JOURNAL OF THEORETICAL AND APPLIED MECHANICS
50, 4, pp. 923-931, Warsaw 2012
50th Anniversary of JTAM

MODIFICATION OF THE LSM CRITERIA TO APPROXIMATE TEST DATA FOR FATIGUE CRACK GROWTH RATE

Sylwester Kłysz
Air Force Institute of Technology, Warsaw, Poland and
University of Warmia and Mazury, Olsztyn, Poland; e-mail: sylwester.klysz@itwl.pl

Janusz Lisiecki, Andrzej Leski
Air Force Institute of Technology, Warsaw, Poland

Tomasz Bąkowski
OLT Express Regional, Warsaw, Poland

Modification of the least squares method criteria is presented. It is intended to enhance the approximation of test data for fatigue crack growth rate. In particular, changes are introduced to improve approximation in the cases when the range of values for at least one of the variables is within the range of several orders of magnitude or differs from the range of values for other variables. Differences between particular criteria and their influence on the approximation of test data is illustrated further on. Calculations are conducted for aluminum alloy 2024 taken from Mi-2 helicopter rotor blades.

Key words: fatigue crack propagation, NASGRO equation, method of least squares (LSM)

1. Introduction

Over the years, a number of relationships have been developed to represent all or parts of a typical range of fatigue crack growth data, $da/dN = f(\Delta K)$. Application of the NASGRO equation derived by Forman and Newman from NASA, de Koning from NLR and Henriksen from ESA, of the general form ([1]; [8])

$$\frac{da}{dN} = C\Big(\frac{1-f}{1-R}\Delta K\Big)^n \frac{\Big(1-\frac{\Delta K_{th}}{\Delta K}\Big)^p}{\Big(1-\frac{K_{max}}{K_c}\Big)^q} \tag{1.1}$$

has significantly extended possibilities of describing the crack growth rate, tested according to the standard [3]. The NASGRO equation represents the most comprehensive formulation of the crack growth law. It is a full-range model that mathematically represents all three regions of the da/dN vs. ΔK curve comprises the mean stress effect, threshold, fast fracture and crack closure (Hudson and Seward, 1983; [9]; Walker, 1970). The coefficients stand for ([2]; Fuchs and Stephens, 1980; [8]): a is the crack length [mm], N – number of load cycles, R – stress ratio, ΔK – stress intensity factor (SIF) range depending on the specimen size, applied loads, crack length, $\Delta K = K_{max} - K_{min}$ [MPa$\sqrt{\text{m}}$], ΔK_{th} – SIF threshold, minimum value of ΔK, at which the crack starts to propagate: K_c – critical value of SIF, f – Newman's function describing closing of the crack (Newman, 1984), C, n, p, q – empirical coefficients.

Determination of the above coefficients for the equation correctly approximating tests data is difficult and causes some singularities when the Least Squares Method (LSM) is used. Curves fitting to the NASGRO equation are obtained using the NASMAT module contained within the NASGRO suite of software [8]. The NASMAT curve fitting algorithms use the least-squares error minimization routines in the log-log domain to obtain the corresponding constants. The constants C and n, i.e. the main fit parameters, are determined through the minimization of the sum of squares of errors, where the error term corresponding to the i-th data pair $(\Delta K, da/dN)_i$ is (Forman *et al.*, 2005)

$$e_i = \log\Big(\frac{da}{dN}\Big)_i - \log C - n\log\Big(\frac{1-f}{1-R}\Delta K\Big)_i - \log\left[\frac{\Big(1-\frac{\Delta K_{th}}{\Delta K}\Big)^p}{\Big(1-\frac{K_{max}}{K_c}\Big)^q}\right]_i \tag{1.2}$$

Values of da/dN are determined using the method of differentiating the dependence $a-N$ with the secant or the polynomial method applied ([1]; [3]; [8]).

The curve fitting of crack growth data is an iterate process that consists in using established values of various constants (other than C and n), specifying the data sets that typify the material, applying the least-squares algorithm to compute C and n, and plotting the data for various R values with the curve fit of each stress ratio. The process is continued by making slight modifications in the entered values until the best fit to the test data is obtained. In general, fitting the NASGRO equation is really a multi-step process involving:

- fitting or defining the threshold region
- fitting or defining the critical stress intensity or toughness to be used at the instability asymptote
- making initial assumptions on key parameters such as p and q
- performing the least squares fit to obtain C and n, and finally
- using engineering judgment to adjust the results for consistency and/or a desired level of conservatism.

