ( m ) are the degree of the numerator and denominator polynomials – ( 1 \leq n \leq 5 ) and ( 1 \leq m \leq 5 )</td>
</tr>
<tr>
<td>Sum of Sines</td>
<td>( y = \sum_{i=1}^{m} a_i \sin(h_i x + c_i) ), where ( a_i ) is the amplitude, ( h_i ) is the frequency, and ( c_i ) is the phase constant for each sine wave term, ( n ) is the number of terms in the series and ( 1 \leq n \leq 8 )</td>
</tr>
<tr>
<td>Weibull</td>
<td>( y = a b x^{n-1} e^{-ax^b} ), where ( a ) is the scale parameter and ( b ) is the shape parameter</td>
</tr>
</tbody>
</table>

Curve Fitting Toolbox software uses the method of least squares when fitting data. The supported types of least squares fitting include: linear, weighted linear, constrained, robust, and nonlinear [1, 2, 3, 4].
III. CURVE FITTING METHODOLOGY

The proposed methodology for curve fitting of sensors’ characteristics is illustrated with the block diagram given on fig. 1. A particular application might dictate still other aspects of model fitting that are important to achieving a good fit, such as a simple model that is easy to interpret. On the first step Importing Data the chosen characteristic must be represented as the predictor (X) data, response (Y) data, and weights. If the weights are not imported, then they are assumed to be 1 for all data points. The data may be taking either from the datasheets of sensor’s manufacturer, or by measurements of the particular sensor.

On the next step Fitting Data one model from the library of parametric models listed in table 1 is selected. According to the shape and specificity of the sensor characteristic the appropriate selection must be done. The fitting procedure can be successful in short time if the implemented.

According to the shape and specificity of the sensor characteristic the appropriate selection must be done. The proposed methodology for curve fitting of sensors’ characteristics is illustrated with the block diagram given on fig. 1. A particular application might dictate still other aspects of model fitting that are important to achieving a good fit, such as a simple model that is easy to interpret. On the first step Importing Data the chosen characteristic must be represented as the predictor (X) data, response (Y) data, and weights. If the weights are not imported, then they are assumed to be 1 for all data points. The data may be taking either from the datasheets of sensor’s manufacturer, or by measurements of the particular sensor.

The sum of squares due to error measures the total deviation of the response values from the fit to the response values. It is also called the summed square of residuals and is given with equation:

\[ SSE = \sum_{i=1}^{n} \omega_i (y_i - \hat{y}_i)^2 \]  

where \( \omega_i \) are the weights. The weights determine how much each response value influences the final parameter estimates. If a value of SSE is closer to 0 indicates that the model has a smaller random error component, and that the fit will be more useful for prediction.

Coefficient of determination (R-square) determining how successful the fit is in explaining the variation of the data. R-square is the square of the correlation between the response values and the predicted response values:

\[ R^2 = \frac{\sum_{i=1}^{n} \omega_i (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} \omega_i (y_i - \bar{y})^2} \]  

where \( \bar{y} \) is mean value of \( y_i \). R-square with a value closer to 1 indicating that the model accounts for a greater proportion of variance.

If the number of fitted coefficients increases in the model, \( R^2 \) will increase although the fit may not improve in a practical sense. To avoid this situation, should be used the degrees of freedom adjusted \( R^2 \) statistic. This statistic is based on the residual degrees of freedom \( v \) defined as the

- Importing Data
- Fitting Data
- Assessing Goodness of Fit
  - Residuals Plot Analysis
  - Goodness of fit statistics
  - Confidence and Prediction Bounds
- Saving the Fit Results

Fig. 1. The block diagram of the curve fitting methodology.
number of response values \( n \) minus the number of fitted coefficients \( m \) estimated from the response values:

\[
v = n - m,
\]

where \( v \) indicates the number of independent pieces of information involving the \( n \) data points that are required to calculate the sum of squares. Note that if parameters are bounded and one or more of the estimates are at their bounds, then those estimates are regarded as fixed. The number of such parameters increases the degrees of freedom. The adjusted \( R^2 \) statistic is given:

\[
Adjusted \ R^2 = 1 - \frac{\sum_{i=1}^{n-1} e_i^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2}.
\]  

The adjusted R-square statistic has a value closer to 1 indicating a better fit. Negative values can occur when the model contains terms that do not help to predict the response.

