Curve Fitting for Sensors' Analog Behavioural Modelling

Boyanka Nikolova[1], Milen Todorov[2], Tihomir Brusev[3]

Abstract – **This article summarizes the main methods for curve fitting of non-linear sensor's characteristics. The results are input data in SPICE model, built according to the basic measuring circuit within the sensor's product information. The curve fitting is achieved by MathWorks MATLAB®'s Curve Fitting Toolbox[TM] and National Instruments Multisim. Curve Fitting Toolbox[TM] supplements MATLAB features with data preprocessing capabilities, using parametric and nonparametric models from a library. On the other hand Multisim provides the necessary ABM (analog behavioral model) sources, which use mathematical and conditional expressions to set their output voltage or current. They may contain mathematical and conditional expressions that consist of circuit voltages, currents, time and other simulation parameters. ABM is an extremely powerful feature which provides an efficient way to macro model signal processes through non-linear mathematical and conditional expressions.**

Keywords – **Curve fitting, Simulation methodology, Intelligent sensors, MATLAB, Multisim.**

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 farther 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 [1, 2, 6].

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. These results can be used as input data for the simulation models.

[1]Boyanka Nikolova is with the Faculty of Telecommunications, Technical University of Sofia, Kl. Ohridski 8, 1797 Sofia, Bulgaria, E-mail: bnikol@tu-sofia.bg

[2]Milen Todorov is with the Faculty of Telecommunications, Technical University of Sofia, Kl. Ohridski 8, 1797 Sofia, Bulgaria, E-mail: brusev@ecad.tu-sofia.bg

[3]TihomirBrusev is with the Faculty of Telecommunications, Technical University of Sofia, Kl. Ohridski 8, 1797 Sofia, Bulgaria, E-mail: brusev@ecad.tu-sofia.bg

II. PARAMETRIC MODELS IN MATLAB

TABLE I
CURVE FITTING TOOLBOX LIBRARY MODELS

Type of Fit	Description
Exponentials $y = ae^{bx}$; $y = ae^{bx} + ce^{dx}$	**a, b, c, d** – model's parameters.
Fourier Series $y = a_0 + \sum_{i=1}^{n} a_i \cos(nwx) + b_i \sin(nwx)$	a_0 – models a intercept term in the data and is associated with the i = 0 cosine term; **w** – fundamental frequency of the signal; **n** – number of terms (harmonics) in the series and $n \in [1; 8]$.
Gaussian $y = \sum_{i=1}^{n} a_i e^{\left[-\left(\frac{x-b_i}{c_i} \right)^2 \right]}$	**a** – amplitude; **b** – centroid (location); **c** – related to the peak width **n** – number of peaks to fit, $n \in [1; 8]$.
Polynomials $y = \sum_{i=1}^{n+1} p_i x^{n+1-i}$	**n+1** – order of the polynomial; **n** – degree of the polynomial, $n \in [1; 9]$.
Power Series $y = ax^b$; $y = a + bx^c$	**a, b, c** – model's parameters.
Rationals $y = \dfrac{\sum_{i=1}^{n+1} p_i x^{n+1-i}}{x^m + \sum_{i=1}^{m} q_i x^{m-i}}$	**n** – degree of the numerator polynomial, $n \in [0; 5]$; **m** – degree of the denominator polynomial, $m \in [1; 5]$.
Sum of Sines $y = \sum_{i=1}^{n} a_i \sin(b_i x + c_i)$	**a** – amplitude; **b** – frequency; **c** – phase constant for each sine wave term; **n** – number of terms in the series.

Parametric fitting involves finding coefficients (parameters) for one or more models that fit to data [4, 8]. The data is assumed to be statistical in nature and is divided into two components: a deterministic component and a random component [3]. The deterministic component is given by a parametric model and the random component is often described as error associated with the data. The model is a function of the independent (predictor) variable and one or more coefficients. The error represents random variations in the data that follow a specific probability distribution. Systematic variations can also exist, but they will lead to a fitted model that does not represent the data well. To improve

the fit it may be necessary to increase the number of iterations. The fit is well behaved over the entire data range when the residuals are randomly scattered about zero.

After fitting data with one or more models is necessary to evaluate the goodness of fit. A visual examination of the fitted curve displayed in Curve Fitting Tool should be first step. Beyond that, the toolbox provides these methods to assess goodness of fit for both linear and nonlinear parametric fits [3, 4, 8, 9]:

- Residual analysis
- Goodness of fit statistics
- Confidence and prediction bounds.

These methods group into two types: graphical and numerical. Plotting residuals and prediction bounds are graphical methods that aid visual interpretation, while computing goodness of fit statistics and coefficient confidence bounds yield numerical measures that aid statistical reasoning.

