ISSN 2229-5518

Assessing the Effect of Minimization Error Functions on Curves and Models Fitting: A SOLVER Approach

Anietie N. Okon*, Uwem E. Inyang, Innocent O. Oboh

Abstract— In science and engineering research, most observed (experimental) data are fit to mathematical model(s): linear or nonlinear. Mostly, the statistical error test – least square method is used to minimize the difference between the experimental and model-predicted data to achieve a good fit. In this study, some statistical error functions: average, absolute, mean square, root mean square, Chi-square and normalized root mean square were used as error minimization approach to fit experimental (i.e., viscosity and shear rate) data to Carreau-Yasuda model using SOLVER. The values of the model's adjustable parameters obtained from average, absolute and Chi-square errors were different while the sum square, mean square, root mean square and normalize root mean square errors resulted in the same values. Furthermore, the goodness and acceptability of the fitted model were established using the coefficient of determination (R^2) and F-test (F_{test}) value. From the fitted model R^2 results obtained, the average error had 0.9585 while other error functions had R^2 of 0.9998. The Ftest values obtained showed that the average error had 0.9756; this was greater than its F-critical value of 0.6365, while the sum of square errors had 1.0002, which was less than their F-critical value of 1.5709. Hence, the fitted model based on sum of square error functions was more fitted and accepted than model fitted with the average error functions. Therefore, sum of square, mean square, root mean square and normalize root mean square error functions are good error minimization approach to fit nonlinear models and curves to experimental data.

Keywords— Experimental data, Models and curves fitting, Error minimization approach, Error functions, SOLVER.

---

1 Introduction

The model fitting to a set of data is one of the common and frequently carried out tasks in many disciplines of science and engineering [1]. Generally, it is the process of constructing a mathematical function (model) that has the best fit to a series of data points, possibly subjected to constraint [2]. Thus, the goodness of fit is an essentially important parameter that estimates how well the model (i.e., the prediction) pronounces the experimental data. Mostly, the least-squares method is used to measure the goodness of fit. This is based on the theory that the scale of the difference between the experimental data points and the model prediction is a good measure of how well the model fits the data [3]. Before now, nonlinear data would be changed into a linear form and consequently analysed by the least-squares fit approach. This analysis could yield inaccurate measurements and predictions of the data and may alter the experimental error or alter the relationship between the 'independent' and 'dependent' variables. Regrettably, this approach of fitting nonlinear data is erroneous and old fashioned which should not be applicable. For the nonlinear data, it is important to apply a protocol that will fit a nonlinear model to the data. With the advent of the computer, a suitable method for this protocol (algorithm) is called iterative nonlinear least-squares fitting.

According to Ahn [1], when we are particularly interested in fitting a set of measurement points to the nonlinear model, the least-squares model fitting with the error measure is called model fitting in the literature. The model-fitting problem is a nonlinear minimization problem to be solved through iteration and has been widely recognized as an analytically and computationally difficult problem. Interestingly, several algorithms had been developed that are used in nonlinear estimation; these include the Gauss-Newton, the Marquardt-Levenberg, the Nelder-Mead and the steepest descent methods [4]. Okon *et al.* [5] added that SOLVER in Microsoft Excel, which is based on the robust and reliable generalized reduced gradient (GRG) method, can be used as an easy iteration protocol to perform the nonlinear iteration. Additionally, all the algorithms have similar properties; they require input initial parameters and use these values to get a better estimation of the parameters used in an iterative process. In fitting the model to experimental data, most of the protocols used the sum of square error to minimize the difference between the model-predicted and experimental data. Interestingly, there are other statistical error measurement tools like; average error, absolute error, mean square error (MSE), root-mean-square error (RMSE), Chi-square (X^2), among others. Therefore, this study looks at the effect (i.e., prediction) of using different statistical error functions on fitting experimental data to a model or curve.