For the LSM approximation of test data, the analytical description thereof and determination of coefficients of approximation equations, according to which the criterion used in the analysis, is the minimum of the square sum

$$S = \sum_{i=1}^{n}(\overline{y}_i - y_i)^2 \tag{1.3}$$

of deviations between values of the test data y_i and those of the approximated function $\overline{y}_i$. This method of approximation is characterized with the following properties that in some cases may be considered as disadvantages (Forman *et al.*, 2005; Huang *et al.*, 2005; Taheri *et al.*, 2003; White *et al.*, 2005; Zhao and Jiang, 2008):

- value of the sum S increases regarding the order of magnitude as approximated values increase, e.g. if values of the test data are of the order 10, 10^3, 10^6, with the scatter of 10%, the summed differences are of the order 1, 10^2, 10^5 respectively, and hence:
  - the same, e.g. the (2-5)-times change in the test value results in different as to the order of values of the summed differences
  - dynamic changes in the total value of the sum S depend on values of the differences – as a quadratic function it is characterized by a linear function of the derivative, which also means that for differences close to zero (e.g. 10^{-5}, 10^{-8}, etc.) this dynamic change is much smaller than for differences of higher magnitudes, which influences the "flexibility" of the performed approximation
  - if the test data differ significantly in magnitude from each other (e.g. from 1 to 10^5 or from 10^{-8} to 10^{-2}), the approximated values near the lower threshold contribute much less to the total sum S than the approximated values near the upper threshold; this means that, e.g. tens or hundreds of test data with differences in magnitude of 100% from value 1 are less significant in the performing approximation than one or a few data points which differ by 1% from value 10^5.

According to the above stated example, the approximation is "asymmetric" since a better approximation will be achieved for higher values of test data, neglecting differences around smaller values – an example of such an approximation is shown in Fig. 1, where one can see a good fit of theoretical description of 9 curves for large values of da/dN (over 10^{-4} mm/cycle), while there is a visible mismatch-fit for the smallest values (below 10^{-5} mm/cycle). The presented approximation has been achieved by satisfying the LSM criterion, i.e. the minimum value of the sum S.

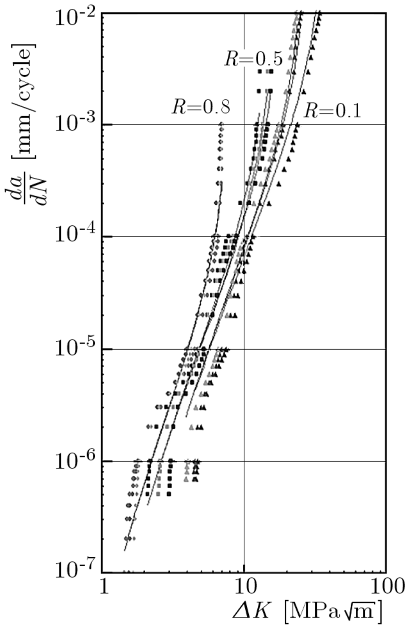

Fig. 1. Example of approximation with the NASGRO equation, LSM according to formula (1.3); test data with large y value scatter

When the test data are within a wide range of values, e.g. 5 orders of magnitude, i.e. from 10^{-2} to 10^{-7} mm/cycle, then differences between the highest values and the approximating function will have the largest influence on the square sum S of deviations while differences for small values, sometimes of 2-3 orders of magnitude, do not contribute much to the total sum S.

The LSM can be flexible when the below proposed modification is introduced, and this is the main aim of this work.