Root mean squared error (RMSE) is also known as the fit standard error and the standard error of the regression. It is an estimate of the standard deviation of the random component in the data, and is defined as

\[
RMSE = \sqrt{\frac{\sum_{i=1}^{n} e_i^2}{v}}.
\]  

Just as RMSE value closer to 0 indicates a fit that is more useful for prediction.

Confidence and prediction bounds define the lower and upper values of the associated interval, and define the width of the interval. The width of the interval indicates how uncertain have about the fitted coefficients, the predicted observation, or the predicted fit. A very wide interval for the fitted coefficients indicate that should be used more data when fitting.

The saved information on the last step Saving the Fit Results, can be used for documentation purposes, or to extend data exploration and analysis.

### IV. APPLICATION OF SENSORS CURVE FITTING

IEEE 1451.4 is a standard that defines a relatively simple, straightforward mechanism for adding intelligent, plug and play capabilities to traditional analog sensors. Without adding any new hardware to the system, these plug and play sensors can bring real, immediate benefits in ease of use and productivity to any measurement and automation system that uses sensors [6, 7]. The underlying mechanism for plug and play identification is the standardization of a Transducer Electronic Data Sheet (TEDS). The information stored in TEDS enables the system to identify, characterize, interface, and properly use the signal from the analog sensor. In addition to such configuration information TEDS can store calibration data specific to an individual sensor. Sensor calibration considers transducers on an individual basis. In other words, the sensor output voltage is mapped to a physical measurement based on metrics obtained from a specific sensor calibration.

The data stored in the sensor’s EEPROM or a virtual TEDS file is compressed to save space [7]. There are sixteen separate standard templates for each type of sensor. In addition to these standard templates, there are three calibration templates: a calibration table, frequency response table and polynomial calibration curve.

The calibration table template describes a lookup table with the electrical-to-physical transfer function of a transducer. The calibration table template provides a simple means of recording a few data points for calibration purposes. This is particularly useful when calibrating around a very narrow range of values.

The frequency response table template specifies the frequency response transfer function of a transducer with a lookup table of frequency-amplitude pairs.

The aim of presented investigation is to precisely characterize a transducer over its entire range. Most appropriate for such initiative is polynomial calibration curve. This may prove difficult, especially in the case of a transducer with a highly nonlinear transfer function, such as a thermocouple or thermistor. To make effective use of the calibration curve template, several calibration points must be taken and then fitted to a polynomial curve. This may take significantly more time than collecting a few data points for a calibration table. However, the tradeoff is greater accuracy and a wider range of calibrated values.

Using the proposed approach the transfer characteristic of thermistor PR222J2 is fitted with appropriate parametric model. For input data of temperature dependence is used the detailed look up table given in datasheet of thermistor’s manufacturer U.S. SENSOR Inc. [8].

![Fig. 2. The data, fit, prediction bounds and residuals for transfer characteristic of thermistor PR222J2.](image)
In TEDS template the overall function is described in a piecewise manner, each segment defined by an array of coefficients and powers. For example, a polynomial curve from table 2 may be described by \((9, -5.91E-18), (8, 5.69E-16), (7, 2.92E-14), (6, -2.46E-12), (5, -1.74E-10), (4, 1.81E-8), (3, -9.11E-8), (2, 4.38E-6), (1, -4.18E-4), (0, 8.14E-3)]\), \((3.1), (2.5), (1, -2), (0, 1)\], where the numbers in parentheses are interpreted as (Curve Power, Curve Coefficient).

In order to achieve better accuracy, especially for more non-linear sensors, TEDS provides an alternative means of calibration and linearization. The calibration curve template specifies a multi-segment polynomial curve and each segment is bounded by sequential values of, closed below and open above [6, 7]. The implementation of this capability in the case of treated thermistor is shown in last three columns in Table 2. As can be seen from the numerical fit results that the SSE and RMSE statistics have values very close to zero. Therefore the parametric models of divided transfer characteristic of the thermistor much better fits the input data.