2 Evaluation of Statistical Error Functions

2.1 Sample Model and Data Acquisition

The model used for this evaluation was developed by Carreau-Yasuda for the viscosity of a non-Newtonian fluid. This model as expanded in Equation 1 has five (5) adjustable parameters, namely, a, n, λ, μ_o and $\mu_{\inf}$. The experimental data to be fitted to this model (Equation 1) were obtained from the study of Morrison [6]; as presented in Table A.1 (Appendix).

$$\mu_{eff} = \mu_{\inf} + \left(\mu_o - \mu_{\inf}\right)\left[1 + \left(\dot{\gamma}\lambda\right)^a\right]^{\left(\frac{n-1}{a}\right)} \tag{1}$$

where μ_o is the viscosity at zero shear rate, $\mu_{\inf}$ is the viscosity at infinite shear rate, λ denotes the relaxation time, $\dot{\gamma}$ is the shear rate, μ_{eff} is the effective viscosity, a is exponent and n is the power index.

2.2 Model Fitting Procedures

The different statistical error functions used to fit the experimental data to the Carreau-Yasuda model (Equation 1) using SOLVER in Microsoft Excel are expressed in Equations 2 through 8.

i. **Average error:**

$$E_{avg} = \frac{1}{N}\sum_{i=1}^{N}\left[\frac{\mu_{\exp.} - \mu_{model}}{\mu_{\exp.}}\right] \tag{2}$$

ii. **Absolute error:**

$$E_{abs} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{\mu_{\exp.} - \mu_{model}}{\mu_{\exp.}}\right| \tag{3}$$

iii. **Sum of square error:**

$$SSE = \sum_{i=1}^{N}\left(\mu_{\exp.} - \mu_{model}\right)^2 \tag{4}$$

iv. **Mean square error**:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}\left(\mu_{\exp.} - \mu_{model}\right)^2 \tag{5}$$

v. **Root mean square error:**

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}\left(\mu_{\exp.} - \mu_{model}\right)^2}{N}} \tag{6}$$

vi. **Chi-square error:**

$$\chi^2 = \sum_{i=1}^{N}\left(\frac{\mu_{\exp.} - \mu_{model}}{\mu_{model}}\right)^2 \tag{7}$$

vii. **Normalized root mean square error:**

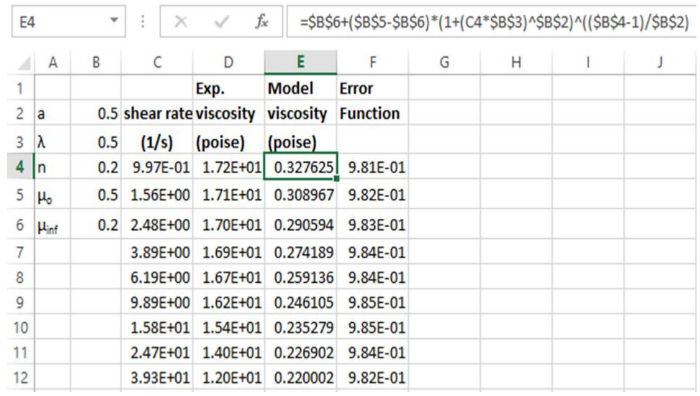
$$N_{RMSE} = \frac{RMSE}{\mu_{\exp.(\max)} - \mu_{\exp.(\min)}} \tag{8}$$

where;

$\mu_{\exp.}$ = experimental viscosity

μ_{model} = model predicted viscosity.