2. Modifications of the LSM method criterion

In order to eliminate the approximation mismatch-fit as shown in Fig. 1 and to improve the quality of approximation, a modification of formula (1.3) is proposed. It should reach the following form

$$S^* = \sum_{i=1}^{n} \Big(\frac{\overline{y}_i - y_i}{y_i}\Big)^2 = \sum_{i=1}^{n} \Big(\frac{\overline{y}_i}{y_i} - 1\Big)^2 \tag{2.1}$$

In this way, the fraction in brackets, as a relative error, is a stable measure of deviation between the approximated and approximating values, i.e. independent of the magnitude of compared values, since:

- each value among the test data y_i has an equal contribution in the sum S^*, independent of its magnitude 10^{-7}, 10^{-2}, 1 or 10^5 (i.e. it fits in any magnitude range) – always a deviation of e.g. 10-, 50-, 200-percent of the approximating value will give a component of the sum S^* equal to 0.01, 0.25, 4, respectively

- the criterion assures that the achieved approximation is "symmetric", i.e. the degree of approximation around lower and higher values is the same
- disadvantages of the criterion described with formula (1.3) are no longer valid.

Figure 2 shows results of approximation achieved with modified LSM criterion (2.1). A visible improvement in the approximation of the 9 curves can be seen within the whole range of da/dN values.

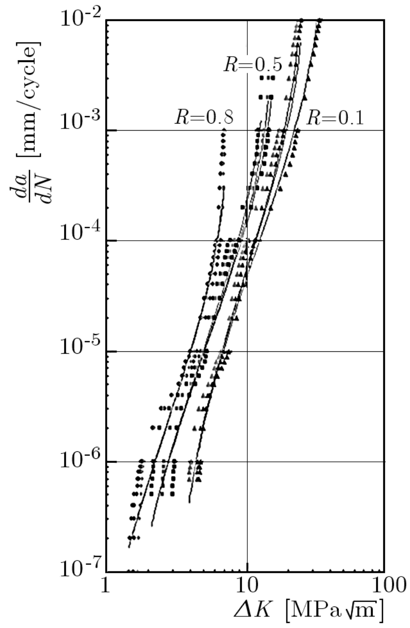

Fig. 2. Approximation of $da/dN = f(\Delta K)$ data with the NASGRO equation, the LSM modified according to formula (2.1)

The criterion described with formula (2.1) has also some specific property: if the approximating value equals zero (i.e. for the approximation smaller by 100%) or it is twice as much as the approximated value (i.e. for approximation larger by 100%), then independently of the approximated value, the component of the sum S^* will be equal to 1.

Both criteria (1.3) and (2.1) have also some disadvantage consisting in that if the approximating value $\overline{y}_i$ is much smaller than the approximated value y_i (i.e. by 3, 5, 7 orders of magnitude) or simply close to zero, then the component of the sum S and S^* is close to the square of value y_i (in the case of (1.3)) or to 1 (in the case of (2.1), independently of how these two values differ from each other.

Obviously, it is important whether the approximation and behavior of the approximating curve near the value y_i at the level of e.g. 10^{-6} and lower (i.e. for strongly decreasing values within the "threshold" range of the graph) take place at the level of 10^{-8}, 10^{-12} or 10^{-20} (what is not hard to achieve for curves showing strong vertical courses on graphs plotted with the logarithmic scale applied); it is much better when the possible difference between values $\overline{y}_i$ and y_i is not too large.

Due to dynamic changes around the value $\overline{y}_i$ equal to zero (completely monotonic as for the second-degree polynomial), functions (1.3) and (2.1) are practically insensitive to the approximated values 10^{-2}, 10^{-5}, 10^{-8} or 10^{-20}. Hence, it is most preferable if the LSM approximating criterion takes such cases into account.

Therefore, a modification is proposed to transform the criterion into the following form

$$S^{**} = \sum_{i=1}^{n} \Big(\frac{\overline{y}_i}{y_i} - 1\Big)\Big(1 - \frac{y_i}{\overline{y}_i}\Big) \tag{2.2}$$

Owing to this, for both large values $\overline{y}_i$ (much different from the approximated value y_i) and small values (approaching zero) with respect to y_i, the components of the sum take significant values, i.e. in both cases, they give a significant (although – as it can be seen – diverse/unsymmetrical for each case) contribution to the total approximation error – as shown in Fig. 3a. In order to make the S_i components of sum (2.2) and the total sum S^{**} as an approximation criterion reach the minimum (not the maximum, as in Fig. 3a, also when reversal of the sign takes place between the approximated y_i and approximating value $\overline{y}_i$), the following form would be better

$$S^{**} = \sum_{i=1}^{n} \Big(\Big| \frac{\overline{y}_i}{y_i} \Big| - 1 \Big) \Big(1 - \Big| \frac{y_i}{\overline{y}_i} \Big| \Big) \tag{2.3}$$

Both extremes of the S_i function for both positive and negative values of are the minima shown in Fig. 3b.