Graphical measures are more beneficial than numerical measures because they allow to view the entire data set at once, and they can easily display a wide range of relationships between the model and the data. The numerical measures are more narrowly focused on a particular aspect of the data nad often try to compress that information into a single number. In practice, depending on data and analysis requirements it is necessary to use both types to determine the best fit.

When fitting data that contains random variations, there are two important assumptions that are usually made about the error:

- The error exists only in the response data, and not in the predictor data.
- The errors are random and follow a normal (Gaussian) distribution with zero mean and constant variance, σ^2.

The errors are assumed to be normally distributed because the normal distribution often provides an adequate approximation to the distribution of many measured quantities.

III. CURVE FITTING AND SIMULATION

On figure 1 is illustrated the proposed methodology for curve fitting and modeling methodology.

On step 1.1 the chosen characteristic must be represented as workspace vectors with the same length. To perform any curve fitting task, must be selected at least one vector of data:

- **X data** – selects the predictor data.
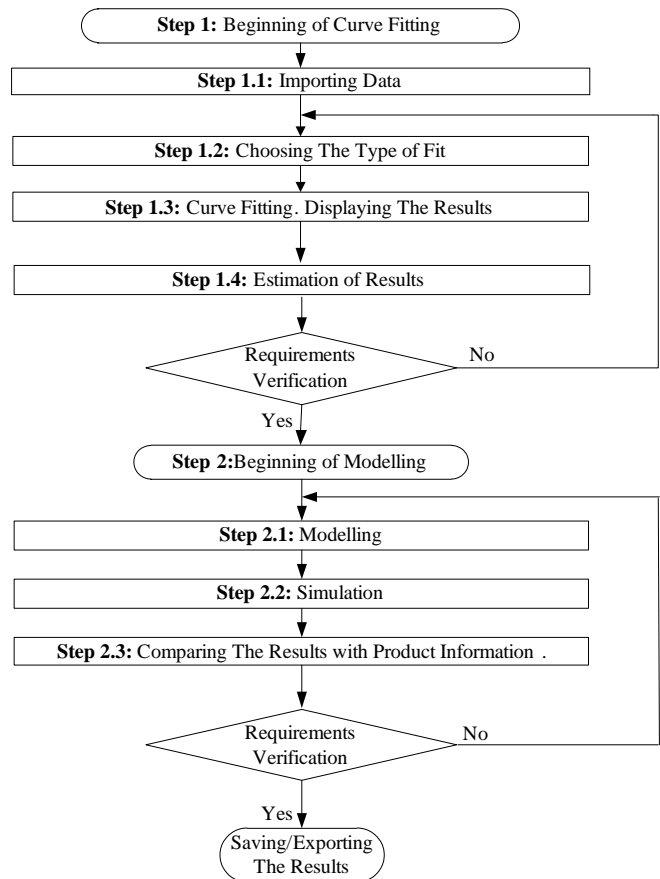- **Y data** – selects the response data.

Fig. 1. Block diagram of curve fitting and modeling methodology.

Also it could be set the weights associated with the response data. If weights are not imported, they are assumed to be 1 for all data points.

On the next step 1.2 we choose one model from the library of parametric models. According to the shape and specificity of the sensor characteristic the appropriate selection must be done.

Step 1.3 displays the results of fitting.

The importance of step 1.4 consists in the analysis results. It must be examine the fitted curve, residuals, goodness of fit statistics, confidence bounds, and prediction bounds for the current fit. It is recommended to compare the current fit and data set to previous fits and data sets by examining the goodness of fit statistics. Figure 2 shows this in detail.

On step 2.1 the parts from sensor's basic measuring circuit should be substitute by ABM sources with. As input data for ABM sources is used the results of fitting. The idea of this is seen in detail on figure 3.

Once the model is ready we can run the simulation with given parameters – step 2.2.