The extracted data from Morrison [6] were entered into the Excel spreadsheet in two columns; COLUMN C contained the shear rate while COLUMN D contained the viscosity data, as shown in Figure 1. The adjustable parameters: a, n, λ, μ_o and $\mu_{\inf}$ in the model (Equation 1) were assumed; as indicated in COLUMN B in Figure 1, to estimate the effective viscosity (i.e.,) in COLUMN E. Then, the difference between the experimental viscosity ($\mu_{\exp.}$) and model-predicted viscosity (μ_{model}) was evaluated in COLUMN F (Figure 1), using the error functions. Again, the summation of the error function was estimated in CELL F59, as visible in Figure 2. Afterwards, the SOLVER function was activated to minimize the summed error function value in CELL $F$59 (i.e., Set Objective) by changing the adjustable parameters in the model (Variable Cells) in COLUMN B (i.e., $B$2:$B$6); as shown in Figure 2. Afterwards, the solving method GRG was selected to handle the nonlinear minimization (iterative) protocol as the Solve icon was clicked to start the SOLVER solving process.



E4 =$B$6+($B$5-$B$6)*(1+(C4*$B$3)^$B$2)^(($B$4-1)/$B$2)

	A	B	C	D	E	F
1				Exp.	Model	Error
2	a	0.5	shear rate	viscosity	viscosity	Function
3	λ	0.5	(1/s)	(poise)	(poise)	
4	n	0.2	9.97E-01	1.72E+01	0.327625	9.81E-01
5	μo	0.5	1.56E+00	1.71E+01	0.308967	9.82E-01
6	μinf	0.2	2.48E+00	1.70E+01	0.290594	9.83E-01
7			3.89E+00	1.69E+01	0.274189	9.84E-01
8			6.19E+00	1.67E+01	0.259136	9.84E-01
9			9.89E+00	1.62E+01	0.246105	9.85E-01
10			1.58E+01	1.54E+01	0.235279	9.85E-01
11			2.47E+01	1.40E+01	0.226902	9.84E-01
12			3.93E+01	1.20E+01	0.220002	9.82E-01

Fig 1, Microsoft Excel screen shot of the model estimation

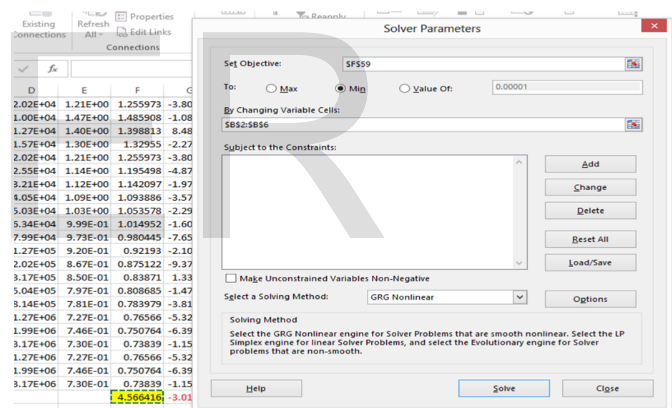


Fig 2, Screen shot of the SOLVER menu

The values obtained for the model's (Equation 1) adjustable parameters: a, n, λ, μ_o and $\mu_{\inf}$ based on the various error functions are presented in Table 1. Then, the predictions of the fitted model based on the various mentioned error functions were evaluated using the coefficient of determination (R^2) and F-test (F_{test}) values. According to Oforkansi and Oduola [7], coefficient of determination (R^2) is considered as one of the main criteria for selecting the best fit model(s). This is because R^2 depicts the degree of explained or accounted for, the variance between the experimental data and model-predicted results. In addition to the coefficient of determination, the acceptability of the goodness of the fitted model was based on the F-test value. That is, the model with the F-test value less than the F-critical value (Tables A.2 and A.3 in Appendix) indicates the best-fitted model. The equations for R^2 and F-test are expanded in Equations 9 and 10 respectively.

ISSN 2229-5518

$$R^2 = \left[1 - \frac{\sum_{i=1}^{N} \left(\mu_{\exp.} - \mu_{model} \right)^2}{\sum_{i=1}^{N} \left(\mu_{\exp.} - \overline{\mu_{model}} \right)^2} \right] \tag{9}$$

$$F_{test} = \frac{\sum_{i=1}^{N} \left(\mu_{\exp.} - \overline{\mu_{\exp.}} \right)^2}{\sum_{i=1}^{N} \left(\mu_{model} - \overline{\mu_{model}} \right)^2} \tag{10}$$

where;

$\overline{\mu_{\exp.}}$ = average experimental viscosity

$\overline{\mu_{model}}$ = average model predicted viscosity.