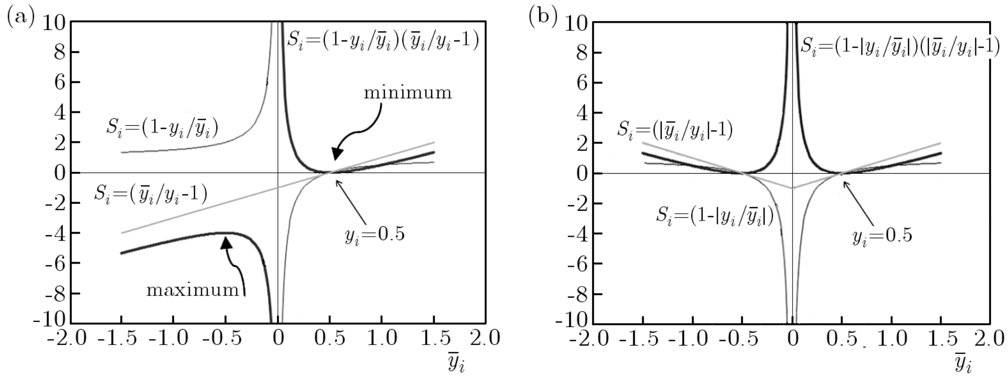

Fig. 3. Component of the sum for approximation criterion (2.2) – (a), and approximation criterion (2.3) – (b)

Obviously, due to the discontinuity of functions (2.2) and (2.3), the reversal of sign of the approximating value $\overline{y}_i$ in relation to the approximated value y_i is not beneficial as the question of achieving the minimum of the total sum S^{**} appears. However, it is not a significant disadvantage, since quite often such a situation means that the test data are not physically logical (e.g. when mass, body height, time or energy are measured), when it should not be expected that the approximated value becomes negative. Then the approximating values $\overline{y}_i$ with signs opposite to those of the approximated y_i have to be neglected as incorrect, and the discontinuity effect for the S_i components does not occur, since the graph shown in Fig. 3 consists of only positive or only negative parts.

In other cases (when the approximating value $\overline{y}_i$ and the approximated value y_i can have opposite signs, due to e.g. temperature measurements, income assessment, value gradient measurement, etc.). The criteria (2.2) and (2.3) show an advantageous feature of reversal of the sign (e.g. the component of the sum S_i rapidly increases). Owing to this, they contain a kind of "mathematical barrier" preventing easy reversal of the sign between the approximating and the approximated values. However, one should remember that a possible iteration step should be kept within some reasonable range so that it does not cause the omission of such discontinuity of the changes in S_i, since it may lead to improper estimation of the approximated value or disturb the convergence of calculations.

Approximation criterion functions for (1.3), (2.1) and (2.2) (and their components) as related to the approximating values $\overline{y}_i$, for:

- different approximated values y_i equal to 5, 2, 1, 0.25, 0.01, 0.00001
- the same range of variability of $\overline{y}_i$, i.e. $(-3y_i, 3y_i)$, in order to show the $\overline{y}_i \to 0$ effect, are shown in Fig. 4.

