2011 International Conference on Transportation, Mechanical, and Electrical Engineering (TMEE)
December 16-18, Changchun, China

Parameter Estimation of Three-parameter Weibull Distribution via Particle Swarm Optimization Algorithm

Xu Wen
Biological Engineering Branch
Jilin Business and Technology College
Changchun,China
Xuwen2004@sina.com

Xue Yu-xia
Biological Engineering Branch
Jilin Business and Technology College
Changchun,China
lace99@yahoo.cn

Shen Gui-xiang
College of Mechanical Science and Engineering
Jilin University
Changchun,China
shenguixiang@sohu.com

Zhang Ying-zhi
College of Mechanical Science and Engineering
Jilin University
Changchun,China
zhangyingzhi@sohu.com

Abstract—**Weibull distribution plays a very important role in reliability engineering. The parameter estimation of its three parameters is very difficult to solve. In this paper, we have briefly discussed this problem and proposed a new approach, i.e. the correlation coefficient optimization method based on the particle swarm optimization algorithm, and the feasibility and effectiveness of this method was demonstrated by two examples.**

Keywords-Three-parameter Weibull distribution, Parameter estimation, correlation coefficient optimization method, Particle swarm optimization(PSO)

I. Introduction

In real life situations, distribution parameters are usually unknown and must be estimated. Especially in reliability studies, parameter estimation is very crucial[1].

The Weibull distribution is widely used in reliability engineering and life data analysis because of its flexibility. This distribution has three parameters, and its parameter estimation process is quite complicated. For simplicity, the location parameter is often assumed to be zero and as a result, the three-parameter Weibull distribution model is turned to two-parameter one, whose parameter estimation process can be realized easily. Many methods have been presented to estimate the parameters of two-parameter Weibull distribution, such as Weibull probability plot[2], Modified Profile Likelihood, Robust estimation method[3], bias correction method and linear regression method [4,5]. However, the three-parameter Weibull distribution is more robust than two-parameter Weibull distribution [6]. To solve the parameter estimation, this paper presents a new method, namely, the correlation coefficient optimization method based on the PSO algorithm. First of all, assume that the location parameter is known, Weibull distribution is transformed into linear equation in accordance with the linear regression method, so that the other two parameters and correlation coefficient can be expressed as functions of location parameter according to the least square method. The PSO algorithm is used to determine the optimal correlation coefficient so as to estimate the three parameters.

II. Correlation coefficient optimization method

The cumulative probability distribution function of three-parameter Weibull distribution is given by

$$F(t)=\begin{cases}1-\exp[-(\frac{t-\gamma}{\eta})^{\beta}] & t>\gamma \\ 0 & elsewhere\end{cases} \tag{1}$$

Where η, β and γ are scale, shape, and location parameters respectively, and η>0, β>0, γ≥0.

After transforming both sides of equation (1), we can get:

$$\ln(-\ln(1-F(t)))=-\beta\ln\eta+\beta\ln(t-\gamma) \tag{2}$$

Let $Y=\ln(-\ln(1-F(t)))$, $X=\ln(t-\gamma)$, $b=\beta$, $a=-\beta\ln\eta$, then equation (2) can be translated into the following linear equation:

$$Y=a+bX \tag{3}$$

Presume that $(t_1,F(t_1)),(t_2,F(t_2)),\cdots(t_n,F(t_n))$ are a set of life data, where $t_1<t_2<\cdots<t_n$,then they can be translated into corresponding ($(x_1,y_1),(x_2,y_2),\cdots(x_n,y_n)$) through equation (1) & (2).

According to the least square method principle, let:

$$\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i,\ \bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i \tag{4}$$

$$L_{xx}=\sum_{i=1}^{n}x_i^2-n\bar{x}^2 \tag{5}$$

$$L_{yy}=\sum_{i=1}^{n}y_i^2-n\bar{y}^2 \tag{6}$$

$$L_{xy} = \sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y} \quad (7)$$

We can get the values of undetermined parameter a,b and correlation coefficient ρ:

$$\hat{b} = \frac{L_{xy}}{L_{xx}}, \hat{a} = \bar{y} - \hat{b}\bar{x} \quad (8)$$

$$\rho = \frac{L_{xy}}{\sqrt{L_{xx} L_{yy}}} \quad (9)$$

It is obviously that a,b and ρ are all functions of γ. Value of γ must meet that the absolute value of linear correlation coefficient |ρ| be the maximum one, namely $\frac{d|\rho|}{d\gamma} = 0$. That is:

$$\frac{d\rho^2}{d\gamma} = 2\rho^2 \left(\frac{1}{L_{xy}} \frac{dL_{xy}}{d\gamma} - \frac{1}{2L_{xx}} \frac{dL_{xx}}{d\gamma}\right) = 0 \quad (10)$$

(10)

This equation is very complicated and is difficult to solve. Literature [7] considers the iteration method, that is, increase from 0 to t1 in accordance with a certain step length of iteration increment, then calculate the linear correlation coefficient ρ of each iteration, where the value corresponding to the maximum |ρ| is the estimated value of the location parameter. However, this iterative method takes long computation time, and the optimal solution may be missed if the step length is too large, it will spend too much time if the step length is too small. The determination of the optimal linear correlation coefficient can be easily achieved via PSO algorithm.