3 Findings from the Evaluation

Table 1 presents the values of the model's adjustable parameters obtained based on the different error functions. The minimized error function values that established the model's adjustable parameter are presented in Table 2. The results obtained indicated that different adjustable parameters' values for the model were established by some of the statistical error functions: average, absolute and Chi-square errors (Table 1). Furthermore, the results indicated that the sum of square, mean square, root mean square and normalize root mean square error functions established the same values for the model's adjustable parameters. The reason for this is attributed to the fact that this error functions simply evaluate the squared difference between the experimental data and model-predicted results. For the average error function, establishing the adjustable parameters of the model required using the "To Value Of" option on the SOLVER. This applied function on the SOLVER handles the challenges of early stopping (converging) of the GRG iteration protocol to drive the "Objective Cell" (i.e., the average error) to the targeted value. Nevertheless, in some cases, the targeted value may not be achieved because when the GRG algorithm establishes the local optimal for the function it truncates the iterative protocol without getting to the targeted value. Based on the established model's adjustable parameters from the various error functions, the predictions of the fitted model were compared with the experimental data (Figure 3). In Figure 3, the results obtained indicated that the fitted model predictions were close to the experimental data, except for the fitted model's prediction from the average error function. The exceptional prediction from the model fitted using the average error function approach is because some evaluated difference between the experimental data and model prediction resulted in negative values. This happened to those model predicted values that are greater than the experimental data. Consequently, these negative values affected the summed difference between experimental data and model-predicted results.

- *Anietie N. Okon, PhD, is a Senior Lecturer in Petroleum Engineering programme, Department of Chemical and Petroleum Enginerring, University of Uyo, Nigeria. E-mail: anietieokon@uniuyo.edu.ng*
- *Dr. Uwem E. Inyamg and Dr. Innocent O. Oboh are Senior Lecturer and Associate Professor, respectively, in Chemical Engineering programme, Department of Chemical and Petroleum Engineering, University of Uyo, Nigerian.*

Furthermore, the validity of the various error functions used to establish the adjustable parameters in the model was evaluated using the coefficient of determination (R^2) to determine the closeness (goodness of fit) of the fitted model predictions to the experimental data. Also, the acceptability of the fitted model was evaluated using the F-test value (F_{test}); as presented in Table 2. The results obtained indicated that almost all the error functions have the same R^2 of 0.9998 except for an average error with R^2 of 0.9585. The R^2 obtained for average error indicated the reason why the fitted model's predictions based on average error were not close to (i.e., align with) the experimental data as the fitted model based on other error functions did; as depicted in Figure 3. Despite the closeness of the fitted models' prediction as established by the R^2 obtained, the Ftest value obtained showed that the fitted model based on the sum of square, mean square, root-mean-square and normalized root-mean-square error functions were more acceptable than the fitted model based on absolute error and Chi-square error functions. This assertion is based on the Ftest values obtained, that is, 1.0002, 1.0037 and 1.0004 for the sum of square error, absolute error and Chi-square error respectively. The acceptability of the fitted model based on the sum of square errors was further supported by the fact that, the Ftest value, 1.0002, obtained was less than their F-critical value of 1.5709 (Tables A.2 and A.3 in Appendix). This implied that the fitted model based on average error is rejected because its Ftest value, 0.9756 was greater than its F-critical value of 0.6365.