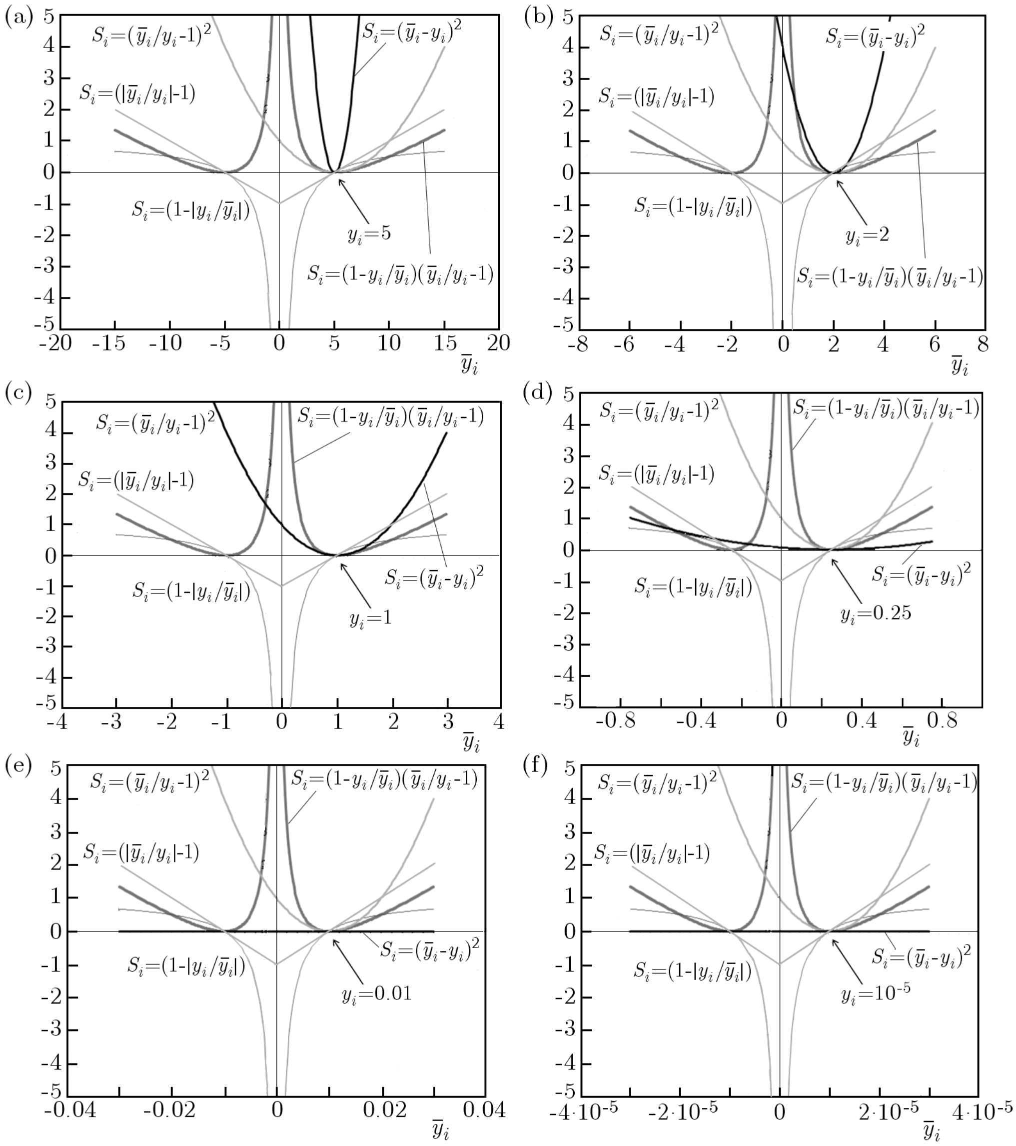

Fig. 4. Approximation criterion functions for (1.3), (2.1) and (2.3) for different approximated values y_i for constant approximating values scatter range within $(-3y_i, 3y_i)$ range

All advantages and disadvantages of the above presented LSM approximation criteria can be seen on the graphs above, in particular:

- significant dependence of values of components of the sum S (formula (1.3)) on the approximated value y_i
- invariability of values of the components of sums S^* (formula (2.1)) and S^{**} (formula (2.2) and (2.3)) on all the graphs, i.e. for any approximated value y_i
- no response of the components of sums S and S^* to the $\overline{y}_i \to 0$ effect, and dynamic change in the components of the sum S^{**} near value $\overline{y}_i = 0$.

The only curve that changes in the graphs presented in Fig. 4 is the plot for components of the sum S graph, i.e. for the standard form of the LSM.

The result of approximation with criterion (2.3) applied is shown in Fig. 5c – for the same set of data as in Figs. 1 and 2. Exemplary results of approximation for different test data (with slightly smaller scatter between individual $da/dN - f(\varDelta K)$ curves) is presented in Figs. 5a and 5b.