The basic structure of the proposed model of gas sensor is shown on figure 4. The main purpose of model is to simulate the change of the output voltage as slope function of the input physical phenomenon. This is achieved with a block "Transfer Function". The environment is presented by block "Externalities". Changing the signal due to external factors (mainly temperature and relative humidity) are calculated in "Externalities" block and added to the sensor's signal by voltage summer.
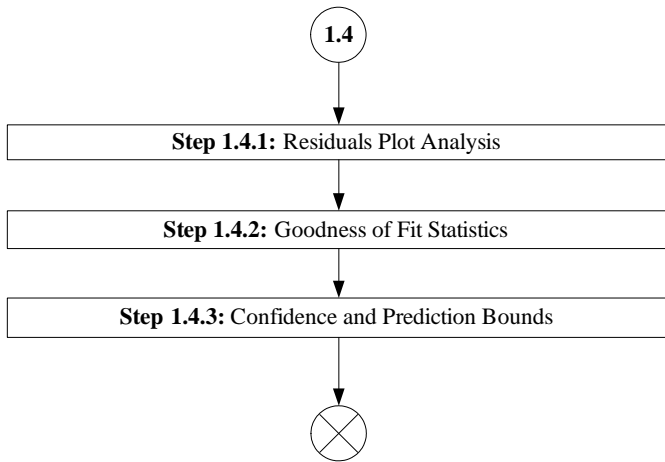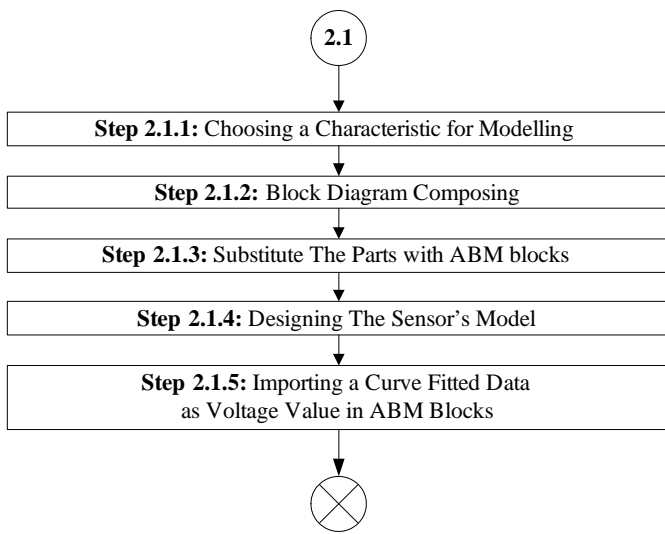
Fig. 2. Estimation of results.

Fig. 3. Modelling.

TABLE II
CURVE FITTING TOOLBOX LIBRARY MODELS

Goodness of fit	Best-case value
Sum of Squares due to Error (SSE)	0
R-square (R^2)	1
Adjusted R-square (Adjusted R^2)	1
Root Mean Squared Error	0

The transfer function and other characteristics, which describe the effects of temperature and humidity, can easily be modeled by appropriate behavioral elements. The behavioral elements use mathematical and conditional expressions to set their output voltage or current. They contain mathematical expressions obtained by curve fitting. The main advantage of using a mathematical function is short simulation time and better convergence [7].
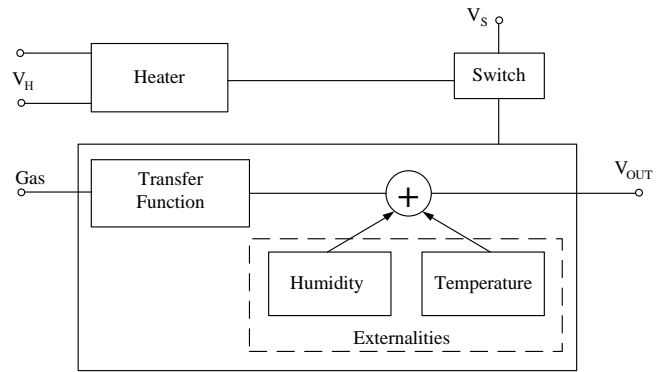

Fig. 4. Gas sensor's model.

Using the existing library of ABM blocks in National Instruments Multisim and figure 4 is developed behavioral model of TGS 2600 gas sensor – figure 5 [5, 7].
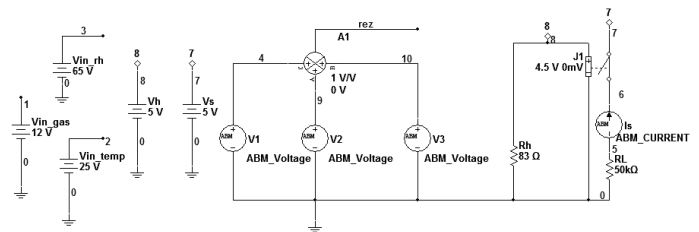

Fig. 5. Gas sensor's model in NI Multisim.

The data, fit, prediction bounds and residuals for fitting of full transfer characteristics are shown in figure 6. In table III are given the fitting numerical results for model's parameters and goodness of fit statistics. For a first fitting model 9th degree polynomial (Poly9) is chosen. For a second model is chosen power model with three parameters (Power2). As shown by results the residuals for Power Fit 2 are randomly scattered around zero and indicating that this model describes the data well. The same conclusion can be done comparing numerical values of goodness of fit statistics.
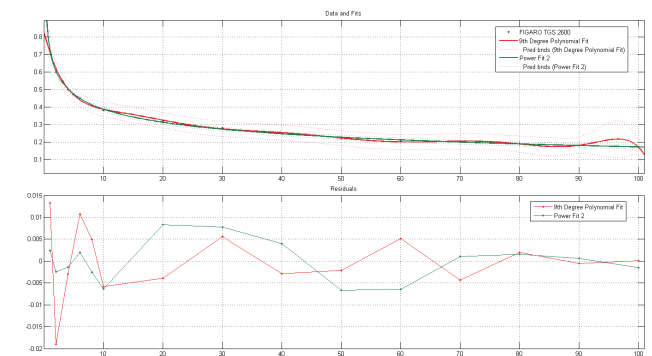

Fig. 6. The data, fit, prediction bounds and residuals for transfer characteristic of TGS 2600 sensor.