Table 1: Values for the model adjustable parameters

Error Functions	Model Adjustable Parameters				
	a	n	λ	μ_o	$\mu_{\inf}$
Average error	2.2122	0.5835	0.1949	19.9711	0.7246
Absolute error	2.3194	0.5179	0.0525	17.0109	0.6804
Sum of square error	2.2173	0.5155	0.0512	17.0514	0.6895
Mean square error	2.2172	0.5155	0.0512	17.0514	0.6895
Root mean square error	2.2173	0.5155	0.0512	17.0514	0.6895
Chi-square error	1.9153	0.5200	0.0520	17.2154	0.6690
Normalize RMSE	2.2172	0.5155	0.0512	17.0514	0.6895

Table 2: Minimised error functions value and their respective validity evaluation results

Error Functions	Error Function Values	Coefficient of Determination (R^2)	F-test Value (F_{test})
Average error	-5.25E-08	0.9585	0.9756
Absolute error	0.01567	0.9997	1.0037
Sum of square error	0.27826	0.9998	1.0002
Mean square error	0.00506	0.9998	1.0002
Root mean square error	0.07113	0.9998	1.0002
Chi square error	0.03258	0.9997	1.0004
Normalize RMSE	0.00432	0.9998	1.0002

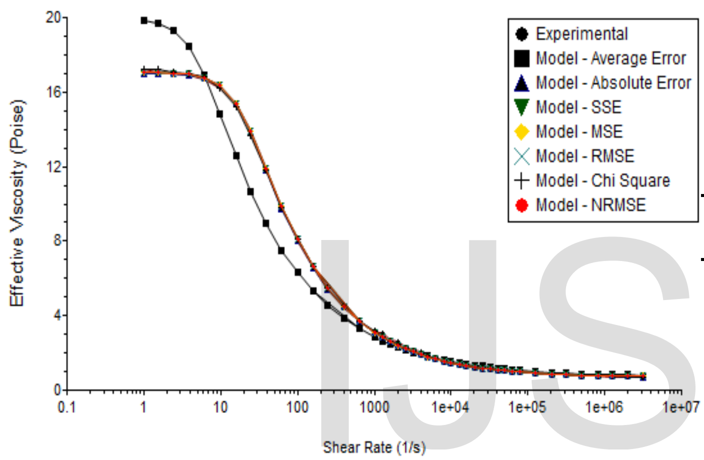


Fig 3, Comparing various error functions model estimations

4 Conclusion

Based on the evaluation results obtained from the SOLVER, the following conclusions were drawn:

i. nonlinear models fitted with the average error function would be less fitted with the experimental data when compared with models fitted using other error functions;
ii. fitting nonlinear models to experimental data using the average error function would require setting the error minimization to the least value possible;
iii. coefficient of determination (R^2) and F-test (F_{test}) establish the fitted model goodness fit and acceptability than other statistical error functions; and
iv. the sum of square, mean square, root mean square and normalize root mean square error functions are good error minimization approach to fit nonlinear models and curves to experimental data.

References

[1] S.J. Ahn, "Geometric Fitting of Parametric Curve and Surfaces". Journal of Information Processing Systems, Vol. 4, No. 4, pp. 153-158, 2008.
[2] https://en.wikipedia.org/wiki/Curve_fitting. Accessed 17th November, 2018.
[3] M.A. Hossain, H.H. Ngo, W. Guo, "Introductory of Microsoft Excel SOLVER Function – Spreadsheet Method for Isotherm and Kinetics Modelling of Metals Biosorption in Water and Wastewater" Journal of Water Sustainability, Vol. 3, No. 3, pp. 223-237, 2013.
[4] M.L. Johnson, "Why, When, and How Biochemists should use Least Squares" Journal of Analytical Biochemistry, Vol. 206, No. 2, pp. 215-225, 1992.
[5] A.N. Okon, F.D Udoh, D. Appah, "Empirical Wellhead Pressure – Production Rate Correlations for Niger Delta Oil Wells". Paper presented at the Society of Petroleum Engineers Nigeria Annual International Conference and Exhibition, Lagos, Nigeria, 4-6 August, 2015.
[6] F.A. Morrison, "Using the SOLVER Add-in in Microsoft Excel" Michigan Technology University, USA, 2005.
[7] B.C. Oforkansi, M.K. Oduola, "Mathematical Model of Thin-Layer Drying Process in a Plantain Sample". International Journal of Engineering Research, Vol. 5, No. 5, pp. 364-367, 2016.

