

# Welch 谱估计的随机误差与置信度

朱学旺, 刘青林, 张思箭

(中国工程物理研究院总体工程研究所, 四川 绵阳 621999)

**摘要:** 目的 分析 Welch 谱估计的随机误差及置信度。方法 首先导出 Welch 谱估计量随机误差的表达式, 然后采用正态标准化处理方法, 将 Welch 谱估计分量转化为  $\chi^2$  变量。结果 获得了 Welch 谱估计的置信度分析方法。结论 以 128 自由度的 Welch 谱估计结果的随机误差(12.5%)作为相对置信区间, 则其置信概率为 68.4%, 若要达到 95.5% 以上置信概率, 则 Welch 谱估计的自由度不能小于 512, 相应的随机误差为 6.25%。

**关键词:** 功率谱估计; Welch 方法; 随机误差; 置信度; 置信概率; 置信区间

**DOI:** 10.7643/issn.1672-9242.2015.01.008

**中图分类号:** TJ01; TP311.13 **文献标识码:** A

**文章编号:** 1672-9242(2015)01-0034-05

## Random Error and Confidence Level of Welch Spectral Estimation

ZHU Xue-wang, LIU Qing-lin, ZHANG Si-jian

(Institute of Systems Engineering, CAEP, Mianyang 621999, China)

**ABSTRACT:** **Objective** To analyze the random error and confidence level of Welch estimation for power spectral density (PSD). **Methods** The expression for Welch PSD estimation random error was firstly deduced, and the Welch spectral estimation component was converted into  $\chi^2$  variable by standardizing as a square sum of normal Gaussians. **Results** An analysis method for the confidence level of Welch spectral estimation was obtained. **Conclusion** The confidence probability for a 128 DOF Welch PSD estimation was 68.4% when the relative confidence interval described by the random error was 12.5%. To increase the confidence probability to 95.5%, the DOF of Welch PSD estimation must be no less than 512, corresponding to a random error of 6.25%.

**KEY WORDS:** PSD estimation; Welch PSD estimation; random error; confidence level; confidence probability; confidence interval

振动环境工程研究中, 通常采用 Welch 方法来获得随机振动信号的功率谱密度估计(Welch 谱估计)<sup>[1-2]</sup>。Welch 谱估计本质上是修正周期图方法的一种, 其做法是通过将振动信号的分段平滑、数据重

叠以及加窗处理等技术, 以达到降低谱估计方差的目的。Welch 谱估计是一致估计, 即当用于谱估计的每段数据足够长且数据段足够多时, Welch 谱估计结果的偏差和方差均等于 0 或趋于 0<sup>[3-4]</sup>。令人遗憾的

收稿日期: 2014-09-25; 修订日期: 2014-10-19

Received: 2014-09-25; Revised: 2014-10-19

作者简介: 朱学旺(1963—), 男, 湖北鄂州人, 硕士, 研究员, 主要研究方向复杂结构动力学及装备振动环境工程。

**Biography:** ZHU Xue-wang(1963—), Male, from Ezhou, Hubei, Master, Research Professor, Research focus: complicated structure dynamics and materiel environmental vibration engineering.

是,工程中用于谱估计的数据总是有限的,既不能保证每段数据足够长,也不能保证数据段足够多,即使对于平稳随机过程也是如此。为了便于操作,GJB 150.16A<sup>[5]</sup>和 MIL-STD-810F<sup>[6]</sup>对 Welch 谱估计给出了推荐要求:统计自由度不少于 120,在试验频带内至少保证 400 线。换句话说,至少要 60 段数据,每段数据的长度要确保谱分析后的带宽小于 5 Hz(当试验频率上限为 2000 Hz 时)。在这种状态(自由度取 128)下获得的 Welch 谱估计,可以证明其随机误差与相同自由度  $\chi^2$  的随机误差相同(约为 12.5%)。也就是说,该 Welch 谱估计的标准差与其均值(真值)的比为 12.5%,但是其在一定置信概率下的置信区间并不知道。

