Proceedings of the *International Conference on Theory and Application of Mathematics and Informatics ICTAMI 2005* - Alba Iulia, Romania

AN ITERATIVE NONLINEAR REGRESSION FOR POLARISATION/DEPOLARISATION CURRENTS

Emiliana Ursianu, Radu Ursianu and Victor Ursianu

Abstract. To define the relationship between the response – a random or dependent variable y and a predetermined or independent variable x, a straight line or concave respectively convex functions affected by random errors must be used.

A locally weighted low-degree polynomial regression i.e. linear is fit to several subsets of data using linear least square fitting and a first-degree polynomial or the exponential curve.

PDC RVM Nova software provides a toolbox to load the dataset, to filter the data and to estimate the parameters of each subset.

1. Introduction

The problem of fitting experimental data by a linear combination of exponential is very common in technical and biological researches.

The corresponding regression model reflecting the effect of random errors on this fitting is non-linear with respect to its parameters:

$$y_i = \alpha + \sum_{j=1}^{k} \beta_j \rho_j^i + e_i, i = 1, ..., n \tag{1}$$

where k is the number of exponentials; n is the number of observations; here $\rho_j = \exp(-\lambda_i)$, α and β_j are constants and ei are random errors with zero mean and variance ρ^2.

The itterative methods of curve-fitting for data to n sum of exponentials require initial estimates of the parameters involving linearized by expanding in a Taylor series.

2. Model assumptions

The model (1) may be written more generally as:

$$\mathbf{y} = \alpha + \beta g(\mathbf{x}) + \mathbf{e} \tag{2}$$

i.e. $y_j = \alpha + \beta g(x_j) + e_j$. where $g(x)$ represents either a concave or convex continous function.

The null hypothesis may be:

$$H_0: \qquad g(x) = x \tag{3}$$

under various assumption concerning the probability distribution of the e's. The random errors ej are independently and symmetrically, but not necessarly identically, distributed around 0 with continous distribution function:

$$F_j(z) = P\{e_j < z\}$$
$$F_j(0) = \frac{1}{2}, j = 1, ..., n$$

For example, in normal regression analysis, the problem is to test for linearity against specific alternatives (i.e. concave or convex) if the errors are i.i.d. (independently and identically distributed) with unknown variance ρ^2 or if they have a covariance structure known except for a scale factor.

However with a sufficient number or replicated points, an independent estimate of or of the scale factor is available and then it is possible to test for linearity against the mentioned alternatives.

3. Test statistics

Let the pairs $\{y_j, x_j\}, j = 1, ..., n$ be ordered so that $x_1 < x_2 < ... < x_n$ and define three disjoint intervals T_1, T_2, and T_3 on the x-axis such that containing the first n_1 observed the x's lie in T_1 , the next n_2 ordered the x's lie in T_2 and the remaining $n_3 = n - n_1 - n_2$, ordered the x's lie in T_3 .

Let us define the statistics:

$$h_{ijk} = h(e_i, e_j, e_k) = 1, if (y_j - y_i)/(x_j - x_i) > (y_k - y_i)/(x_k - x_i)$$
$$h_{ijk} = h(e_i, e_j, e_k) = -1, if (y_j - y_i)/(x_j - x_i) < (y_k - y_i)/(x_k - x_i) \tag{4}$$

$$h_{ijk} = h(e_i, e_j, e_k) = 0, \; otherwise$$

The proposed statistics are:

$$H_n = (n_1 n_2 n_3)^{-1} \sum h_{ijk} \tag{5}$$

all $n_1 n_2 n_3$ sets of pairs.

The decision: the regression is concave if $H_n > 0$ respectively convex if $H_n < 0$.