Appendix

Table A.1: Experiment viscosity- shear rate data from Morrison [6]

shear rate (1/s)	viscosity (poise)	shear rate (1/s)	viscosity (poise)	shear rate (1/s)	viscosity (poise)
9.97E-01	1.72E+01	6.40E+02	3.68E+00	1.57E+04	1.30E+00
1.56E+00	1.71E+01	1.01E+03	3.12E+00	2.02E+04	1.21E+00
2.48E+00	1.70E+01	1.27E+03	3.06E+00	2.55E+04	1.14E+00
3.89E+00	1.69E+01	1.61E+03	2.67E+00	4.05E+04	1.09E+00
6.19E+00	1.67E+01	2.03E+03	2.64E+00	5.03E+04	1.03E+00
9.89E+00	1.62E+01	2.56E+03	2.28E+00	6.34E+04	9.99E-01
1.58E+01	1.54E+01	3.23E+03	2.15E+00	7.99E+04	9.73E-01
2.47E+01	1.40E+01	4.01E+03	2.05E+00	1.27E+05	9.20E-01
3.93E+01	1.20E+01	4.02E+03	1.94E+00	2.02E+05	8.67E-01
6.26E+01	9.86E+00	4.99E+03	1.88E+00	3.17E+05	8.50E-01
9.96E+01	7.98E+00	6.30E+03	1.67E+00	5.04E+05	7.97E-01
1.58E+02	6.54E+00	8.08E+03	1.60E+00	8.14E+05	7.81E-01
2.49E+02	5.40E+00	1.00E+04	1.47E+00	1.27E+06	7.27E-01
4.00E+02	4.39E+00	1.27E+04	1.40E+00	1.99E+06	7.46E-01
6.40E+02	3.68E+00	1.57E+04	1.30E+00	3.17E+06	7.30E-01
1.58E+02	6.54E+00	2.02E+04	1.21E+00	1.27E+06	7.27E-01
2.49E+02	5.40E+00	1.00E+04	1.47E+00	1.99E+06	7.46E-01
4.00E+02	4.39E+00	1.27E+04	1.40E+00	3.17E+06	7.30E-01

Table A.2: Experimental and model predicted F-Test Analysis

	Average Error		*Absolute Error*		*Sum of Square Error (SSE)*		*Mean Square Error (MSE)*	
	Experimental	*Predicted*	*Experimental*	*Predicted*	*Experimental*	*Predicted*	*Experimental*	*Predicted*
Mean	4.5664	4.3784	4.5664	4.5464	4.5664	4.5664	4.5664	4.5664
Variance	29.4970	30.2334	29.4970	29.3891	29.4970	29.4919	29.4970	29.4918
Observations	55	55	55	55	55	55	55	55
df	54	54	54	54	54	54	54	54
F	0.9756		1.0037		1.0002		1.0002	
P(F<=f) one-tail	0.4640		0.4947		0.4997		0.4997	
F Critical one-tail	0.6365		1.5709		1.5709		1.5709	

Table A.3: Experimental and model predicted F-Test Analysis continue

	Root Mean Square Error (RMSE)		*Chi-Square (X^2)*		*Normalize Root Mean Square Error (NRMSE)*	
	Experimental	*Predicted*	*Experimental*	*Predicted*	*Experimental*	*Predicted*
Mean	4.5664	4.5664	4.5664	4.5693	4.5664	4.5664
Variance	29.4970	29.4918	29.4970	29.4839	29.4970	29.4918
Observations	55	55	55	55	55	55
df	54	54	54	54	54	54
F	1.0002		1.0004		1.0002	
P(F<=f) one-tail	0.4997		0.4994		0.4997	
F Critical one-tail	1.5709		1.5709		1.5709	