V. CONCLUSION

The present paper suggests an systematic approach for implementing curve fitting models and methods in order to achieve equation that precisely describe sensor’s transfer function. The aim of such equation is to implement it in calibration TEDS template and in such a way to improve measurement accuracy. To proof usability of presented approach a number of equation for thermocouples, RTDs and thermistors have been achieved. As illustration experimental results from curve fitting of thermistor’s transfer characteristic is presented. The obtained parametric models are suitable for calibration and transfer functions in measurement system of non-electric quantities or for behavior modeling of various sensors.

This investigation has been carried out in the framework of the research projects № 091 НН 109-07 and № Д002-126/2008.

### TABLE 2. RESULTS FROM CURVE FITTING OF THERMISTOR’S TRANSFER CHARACTERISTIC

<table>
<thead>
<tr>
<th>Model</th>
<th>Full Characteristic</th>
<th>Divided Characteristic</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Equation</td>
<td>9th Degree Polynomial (Poly9)</td>
</tr>
<tr>
<td></td>
<td>( f(x) = ae^{bx} + ce^{dx} )</td>
<td>( f(x) = p_1x^9 + p_2x^8 + p_3x^7 + p_4x^6 + p_5x^5 + p_6x^4 + p_7x^3 + p_8x^2 + p_9x + p_{10} )</td>
</tr>
<tr>
<td>Parameters</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>0.0002613</td>
<td>( p_1 )</td>
</tr>
<tr>
<td>b</td>
<td>-0.1048</td>
<td>( p_2 )</td>
</tr>
<tr>
<td>c</td>
<td>0.006674</td>
<td>( p_3 )</td>
</tr>
<tr>
<td>d</td>
<td>-0.0545</td>
<td>( p_4 )</td>
</tr>
<tr>
<td>( p_6 )</td>
<td>1.181.10^{-8}</td>
<td>36.79</td>
</tr>
<tr>
<td>( p_r )</td>
<td>-9.112.10^{-8}</td>
<td>1810.0</td>
</tr>
<tr>
<td>( p_s )</td>
<td>4.382.10^{-6}</td>
<td>5.725.10^{-6}</td>
</tr>
<tr>
<td>( p_b )</td>
<td>-0.0004184</td>
<td>1.056.10^{-6}</td>
</tr>
<tr>
<td>( p_{10} )</td>
<td>0.008147</td>
<td>8.648.10^{-6}</td>
</tr>
<tr>
<td>Goodness of fit</td>
<td>SSE</td>
<td>2.936.10^{-5}</td>
</tr>
<tr>
<td>( R^2 )</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Adj. ( R^2 )</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>RMSE</td>
<td>4.325.10^{-4}</td>
<td>0.8947.10^{-3}</td>
</tr>
</tbody>
</table>

In TEDS template the overall function is described in a piecewise manner, each segment defined by an array of coefficients and powers. For example, a polynomial curve from table 2 may be described by \((9, -5.91E-18), (8, 5.69E-16), (7, 2.92E-14), (6, -2.46E-12), (5, -1.74E-10), (4, 1.18E-8), (3, -9.11E-8), (2, 4.38E-6), (1, -4.18E-4), (0, 8.14E-3)]\), \((3.1), (2.5), (1, -2), (0, 1)\], where the numbers in parentheses are interpreted as (Curve Power, Curve Coefficient).

In order to achieve better accuracy, especially for more non-linear sensors, TEDS provides an alternative means of calibration and linearization. The calibration curve template specifies a multi-segment polynomial curve and each segment is bounded by sequential values of, closed below and open above [6, 7]. The implementation of this capability in the case of treated thermistor is shown in last three columns in Table 2. As can be seen from the numerical fit results that the SSE and RMSE statistics have values very close to zero. Therefore the parametric models of divided transfer characteristic of the thermistor much better fits the input data.

### REFERENCES