为了提高 welch 估计的精度,加窗技术的研究一直受到相关研究人员的关注<sup>[7]</sup>。合理的窗函数,可以减小因为数据截断带来的能量泄露,继而减小谱估计的系统误差。谱估计的精度可以采用估计的偏差与估计的方差表示,小的方差表明单次估计接近真值的可能性大,是一种定性描述<sup>[8]</sup>。Welch 谱估计的精度也可以定量描述,此时将谱估计结果作为一个随机变量的一次实现,可采用置信度分析方法获得其置信度与置信区间。文献[8—10]采用定量方法讨论了 Welch 谱估计的精度。针对不同的分布概率,其置信度分析的重点各异,文献[11]对置信度分析的现状进行了总结,而文献[12—13]对  $t$  分布变量的置信区间开展了专门研究。形式上,Welch 谱估计可以表述为多个变量的平方和,与数学上定义的  $\chi^2$  分布变量具有类似的形式。一些经典的谱分析论著在讨论谱估计时也直接给出了与  $\chi^2$  分布变量相同的随机误差<sup>[3,14—15]</sup>,却没有进行详细的说明。文中给出了 Welch 谱估计随机误差的导出过程,试图证明其正确性。另外,为了进一步理解 Welch 谱估计随机误差的物理意义,研究了 Welch 谱估计结果的置信概率及对应的置信区间。通过正态标准化处理,能够将  $\chi^2$  变量的置信度分析方法应用于 Welch 谱估计结果的置信度分析。

## 1 $\chi^2$ 变量统计分析方法<sup>[14]</sup>

数学上定义的  $\chi^2(n)$  变量为  $n$  个独立的  $N(0,1)$  变量  $x_1, x_2, \dots, x_n$  的平方和。即:

$$\chi^2(n) = x_1^2 + x_2^2 + \dots + x_n^2$$

获得  $\chi^2(n)$  的均值  $m$  和方差  $\sigma^2$ :

$$m = E[\chi^2(n)] = nE[x^2] = n$$

$$\sigma^2 = E[(\chi^2(n))^2] - (E[\chi^2(n)])^2 = 2n$$

以均值和方差表述随机过程  $\chi^2(n)$  的随机误差为:

$$\frac{\sigma}{m} = \sqrt{\frac{2}{n}} \tag{1}$$

$\chi^2(n)$  变量的概率分布  $p(Z = \chi^2(n))$  为:

$$p(Z = \chi^2(n)) = [2^{n/2} \Gamma(n/2)]^{-1} (Z)^{(n/2-1)} e^{-Z/2}, Z \geq 0 \tag{2}$$

式中:  $\Gamma(n/2)$  为 Gamma 函数,  $\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$ 。

置信概率与置信区间存在以下关系:

$$P\{a \leq Z \leq b\} = P\{a < Z\} - P\{b < Z\} = \int_a^b p(Z) dZ - \int_b^\infty p(Z) dZ = \int_a^\infty \left[ \Gamma\left(\frac{n}{2}\right) \right]^{-1} (Z)^{(n/2-1)} e^{-Z/2} dZ - \int_b^\infty \left[ \Gamma\left(\frac{n}{2}\right) \right]^{-1} (Z)^{(n/2-1)} e^{-Z/2} dZ = I(a, n) - I(b, n) \tag{3}$$

式中:  $a, b$  分别为置信区间的下限和上限,可查表得到  $I(a, n), I(b, n)$  后反推获得。自由度数较大的  $\chi^2(n)$  变量的分位数见表 1, 不同置信概率 80%, 95%, 99% 下的置信区间结果(利用了对称性假设)如图 1 所示<sup>[16]</sup>。

已知置信区间的上下限  $a, b$  时,可直接求得其对应的置信概率;反之,已知置信概率也可以导出对应的置信区间,只是此时需要假定上下限之外的概率分布。通常的做法是假定对称<sup>[16]</sup>,即小于置信区间下限的概率与大于置信区间上限的概率相等。

表 1  $\chi^2(n)$  分布的概率分布

Table 1 Probability contribution of large D.O.F.  $\chi^2(n)$  variable

$n$	$p$													
	0.99	0.98	0.95	0.9	0.8	0.7	0.5	0.3	0.2	0.1	0.05	0.02	0.01	0.001
64	40.65	42.96	46.59	49.99	54.34	57.62	63.34	69.42	73.28	78.86	83.68	89.34	93.22	104.72
128	93.74	97.32	102.87	107.97	114.38	119.15	127.33	135.88	141.24	148.88	155.40	162.96	168.13	183.19
256	206.32	211.70	219.95	227.46	236.79	243.67	255.33	267.36	274.82	285.39	294.32	304.59	311.56	331.66
512	440.51	448.44	460.53	471.44	484.89	494.75	511.33	528.28	538.72	553.41	565.75	579.84	589.37	616.61