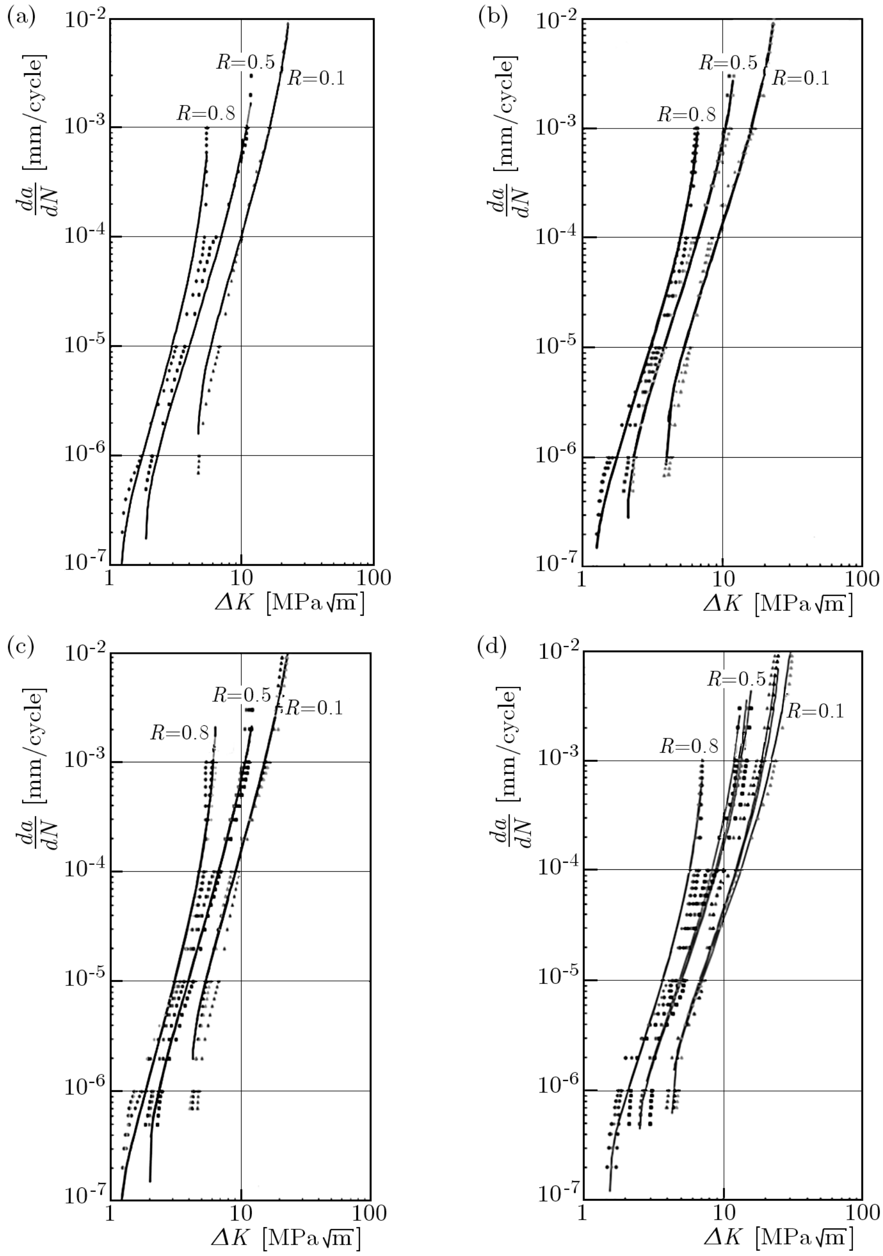

Fig. 5. Approximation of $da/dN = f(\Delta K)$ data in different variants with the NASGRO equation, LSM modified according to formula (2.3)

Favourable effects of the approximation (in comparison with the results shown in Fig. 1) after implementation of the modified LSM criterion can be easily seen. They tend to represent all the test data within the whole range of data variability, independently of their absolute values, independently of the number of described curves – 3 (Fig. 5a), 6 (Fig. 5b), 9 (Fig. 5c) and variant of the approximation (Fig. 5d – for each curve being described individually).

The above described test data come from the examination of aluminum alloy 2024 taken from Mi-2 helicopter's rotor blades [2].

This effect has been achieved only by modifying the LSM criterion, since the idea underlying the approximation method for all the presented graphs is identical – the minimum of the sum of squared deviations between the approximated test data and the approximating values.