The estimators r_i^* of ρ_i are the roots of the kth degree polynomial:

$$r^{*k} - \frac{|A_1^*|}{|A^*|} r^{*k} - ... - \frac{|A_k^*|}{|A^*|} = 0 \tag{6}$$

with coefficients the difference of successive pairs of the experimental data y_i , $i = 1, \ldots, n$:

$$Y_h' = Y_{2h} - Y_{2h-1}, h = 1, 2, ..., 2k \tag{7}$$

where

$$Y_{2h} = \sum_{i=m+1+h-k}^{n+h-2k} y_i ; Y_{2h-1} = \sum_{i=h}^{m+h-k-1} y_i, \qquad n = 2m \tag{8}$$

respectively,

$$Y_{2h} = \sum_{i=m+1+h-k}^{n+h-2k} y_i ; Y_{2h-1} = \sum_{i=h}^{m+h-k} y_i, \qquad n = 2m+1 \tag{9}$$

Remark. For the small size of experimental data or the value of σ^2 too large, the lineary interpolated values, is introduced between each successive pair points: (t_i^*, y_i^*),with $t_i^* = x_i^*$, where $y_i^* = \frac{y_i + y_{i+1}}{2}$ respectively $t_i^* = \frac{t_i + t_{i+1}}{2}$

4. Iterative method

Iterative method using initial estimates involving l.s. process after linearization by expanding in a Taylor series. ([5]).

A method for obtaining initial estimates of the parameters in (2) is applied using a least square (l.s.) "peeling-off" technique. ([3])

By the l.s. method the estimates of a_k and b_k can be obtained by fitting a straight line to the last three data points. Then :

$$y_{residual} = y - a_k e^{-b_k t} \tag{10}$$

By taking the next three points for $y_{residual}$ against $t(= x)$ and fitting a straight line to them by l.s. method , the new $y_{residual}$ is obtained corresponding to the determined a_{k-1} and b_{k-1}, etc.

New

$$y_{residual} = y_{residual} - a_{k-1} e^{-b_{k-1} t} \tag{11}$$

The remaining data points are used to calculate a_1 and b_1.

The ordinary l.s (OLS) estimator respectively the weighted l.s. (WLS) estimator of β – the vector of regression parameters are given in the literature: ([6], [7]).

$$\hat{\beta}_{OLS} = (\mathbf{X'X})^{-1}\mathbf{X'y} \tag{12}$$

respectively

$$\hat{\beta}_{WLS} = (\mathbf{X'WX})^{-1}\mathbf{X'Wy} \tag{13}$$

with $\mathbf{W}$ = block diag $(w_1 I_{n_1}, ..., w_1 I_{n_k})$

$$w_i = u_i^{-1} = (\frac{1}{n_i}\sum_{j=1}^{n_i} \hat{e}_{ij}^2)^{-1} \tag{14}$$

the weights

I_t is the $t \times t$ identity matrix.

For nonlinear least squares fitting to a number of unknown parameters, linear least squares fitting may be applied iteratively to a linearized form of the function until convergence is achieved. However, it is often also possible to linearize a nonlinear function at the outset and still use linear methods for determining fit parameters without resorting to iterative procedures. This approach does commonly violate the implicit assumption that the distribution of errors is normal, but often still gives acceptable results using normal equations, a pseudoinverse, etc. Depending on the type of fit and initial parameters chosen, the nonlinear fit may have good or poor convergence properties. If uncertainties (in the most general case, error ellipses) are given for the points, points can be weighted differently in order to give the high-quality points more weight.

Vertical least squares fitting proceeds by finding the sum of the squares of the vertical deviations R^2 of a set of n data points

$$R^2 \equiv \sum [y_i - f(x_i, a_1, ..., a_n)]^2 \tag{15}$$

from a function f. Note that this procedure does not minimize the actual deviations from the line (which would be measured perpendicular to the given function). In addition, although the unsquared sum of distances might seem a more appropriate quantity to minimize, use of the absolute value results in discontinuous derivatives which cannot be treated analytically. The square deviations from each point are therefore summed, and the resulting residual is then minimized to find the best fit line. This procedure results in outlying points being given disproportionately large weighting.