III. CORRELATION COEFFICIENT OPTIMIZATION METHOD BASED ON PSO ALGORITHM

A. PSO theory

PSO theory was derived from the social behavior of organisms such as bird flocking. It was first proposed by Kennedy and Eberhart in 1995 [8]. Because of its easy realization, high precision and rapid convergence speed, it draws much attention from the academic community, and shows its superiority in solving practical problems. It has been successfully applied in many fields [9-11]. PSO is an evolutionary algorithm, and it has turned out to be a worthy alternative to the Genetic Algorithm (GA) and other optimization techniques. Similar to genetic algorithm(SA), PSO starts from a random solution and find the optimal solution through iteration; it also uses the fitness to evaluate the quality of the solutions. But its rule is simpler than that of SA. PSO has no "cross" and "mutation" operator. Furthermore, PSO requires less processing time than SA[12] .

The basic idea of the particle swarm optimization algorithm is: each solution of the optimization issue is called a particle; define a fitness function to measure the fitness degree of each particle solution. The particles swarm according to the "flight experience" from their own and other particles, so as to achieve the purpose of searching the optimal solution within the whole solution space. The specific search process is as follows:

The system is initialized into a group of random solutions (random particles): $X(0) = (X_1(0), X_2(0), \cdots, X_p(0))$. Each particle in the solution space approaches towards two solutions at the same time, one is the global optimal solution *gbest*, which is the optimal solution found by all the particles of the whole population in all the past searching processes; the other one is individual optimal solution *pbest*, which is found by the particles themselves in all the past searching processes. In each evolution, the particles are updated by tracking these two optimal solutions. After finding the two optimal solutions, the speed and the position of the ith particle in the kth generation are updated according to the following equations [13-14]:

$$V_i(k+1) = K[wV_i(k) + c_1 r_1 (X_i^{pbest}(k) - X_i(k)) + c_2 r_2 (X_g^{gbest}(k) - X_i(k))] \quad (11)$$

$$X_i(k+1) = X_i(k) + V_i(k+1) \quad (12)$$

Where,

K is convergence factor, and

$$K = \frac{2}{\left|2 - c - \sqrt{c^2 - 4c}\right|}, c = c_1 + c_2 > 4 \quad (13)$$

w is inertial weight, and

$$w = w_{\max} - k * \frac{w_{\max} - w_{\min}}{N} \quad (14)$$

N is the total evolution generation number,

c1 and c2 are acceleration constants, usually taken non-negative constants between 0 and 4,

r1 and r2 are random numbers from 0 to 1.

B. Procedures

The Procedures of parameter estimation of three-parameter Weibull distribution via PSO-based correlation coefficient optimization method is as follows:

Step 1. Initialize a particle group with size p, set the initial position and initial velocity.

Step 2. Calculate the fitness value of each particle, let the fitness function $Fitness = -|\rho(\gamma)|$.

Step 3. Compare the fitness value of each particle with that of its best individual position pbest, choose the better one as the best individual position.

Step 4. Compare the fitness value of each particle with that of its best global position gbest , choose the better one as the best global position.

Step 5. Update the particles according to the position evolution equation (11) and the velocity evolution equation (12).

Step 6. If the halting condition is satisfied, output the solution, or else return to Step 2.

See Fig. 1 for the flowing chart of this method.

IV. NUMERICAL EXAMPLES

Two examples are considered to illustrate the feasibility of applying the proposed approach to parameter estimation of three-parameter Weibull distribution. In addition, to know the effects of the sample size on the

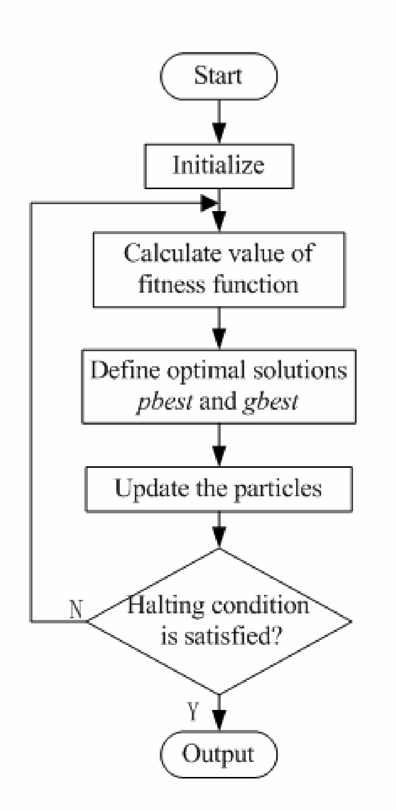

Fig.1 Flowing chart of correlation coefficient optimization method based on PSO algorithm

algorithm, samples of size20, 50, 100, 500, and 1000 have been taken and used to estimate the parameters. For both examples, let $c_1 = c_2 = 2.05$, $w_{\max} = 0.9, w_{\min} = 0.4$, p=30 and N＝10. All the coding has been done using Matlab 7, and has been run on a computer with 1.73GHz of CPU and 2 GB of RAM. It is straightforward from the primary estimation theory that the bigger the sample size the better the estimation. However, here as the sample size increases the fitness function will be more complicated. Therefore, the selection of the sample size is a matter of compromise.