TABLE III
RESULTS FROM CURVE FITTING OF TGS 2600 TRANSFER CHARACTERISTIC

Type of fit	9th Degree Polynomial		Power Fit 2	
Model	$f(x) = p1.x^9 + p2.x^8 + p3.x^7 + p4.x^6 +$ $+p5.x^5 + p6.x^4 + p7.x^3 + p8.x^2 + p9.x + p10$		$f(x) = a.x^b + c$	
Coefficients	p1	-2.115e-015	a	0.9434
	p2	9.705e-013	b	-0.1878
	p3	-1.878e-010	c	-0.2257
	p4	1.997e-008		
	p5	-1.273e-006		
	p6	5e-005		
	p7	-0.001195		
	p8	0.01671		
	p9	-0.1302		
	p10	0.8215		
Goodness of fit	SSE	0.0008261		0.0003024
	R-square	0.9979		0.9992
	Adjusted R-square	0.9942		0.9991
	RMSE	0.01285		0.00502

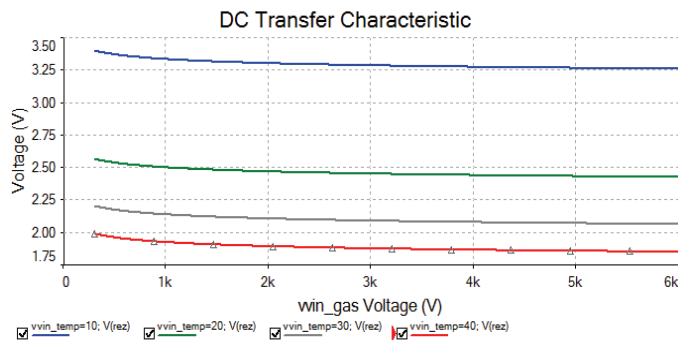Results of the simulation process in NI Multisim are shown in figures 7 and 8.

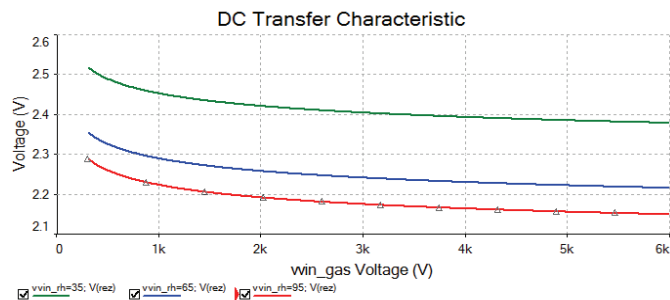Fig. 7. Sensor's transfer function under the influence of temperature.

Fig. 8. Sensor's transfer function under the influence of relative humidity.

IV. CONCLUSION

This paper suggest an systematic approach for implementing curve fitting models and methods in order to achieve equation that precisely describe sensor's transfer characteristic and used as initial conditions for simulation process. As illustration experimental results from the whole process are presented.

REFERENCES

[1] B. Nikolova, G. Nikolov, M. Todorov. "Curve Fitting of Sensors' Characteristics", Annual Journal of Electronics, ISSN 1313-1842, Vol. 3, Number 1, pp.188-191, Sofia, 2009.
[2] B. Nikolova, G. Nikolov. "Design, Development and Calibration of Virtual System for Relative Humidity Measurement. Sensors & Transducers Journal, ISSN 1726-5479, Vol.93, Issue 6. Pp. 1-14, June 2008.
[3] *Curve Fitting Toolbox*[TM] *3 User's Guide.* The MathWorks, Inc., 2001-2010.
[4] F. Riley, P. Hobson, J. Bence, *Mathematical Methods for Physics and Engineering.* Cambridge University Press, New York, U.S.A., 2006.
[5] FIGARO. TGS 2600 – for the detection of Air Contaminants. Product Information.
[6] J. Fraden, *Handbook of Modern Sensors: Physics, Designs, and Applications.* Springer, San Diego, California, U.S.A., 2004.
[7] Multisim. Component Reference Guide. National Instruments Corporation, January 2007.
[8] S. T. Karris, *Mathematics for Business, Science, and Technology with MATLAB*[®] *and Excel*[®] *Computations.* Orchard Publications, California, U.S.A., 2007.
[9] S. Weisberg, *Applied Linear Regression,* John Wiley & Sons, Inc., New Jersey, U.S.A., 2005.