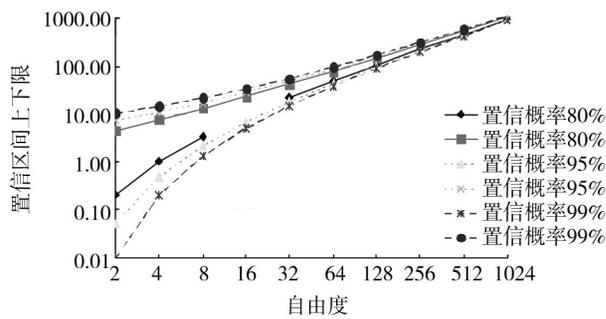


图1  $\chi^2(n)$ 变量的置信区间随统计自由度的变化

Fig.1 Relationship between confidence interval and D.O.F. of  $\chi^2(n)$  variable

## 2 Welch 谱估计的随机误差

为了导出 Welch 谱估计随机误差的表达式, 首先回顾 Welch 谱估计方法。本质上, Welch 谱估计是  $B-T$  估计的一种修正, 其技术特点是加窗、分段平滑和重叠。具体过程是: 将  $N$  个数据  $\{x(0), x(1), \dots, x(N-1)\}$  分为  $K$  段, 每段数据长度为  $L$ , 其中有  $L-D$  个数据为相邻段的重叠数据 ( $D \leq L$ ), 则:

$$N=L+D(K-1) \quad (4)$$

第  $i$  段的  $L$  个数据为:

$$x_i(n)=x(n+iD), n=0, 1, \dots, L-1, i=0, 1, \dots, K-1$$

对每一段数据进行加窗处理, 并分别计算其  $B-T$  周期图, 有:

$$\hat{S}_M^{(i)}(f) = \frac{1}{LU} \sum_{n=0}^{L-1} |w(n)x(n+iD)e^{-j2\pi fn}|^2 \quad (5)$$

式中:  $U = \frac{1}{L} \sum_{n=0}^{L-1} |w(n)|^2$  为数据窗  $w(n)$  的能量参数, 引入它的目的是为了确保周期图  $\hat{S}_M^{(i)}(f)$  为渐近无偏的。

然后对各周期图进行平滑处理, 获得  $N$  个数据的 Welch 谱估计:

$$\hat{S}_w(f) = \frac{1}{K} \sum_{i=0}^{K-1} \hat{S}_M^{(i)}(f) \quad (6)$$

由公式(5), (6)可知, Welch 谱估计与原始信号 (已经加窗处理) 傅立叶变换的实部与虚部的平方和成正比。这样, 公式(6)可改写为:

$$\hat{S}_w(f) = \gamma \sum_{m=0}^{K-1} \{ [X_\alpha^m(f)]^2 + [X_\beta^m(f)]^2 \} \quad (7)$$

式中:  $\gamma$  为常数,  $\gamma = \frac{1}{KLU}$ ;  $X_\alpha^m(f), X_\beta^m(f)$  为经加窗处理的第  $m$  段数据对应的傅立叶变换之实部与虚部 ( $m=0, 2, \dots, K-1$ )。

可以证明, 当原始信号满足零均值正态分布时, 其傅立叶变换的实部与虚部也满足零均值正态分布, 且相互独立<sup>[15]</sup>, 但是, 不能保证其方差为 1。为了能够利用第 1 节的结论, 对傅立叶变换的实部与虚部进行正态标准化处理, 此时, 式(7)改写为:

$$\hat{S}_w(f) = \theta \hat{Q}_x(f) = \theta \sum_{m=0}^{K-1} \left\{ \left[ \frac{X_\alpha^m(f)}{\lambda} \right]^2 + \left[ \frac{X_\beta^m(f)}{\lambda} \right]^2 \right\} \quad (8)$$

式中:  $\theta = \gamma \lambda^2$ ,  $\lambda$  为傅立叶变换的实部与虚部的标准差 (假设其相同)。显然,

$$\hat{Q}_x(f) = \sum_{m=0}^{K-1} \left\{ \left[ \frac{X_\alpha^m(f)}{\lambda} \right]^2 + \left[ \frac{X_\beta^m(f)}{\lambda} \right]^2 \right\} \quad (9)$$