The condition for R^2 to be a minimum is that

$$\frac{\partial(R^2)}{\partial a_i} = 0 \tag{16}$$

for $i = 1, ..., n$.

5. Software description

Transformer is one of the most important and costly apparatus in power system. The reliable and efficient fault-free operation of the high-voltage transformer has a decisive role in the availability of electricity supply. The transformer oil/paper insulation gets degraded under a combination of thermal, electrical, chemical, mechanical, and environmental stresses during its operation. In recent years, there has been growing interest in the condition assesment of transformer insulation.

To estimate the status of transformer insulation it has been elaborated a modern diagnose method based on measuring and processing of polarisation/depolarisation currents which characterise the transformer.

The polarisation/depolarisation currents is affected by noise so it must be processed. After this, it must be transformed into a computed function of this type : $Ipol(t) = I_0 + \sum_{i=1}^{5} Ip_i * e^{\frac{-t}{Tp_i}}$.

The processing of the currents it is made in 2 stages. First the signal is filtered, eliminating the noise partially or even totally and then to the next stage, where the signal is parameterized.

The 4 methods of signal filtering are: Moving Average, Lowess, Loess and Savitzky-Golay. These methods are also used in Curve Fitting Tool software included in MATLAB suite.

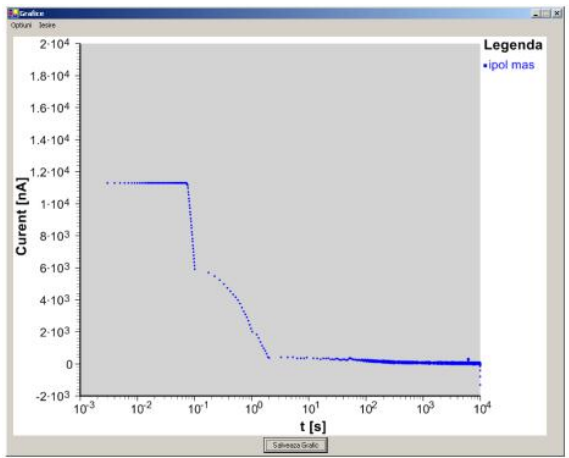

Figure 1: Moving average with span=51

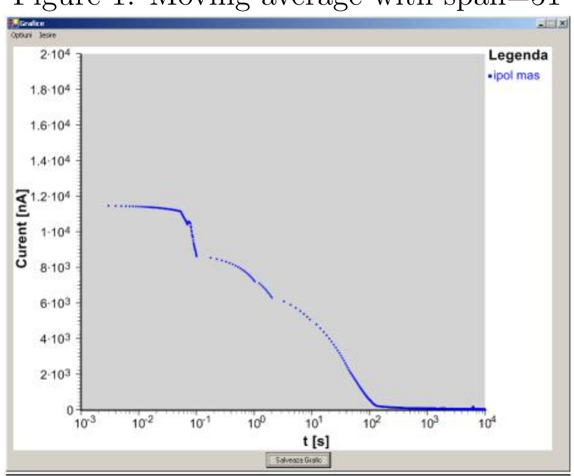

Figure 2: Lowess robust with span = 100

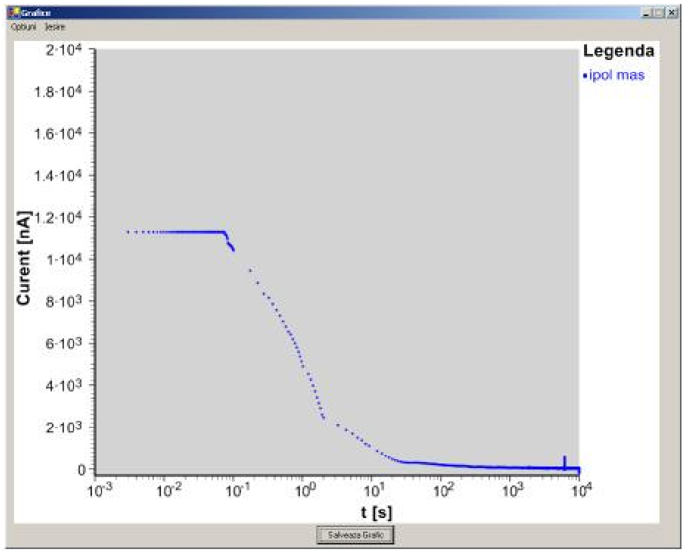

Figure 3: Loess robust with span = 50

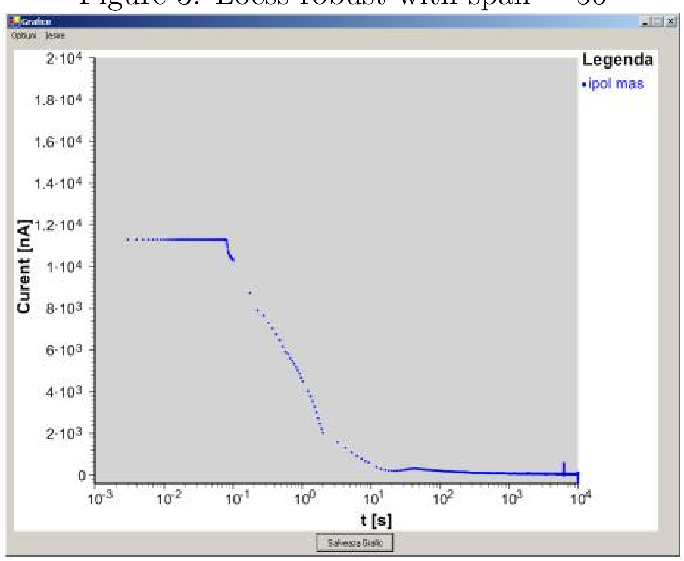

Figure 4: Savitzky-Golay span=51, degree=3, weighted and robust

After the signal is filtered, it must be parameterized by transforming it into a computed function using a mathematical procedure for finding the best-fitting curve to a given set of points by minimizing the sum of the squares of the offsets ("the residuals") of the points from the curve.

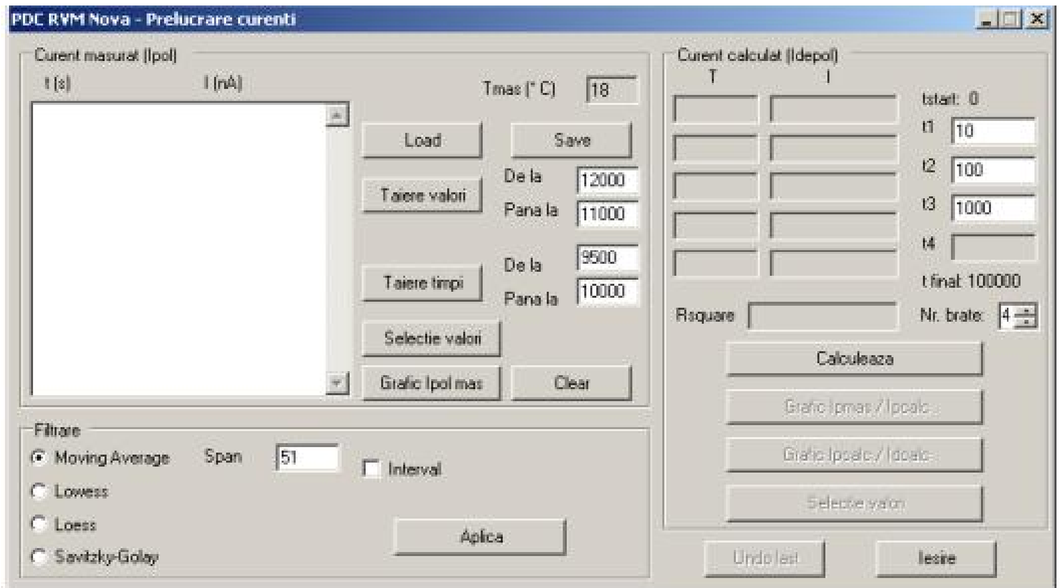

Figure 5: Processing window

6. Experimental data

Measured polarisation current:

Polarisation current

T [s]	Ipol [A]
1,224	1413,258*10^{-09}
1,324	1408,688*10^{-09}
1,423	1404,119*10^{-09}
1,523	1399,549*10^{-09}
1,623	1394,979*10^{-09}
1,723	1390,409*10^{-09}
1,823	1385,839*10^{-09}
1,923	1381,269*10^{-09}
2,022	1376,700*10^{-09}
3,270	1319,577*10^{-09}
4,268	1273,878*10^{-09}
5,267	1228,180*10^{-09}
6,265	1182,482*10^{-09}
7,263	1136,783*10^{-09}
8,261	1091,085*10^{-09}
9,260	1045,387*10^{-09}
10,258	999,688*10^{-09}
11,256	954,411*10^{-09}
12,254	911,769*10^{-09}
13,253	873,510*10^{-09}
14,251	839,703*10^{-09}
15,249	809,928*10^{-09}
16,247	783,597*10^{-09}
17,246	760,007*10^{-09}
18,244	738,376*10^{-09}
19,242	718,296*10^{-09}
20,240	699,552*10^{-09}
22,237	665,522*10^{-09}
24,233	634,301*10^{-09}
26,230	604,354*10^{-09}
28,226	576,383*10^{-09}
30,223	550,447*10^{-09}
33,217	514,321*10^{-09}
36,212	476,633*10^{-09}
39,207	435,187*10^{-09}
42,201	399,729*10^{-09}
46,194	358,577*10^{-09}
50,187	320,050*10^{-09}
55,178	299,794*10^{-09}
61,168	270,940*10^{-09}
65,161	253,329*10^{-09}
71,150	230,056*10^{-09}
75,143	215,811*10^{-09}
81,133	196,829*10^{-09}
85,126	186,289*10^{-09}
91,115	171,613*10^{-09}
95,108	161,278*10^{-09}
101,097	149,926*10^{-09}
111,080	140,172*10^{-09}
121,062	135,233*10^{-09}
131,044	133,004*10^{-09}
141,027	131,709*10^{-09}
151,009	129,132*10^{-09}
160,992	125,897*10^{-09}
170,974	123,199*10^{-09}
180,956	120,886*10^{-09}
190,939	119,028*10^{-09}
200,921	117,059*10^{-09}
220,886	112,838*10^{-09}
240,850	109,047*10^{-09}
260,815	105,320*10^{-09}
280,780	101,814*10^{-09}
300,745	98,232*10^{-09}
330,692	92,884*10^{-09}
360,639	87,880*10^{-09}
390,586	83,655*10^{-09}
420,533	79,660*10^{-09}
460,462	74,586*10^{-09}
500,392	70,223*10^{-09}
550,304	65,856*10^{-09}
600,215	63,001*10^{-09}
650,127	61,353*10^{-09}
700,039	60,153*10^{-09}
751,947	58,891*10^{-09}
801,859	57,518*10^{-09}
851,771	55,923*10^{-09}
901,683	54,560*10^{-09}
951,594	53,354*10^{-09}
1001,506	52,198*10^{-09}
1101,330	49,951*10^{-09}
1201,153	48,331*10^{-09}
1300,977	47,196*10^{-09}
1400,801	46,301*10^{-09}
1500,624	45,599*10^{-09}
1600,448	44,760*10^{-09}
1700,271	44,142*10^{-09}
1800,095	43,663*10^{-09}
1901,915	43,294*10^{-09}
2001,739	43,019*10^{-09}
2201,386	42,404*10^{-09}
2401,033	41,965*10^{-09}
2600,680	41,583*10^{-09}
2800,327	41,157*10^{-09}
3001,971	40,616*10^{-09}
3301,442	39,586*10^{-09}
3600,91	38,368*10^{-09}
3900,38	37,076*10^{-09}
4201,85	35,792*10^{-09}
4601,15	34,343*10^{-09}
5000,44	33,495*10^{-09}
5501,55	33,132*10^{-09}
6000,67	32,891*10^{-09}
6501,79	32,649*10^{-09}
7000,9	32,407*10^{-09}
7500,02	32,166*10^{-09}
8001,14	31,924*10^{-09}
8500,26	31,683*10^{-09}
9001,37	31,441*10^{-09}
9500,49	31,200*10^{-09}