Example 1. The first example here is to estimate the parameters of a Weibull distribution with η=2, β=5, γ=10, which is hereafter referred to as Wei(2,5,10).To get a better understanding of the algorithm, there has been listed the estimated values of the three parameters for different sample sizes and the run time. Table 1 shows the results. It is quite clear that as the sample size increases the estimation will be better. But as noted before because of fitness function will be more complicated with bigger sample sizes, more run time will be taken for the algorithm to settle down to the estimates.

Table1 Simulation results for Wei(2,5,10)

Sample size	20	50	100	500	1000
η	2.1777	2.0547	1.9513	1.9881	1.9969
β	5.5381	5.1821	4.8426	4.9603	4.9903
γ	9.8243	9.9457	10.0482	10.0118	10.0030
Run time [s]	2.61	2.688	3.047	3.875	5.579

Example 2. This example considers a Weibull distribution with η=5, β=35, γ=8, which is hereafter referred to as Wei(5,35,8). The results contained in Table 2 are same as that of the first example.

Table2 Simulation results for Wei(5,35,8)

Sample size	20	50	100	500	1000
η	4.7502	4.8391	4.8348	4.8980	4.9050
β	33.2035	33.8386	33.8011	34.2589	34.3123
γ	8.2498	8.1608	8.1651	8.1020	8.0950
Run time [s]	2.75	3.157	3.312	4.281	5.328

According to the above two examples, it can be known that the estimated parameter values are approximately equal to the actual values after 10 generations of evolution within a very short period of time.

V. CONCLUSION

PSO-based correlation coefficient optimization method for parameter estimation of three-parameter Weibull distribution is processed of simple calculation, short computation time, good robustness. It applies to not only large size of sample data, but also small size of sample data.

REFERENCES

[1] Ling J, Pan J. A new method for selection of population distribution and parameter estimation[J]. Reliab Eng Syst Saf 1998;60:247-55.

[2] L.F. Zhang, M. Xie, L.C. Tang. A study of two estimation approaches for parameters of Weibull distribution based on WPP[J]. Reliability Engineering and System Safety,2007;92:360–368.

[3] Yang ZL, Xie M. Efficient Estimation of the Weibull Shape Parameter Based on a Modified Profile Likelihood[J]. J Stat Comput Simulation 2002;73(2):115–123.

[4] Ho LL, DA Silva AF. Bias Correction for Mean Time to Failure and p-quantiles in a Weibull Distribution by Bootstrap Procedure[J]. Communi Stat-Simulation Comput,2005;34(3):617–629.

[5] Ahmed ES, Volodin AI, Hussein AA. Robust Weighted Likelihood Estimation of Exponential Parameters[J]. IEEE Trans Reliab, 2005;54(3):389–395.

[6] Maha Alqam, Richard M. Bennett, Abdul-Hamid Zureick. Three-parameter vs. two-parameter Weibull distribution for pultruded composite material properties[J]. Composite Structures,2002; 58:497-503.

[7] Yu Jie,Shen Gui-xiang, Jia Ya-zhou. Reliability evaluation on CNC lathes based on three Weibull distributions[J], MODERN MANUFACTURING ENGINEERING,2007;(5):18-20.in Chinese.

[8] Kennedy J, Eberhart R C. Particle swarm optimization[C]. Proc. IEEE Int. Conf. on Neural Networks, Perth, WA,Australia, 1995;1942-1948.

[9] J.P. Coelho, P.B. de Moura Oliveira, J. Boaventura Cunha. Greenhouse air temperature predictive control using the particle swarm optimization algorithm[J].Computers and Electronics in Agriculture, 2005; 49(3):330-344.

[10] Ammar W. Mohemmed, Nirod Chandra Sahoo, Tan Kim Geok. Solving shortest path problem using particle swarm optimization[J]. Applied Soft Computing, 2008; 8(4):1643-1653.

[11] Wen-Shing Lee, Chung-Kuan Kung .Optimization of heat pump system in indoor swimming pool using particle swarm algorithm[J].Applied Thermal Engineering, 2008; 28(13):1647-1653.

[12] Te-Hsiu Sun.Applying particle swarm optimization algorithm to roundness measurement[J]. Expert Systems with Applications, 2009;36:3428–3438.

[13] M. Clerc, J. Kennedy. The particle swarm-explosion, stability and convergence in a multidimensional complex space[J], IEEE Transactions on Evolutionary Computation, 2002;6(1):58-73.

[14] M.M. Ali, P. Kaelo. Improved particle swarm algorithms for global optimization[J]. Applied Mathematics and Computation, 2008;196(2): 578-593.