式中:  $\hat{Q}_x(f)$  是自由度为  $2K$  的  $\chi^2$  变量。因此, 容易导出:

$$\begin{aligned} m_w &= E[\hat{S}_w(f)] = \theta E[\hat{Q}_x(f)] = 2K\theta \\ \sigma_w^2 &= E[(\hat{S}_w(f))^2] - (E[\hat{S}_w(f)])^2 = \\ &= E[(\theta \hat{Q}_x(f))^2] - (E[\theta \hat{Q}_x(f)])^2 = \\ &= \theta^2 (E[(\hat{Q}_x(f))^2] - (E[\hat{Q}_x(f)])^2) = 4K\theta^2 \end{aligned}$$

于是, Welch 谱估计  $\hat{S}_w(f)$  的随机误差为:

$$\varepsilon = \frac{\sqrt{\sigma_w^2}}{m_w} = \sqrt{\frac{1}{K}} \quad (10)$$

这正是  $\chi_n^2$  变量的随机误差表达式。

## 3 Welch 谱估计的置信分析

依据公式(8)不难看出,  $a \leq \hat{S}_w(f) \leq b$  的置信概率与  $\frac{a}{\theta} \leq \hat{Q}_x(f) \leq \frac{b}{\theta}$  的置信概率是完全相同的, 而  $\hat{Q}_x(f)$  的置信分析可以按照第 1 节介绍的方法进行, 因此, Welch 谱估计的置信度也能够相应导出。即:

$$\begin{aligned} P\{a \leq \hat{S}_w(f) \leq b\} &= P\left\{\frac{a}{\theta} \leq \hat{Q}_x(f) \leq \frac{b}{\theta}\right\} = \\ &= I\left(\frac{a}{\theta}, 2K\right) - I\left(\frac{b}{\theta}, 2K\right) \end{aligned} \quad (11)$$

一般而言,  $\theta$  值的确定并非易事, 因此, 给定置信概率下, Welch 谱估计  $\hat{S}_w(f)$  的置信区间  $[a, b]$  无法直接导出<sup>[17]</sup>。

$$\begin{aligned} \text{由 } \frac{a}{E[\hat{S}_w(f)]} \leq \frac{\hat{S}_w(f)}{E[\hat{S}_w(f)]} \leq \frac{b}{E[\hat{S}_w(f)]} \text{ 可得:} \\ \frac{a/\theta}{E[\hat{Q}_x(f)]} \leq \frac{\hat{S}_w(f)}{E[\hat{Q}_x(f)]} \leq \frac{b/\theta}{E[\hat{Q}_x(f)]} \end{aligned} \quad (12)$$

式(12)表明,尽管无法确定因为标准化处理而引入的常数因子 $\theta$ 的具体数值(当然也无法定量给出Welch谱估计的置信区间),但是却可以定量描述Welch谱估计与其均值的比的置信范围(不妨称之为相对置信区间),而且取值恰好是标准 $\chi_n^2$ 变量的置信区间结果。

同样地,如果给定置信区间 $[a, b]$ ,直接求解Welch谱估计 $\hat{S}_w(f_i)$ 的置信概率是困难的。这时可以将置信区间 $[a, b]$ 转化为的相对置信范围 $[a_1, b_1]$ ,即:

$$a_1 = \frac{a/\theta}{E[\hat{Q}_x(f_i)]}, b_1 = \frac{b/\theta}{E[\hat{S}_w(f_i)]} \Rightarrow a/\theta = 2Ka_1, \\ b/\theta = 2Kb_1 \quad (13)$$

这样,便可以获得相应的置信概率。

## 4 算例及分析

算例一<sup>[14]</sup>:假定谱估计结果 $S_0$ 满足 $\chi^2$ 分布,其分析自由度为10,则其在置信概率为80%的置信区间为(4.87, 15.99),即:

$$\frac{4.87}{10} < \frac{S_0}{S_M} < \frac{15.99}{10} \Rightarrow 0.625S_0 < S_M < 2.05S_0$$