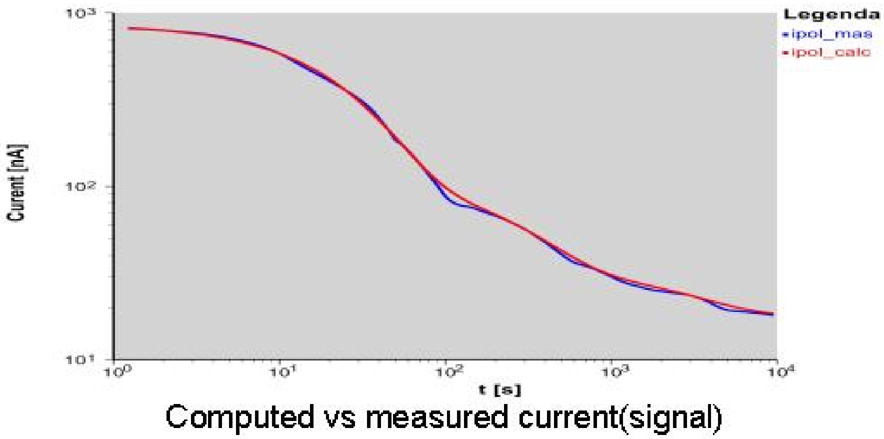

Figure 6: Computed vs measured current(signal)

The parameters of computed function $Ipol(t) = I_0 + \sum_{i=1}^{5} Ip_i * e^{\frac{-t}{Tp_i}}$ where I_0 is the stable value of measured current.

Ip_i [A]	Tpi [s]
77,3*10^{-09}	7,19
677,5*10^{-09}	25,4
76,3*10^{-09}	264
15,7*10^{-09}	2875

are given in ([10]).

References

[1] Agha,M.[1971].*A direct method for fitting linear combinations of exponentials*.Biometrics 27(2),399-413

[2] Della Corte,M.,Buricchi,L.,and Romano,S.[1974].*On the fitting of linear combinations of exponentials.*Biometrics 30,367-369

[3] Foss,S.D.[1969].*A method for obtaining initial estimates of the parameters in exponential curve fitting.*Biometrics 25(3),580-584

[4] Iosifescu,M.et al.[1985].*Mica enciclopedie de statistica.*Ed.st.encicl.,Bucuresti

[5] Postelnicu,T.,Ursianu,Emiliana.[1980].*Application of nonlinear estimation in the exploration of dose-response relationship.*Biom.J.,22(5),425-431

[6] Scheffe',H.[1959].*The Analysis of Variance.*J.Wiley,New Jork

[7] Shao,Jun;Rao,J.N.K.[1993].*Jacknife inferencefor heteroscedastic linear regression models.*Canadian J.of Statistics.21(4),377-395

[8] Thornby,J.I.[1972].*A robust test for linear regression.*Biometrics 28(2),533-543

[9] Vaduva,I.[1970].*Analiza dispersionala.*Ed.tehnica,Bucuresti

[10] Ursianu,V.[2005]. *A modern diagnosising method based on measuring and processing of polarization/depolarization currents.*UPB-Nova Industrial S.A.

Emiliana Ursianu – ISMMA - Romanian Academy
Radu Ursianu – UPB - Politehnical University of Bucharest
Victor Ursianu – Nova Industrial SA