3. Conclusion

The application of the Least Square Method in its classical form for determination of coefficients of the NASGRO equation that describes the fatigue crack propagation curve is ineffective, since data of the approximated function $da/dN = f(\Delta K)$ take values from the range of a few orders of magnitude.

The paper offers some technique to modify the LSM criterion for significant improvement of the approximation results.

The proposed modification of the LSM criterion brings favourable effects for results of the test data approximation (unachievable with the classical LSM method). These effects are as follows:

- the provision of equal "weights" of each test data point in the total sum that determines this criterion (i.e. the sum of deviations between the approximated and approximating values) – independently of the magnitude of difference between the values of data subject to approximation and that of the difference between the approximated and approximating values
- flexibility of the process of approximating in response to the reversal of sign between the approximated value and the approximating one, or to the approximation with values close to zero.

References

1. AFGROW Users guide and technical manual, AFRL-VA-WP-TR-2002-XXXX, Version 4.0005.12.10, James A. Harter, Air Vehicles Directorate, Air Force Research Laboratory, WPAFB OH 45433-7542
2. AFIT Report No 8c/07, Fatigue crack growth rate testing report, Warsaw, 2007 [in Polish]
3. ASTM E647 Standard test method for measurement of fatigue crack growth rates
4. Forman R.G., Shivakumar V., Cardinal J.W., Williams L.C., McKeighan P.C., 2005, *Fatigue Crack Growth Database for Damage Tolerance Analysis*, DOT/FAA/AR-05/15, U.S. Dep. of Transportation, FAA Office of Aviation Research, Washington, DC 20591
5. Fuchs H.O., Stephens R.I., 1980, *Metal Fatigue in Engeneering*, A Willey-Interscience Publication
6. Huang X.P., Cui W.C., Leng J.X., 2005, A model of fatigue crack growth under various load spectra, [In:] *Proc. of Sih G.C., 7th Int. Conf. of MESO*, Montreal, Canada, 303-308
7. Hudson M.C., Seward S.K., 1982, Compendium of sources of fracture toughness and fatigue crack growth data for metallic alloys, *International Journal of Fatigue*, **20**
8. NASGRO® Fracture mechanics and fatigue crack growth analysis software, v5.0, NASA-JSC and Southwest Research Institute, July 2006
9. nCode, User guide manual, ICE-FLOW cracks growth, 2003
10. Newman Jr. J.C., 1984, A crack opening stress equation for fatigue crack growth, *International Journal of Fracture*, **24**, 3, R131-R135
11. Taheri F., Trask D., Pegg N., 2003, Experimental and analytical investigation of fatigue characteristics of 350WT steel under constant and variable amplitude loadings, *Marine Structures*, **16**, 69-91
12. Walker K., 1970, The effect of stress ratio during crack propagation and fatigue for 2024-T3 and 7075-T6 aluminum, *ASTM STP*, **462**

13. White P., Molent L., Barter S., 2005, Interpreting fatigue test results using a probabilistic fracture approach. *International Journal of Fatigue*, **27**, 752-767

14. Zhao T., Jiang Y., 2008, Fatigue of 7075-T651 aluminum alloy, *International Journal of Fatigue*, **30**, 834-849

Modyfikacja kryterium Metody Najmniejszych Kwadratów do aproksymacji wyników badań prędkości propagacji pęknięć zmęczeniowych

Streszczenie

W pracy przedstawiono modyfikację kryterium Metody Najmniejszych Kwadratów do poprawienia dokładności aproksymacji wyników badań prędkości propagacji pęknięć zmęczeniowych. W szczególności wprowadzono zmiany do poprawy aproksymacji w przypadku, gdy przedział wartości przynajmniej jednej ze zmiennych jest w zakresie kilku rzędów wielkości lub różni się znacznie od zakresu zmienności drugiej zmiennej. Przedstawiono również różnice pomiędzy poszczególnymi kryteriami i ich wpływ na aproksymację wyników z badań. Obliczenia przeprowadzono dla danych uzyskanych z badań próbek ze stopu aluminium 2024 pobranego z łopat wirnika helikoptera Mi-2.

Manuscript received September 23, 2011; accepted for print November 23, 2011