式中: $S_M$ 为谱估计变量的统计均值。

相应地,99%的置信概率对应的置信区间为: $0.397S_0 < S_M < 4.63S_0$ 。

算例二:假定 $S_0$ 的自由度为64,重复算例一的计算过程。其结果为: $0.82S_0 < S_M < 1.28S_0$ 对应于置信概率80%; $0.66S_0 < S_M < 1.66S_0$ 对应于置信概率99%。

算例三:假定 $S_0$ 的自由度为128,重复算例一的计算过程。其结果为: $0.86S_0 < S_M < 1.19S_0$ 对应于置信概率80%; $0.74S_0 < S_M < 1.41S_0$ 对应于置信概率99%。

从上述3个算例可以看出,随着自由度的增加,Welch谱估计的置信区间越来越趋向于均值集中。应用公式(12)可知,上述结果同样适用于工程中经常出现的高斯过程(均值为0,方差不等于1)的Welch谱估计。

GJB 150.16A和MIL-STD-810F都推荐谱估计的自由度要大于120,下面针对算例三进行专门分析。其结果表明,对应于置信概率80%,128自由度的Welch谱估计的均值在估计值的0.86~1.19倍之间,换句话说,估计值可能是均值的0.84~1.16倍;对应于置信概率99%,128自由度的Welch谱估计的均值在估计值的0.74~1.41倍之间,即估计值可能是均

值的0.71~1.35倍。这一结果应该引起工程界的重视,因为获得的Welch谱估计可能比真实的谱高,也可能低。

128自由度的Welch谱估计的随机误差为12.5%,用此构造相对置信区间,则具有相同置信概率的 $\chi^2$ 变量的置信区间为(87.5%×128, 112.5%×128),对应的置信概率为68.4%。这是一个令人遗憾的结果,其表明128自由度的Welch谱估计结果仅有68.4%的置信概率能够实现其结果在真值的上下12.5%的范围内。

为了提高置信概率(依然以均值的上下波动12.5%作为相对置信区间),增加Welch谱估计的自由度,例如自由度取512,则其置信概率为95.5%。即有95.5%的可能,保证512自由度的Welch谱估计结果在均值的上下12.5%的范围内,此时,Welch谱估计的随机误差为6.25%。从这个意义上说,也许标准推荐的Welch谱估计的自由度120未必合适,尤其对于具有长数据的平稳随机振动的项目。

## 5 结论

1) Welch谱估计的随机误差与标准 $\chi^2$ 变量的随机误差相同,仅与统计自由度有关。

2) 基于 $\chi^2$ 变量的置信分析方法可以应用于Welch谱估计的置信分析。

3) 128自由度的Welch谱估计的随机误差为12.5%,以此作为谱估计的相对置信区间,其置信概率仅为68.4%,若要提高置信概率到95%,则Welch谱估计的自由度要512以上,此时的随机误差小于6.25%。

4) 结论未考虑 $\chi^2$ 变量概率密度函数的不对称性。

### 参考文献:

- [1] 邓泽怀,刘波波,李彦良. 常见的功率谱估计方法及其Matlab仿真[J]. 电子科技,2014,27(2):50—52.  
DENG Ze-huai, LIU Bo-bo, LI Yan-liang. Common Power Spectrum Estimation Methods and Matlab Simulation[J]. Electronic Sci & Tech, 2014, 27(2): 50—52.
- [2] 姚武川,姚天任. 经典谱估计方法的Matlab分析[J]. 华中理工大学学报,2000,28(4):45—47;  
YAO Wu-chuan, YAO Tian-ren. Analyzing Classical Spectral Estimation by MATLAB[J]. J Huazhong Univ of Sci & Tech, 2000, Vol.28(4): 45—47
- [3] 姚天任,江太辉. 数字信号处理[M]. 武汉:华中科技大学

出版社,2000.

YAO Tian-ren, JIANG Tai-hui. Digital Signal Processing [M]. Wuhan: Huazhong Univ of Sci & Tech Press, 2000.

[4] LINDA S, MALENKA M, WOLFGANG M, et al. Optimized Spectral Estimation for Nonlinear Synchronizing Systems[J]. Physical Review, 2014, 89(3):032912.

[5] GJB 150.16A—2009, 军用装备实验室环境试验方法 第16部分 振动试验[S].  
GJB 150.16A—2009, Military Equipment Laboratory Environmental Test Methods Part 16: Vibration Test[S].

[6] MIL-STD-810F, Environmental Engineering Considerations and Laboratory Tests, Method 514.5 vibration[S].

[7] TURGAY K, MELIH C I. The Obtaining of Window Function Having Useful Spectral Parameters by Helping of Genetic Algorithm[J]. Procedia-Social and Behavioral Sciences, 2013, 83:563—568.

[8] ALEKSEEV V G. "Welch-Type Estimator for a Spectral Density Function, the Case of Discrete-Time Parameter", [J]. Avtometriya, 2001(6):77—91.

[9] ALEKSEEV V G, SUKHODOEV V A. Welch-Type Spectral Density Estimator, Additional Recommendations[J]. Optoelectronics, Instrumentation and Data Processing, 2008, 44(4):302—305.

[10] ALEKSEEV V G, SUKHODOEV V A. Welch-Type Estimator for a Spectral Density Function, Case of a Continuous-time[J]. Optoelectronics, Instrumentation and Data Processing, 2009, 45(2):107—112.

[11] RICHARD K B, JORGE Q, HARIHARAN K I. The Present Status of Confidence Interval Estimation for One-factor Random Models[J]. Journal of Statistical Planning and Inference, 2006, 136:4307—4325

[12] JOANNA T. Confidence Intervals for the Power of Student's t-test[J]. Statistics & Probability Letters, 2005, 73: 125—130.

[13] DENNIS G, MINGFEI L. A Note on Confidence Intervals for the Power of T-test[J]. Statistics & Probability Letters, 2008, 78:488—489.

[14] 纽兰 D E. 随机振动与谱分析概论[M]. 北京:机械工业出版社, 1980.  
NEWLAND D E. An Introduction to Random Vibration and Spectral Analysis[M]. Beijing: China Machine Press, 1980.

[15] 戴诗亮. 随机振动实验技术[M]. 北京:清华大学出版社, 1984.  
DAI Shi-liang. Random Vibration Experiment Technique [M]. Beijing: Tsinghua University Press, 1984.

[16] 蒋书法. 最短置信区间的求解[J]. 上海电力学院学报, 2014, 30(2): 188—192.  
JIANG Shu-fa. Solution to the Shortest Confidence Interval [J]. Journal of Shanghai University of Electric Power, 2014, 30(2):188—192.

[17] 吴昌昊, 龚俊, 刘子琪. 基于 C U D A 实现经典功率谱估计[J]. 四川兵工学报, 2013(10):98—101.  
WU Chang-hao, GONG Jun, LIU Zi-qi. Achievement of Classic Power Spectrum Estimation Based on CUDA[J]. Journal of Sichuan Ordnance, 2013(10):98—101.

(上接第30页)

on Least Squares Function[J]. Journal of Safety Science and Technology, 2013, 11(9):83—90.

[22] ZOU Li-hua, DAI Su-liang, JOHN B, et al. Grey for Forecasting Model Active Vibration Control Systems[J]. Journal of Sound and Vibration, 2009, 322:690—706.

[23] 倪凯. 灰色系统在量测数据处理中的应用[J]. 中国水运, 2012, 12(8):47—49.  
NI Kai. Grey System in Measure Ment Data Handl[J]. Chinese Waterway, 2012, 12(8): 47—49.

[24] 彭涛, 杨岸英, 梁杏, 等. BP-神经网络与灰色系统法的联合模型来预测软基沉降量[J]. 岩土力学, 2005, 26(11): 1810—1814.  
PENG Tao, YANG An-ying, LIANG Xing, et al. The Joint Model of BP-neural Net Work and Grey System Method to Predict Soft Ground Settlement[J]. Rock and Soil Mechanics, 2005, 26(11):1810—1814.

[25] 唐五湘. GM(1,1)时序修正模型及其应用[J]. 云南环境科学, 2003, 22(4):20—21, 45.  
TANG Wu-xiang. GM(1,1) Timing Correction Model and Its Application[J]. Yunnan Environment Science, 2003, 22(4):20—21, 45.

[26] 李希灿, 李丽. 时序残差 GM(1,1)模型[J]. 系统工程与理论, 1998, 18(10):59—63.  
LI Xi-chan, LI Li. Time Residuals GM(1,1) Model[J]. System Engineering and Theory, 1998, 18(10):59—63.

勘误:《装备环境工程》2014年第11卷第6期第6页第2栏第16行“10~15天”更正为“20~30天”。