

DOI: 10.13476/j.cnki.nsbdtqk.2015.06.003

三种拟合评价法计算水文频率的比较

孟彩侠^{1,2}, 王平义^{1,2}, 张晓伟³, 喻涛^{1,2}

(1. 重庆交通大学 水利水运工程教育部重点实验室, 重庆 400074;

2. 重庆交通大学 河海学院, 重庆 400074; 3. 水利部 珠江水利委员会, 广州 510610)

摘要: 水文频率计算中优选参数的原则为: 理论频率曲线和经验点配合最好。介绍了 Excel、频率计算软件、集对分析三种拟合评价在水文频率计算中的应用, 分析了各种方式的优缺点。Excel 法、频率计算软件以离差平方和为量化准则优选参数, 但是该指标不灵敏。集对分析法根据联系数及 a, b, c 的数量综合分析水文频率曲线与实测点的拟合程度, 该法为适线拟合定量评价开辟了一条新途径, 一定程度上克服了常规目估适线法的主观任意性。

关键词: Excel; 频率计算软件; 集对分析; 水文频率

中图分类号: TV 121.4 **文献标志码:** A **文章编号:** 1672-1683(2015)06-1036-04

Comparison of three evaluation applications in hydrologic frequency computation

MENG Cai xia^{1,2}, WANG Ping yi^{1,2}, ZHANG Xiao wei³, YU Tao^{1,2}

(1. Key Laboratory of Hydraulic and Waterway Engineering of the Ministry of Education, Chongqing Jiaotong University, Chongqing 400074, China; 2. School of River & Ocean Engineering Chongqing Jiaotong University, Chongqing 400074, China; 3. Pearl River Water Resources Commission of the Ministry of Water Resources, Guangzhou 510610, China)

Abstract: The principle of parameter selections in the hydrologic frequency calculation is that the theoretical frequency curve matches with the empirical points. In this paper, three fitting evaluation applications and their advantages and disadvantages in the hydrologic frequency calculation are investigated, including Excel, frequency calculation software, and set pair analysis. The Excel method and frequency calculation software use the sum of deviation square as the parameter selection criteria but it is insensitive. According to connection number and the amount of a, b and c , the fitting degree between the frequency curve and measured points is analyzed by set pair analysis. This is a new way to evaluate the fitting of frequency curve quantitatively, which can avoid the subjective arbitrariness of current curve fitting methods.

Key words: Excel; frequency calculation software; set pair analysis; hydrologic frequency

水文频率计算的两个基本内容为分布线型与参数估计^[1]。分布线型的选择是一个不确定性问题, 中国规范推荐的常用线型是皮尔逊Ⅲ型(以下简称 P-Ⅲ型)。P-Ⅲ曲线需要估计的参数有三个: 均值 \bar{x} 、变差系数 C_v 、偏态系数 C_s , 这些参数的估计方法大体可以分为两类: 一类是参数统计法; 另一类是适线法, 其本质是一种优选参数的方法, 优选的原则是估计的频率分布曲线和样本经验点据分布配合最好(因而也叫配线法)。根据最优准则函数能否定量表示, 可分为目估配线法和准则配线法。常见优化准则函数有离差绝对值和准则(ABS)、离差平方和准则(OLS)等^[2]。耿鸿江^[3](2002)提出了用 Excel 求解 P-Ⅲ型分布频率计算中的离均系数数值、参数估计、推求设计值。王红芳等^[4](2007)将集对分析法(SPA)引入频率曲线拟合评估。李宏伟等^[5](2009)

用蚁群算法, 按照 OLS 准则对 P-Ⅲ型分布参数进行了优化计算。桑燕芳等^[6](2009)将模拟退火遗传算法(SAGA)与极大似然法(ML)相结合建立了 SAGA-ML 法, 该法具有较好的适用性。刘力等^[7](2009)提出了基于粒子群优化算法的优化适线法。王柏勇^[8](2010)采用矩法、概率权重矩法和双权函数法等三种方法估计 P-Ⅲ型曲线参数。许义和等^[9](2010)研发了基于 Matlab 的 P-Ⅲ型曲线绘制软件。梁忠民等^[10](2010)考虑参数和线型不确定性提出了基于贝叶斯理论的水文频率分析方法。孟彩侠等^[11](2011)研究了 Excel 在水文频率分析中的应用。冯平等^[12](2013)进行了基于混合分布的非一致性洪水频率分析计算。李扬等^[13](2013)采用高阶概率权重矩法, 估计不同阶 PWM 下的 GEV 分布参数, 并对相应理论频率曲线对经验点据的拟合情况进行评价。

收稿日期: 2015-01-19 修回日期: 2015-10-12 网络出版时间: 2015-11-30

网络出版地址: <http://www.cnki.net/kcms/detail/13.1334.TV.20151130.2012.016.html>

基金项目: 国家自然科学基金面上项目(51079165); 国家自然科学基金(51479015); 西部交通建设科技项目(2009328814012)

作者简介: 孟彩侠(1979-), 女, 陕西渭南人, 讲师, 博士, 主要从事港口航道及近海工程、水文学及水资源方面教学及科研工作。E-mail: 652823106@qq.com

1 资料与方法

1.1 资料

本文计算采用长江寸滩水文站 1954 年- 2008 年实测流量资料,采用“年最大值法”抽样,并且经过调查、考证,得到 1870 年(流量 100 000 m³/s),1905 年(流量 85 100 m³/s)两次历史特大洪水^[14],实测期有一项特大洪水,为 1981 年洪水(流量 85 700 m³/s)。

为了确保研究成果的实用性,在使用资料之前,先要对原始资料的可靠性、一致性、代表性进行审查^[1]。

(1) 可靠性审查:使用资料全部来源于 1949 年以后寸滩水文站实际观测资料,且水尺位置、零点高程、水准基面均未变动,所用数据资料是可靠的。

(2) 一致性审查:从气候条件、下垫面条件两个方面进行审查,气候条件基本可满足一致性要求;上游已建水利工程对水文站径流有一定影响、上游用水对水文站径流亦有一定影响,且 55 年期间,下垫面条件不断在改变,故而下垫面条件不能满足一致性要求。综合考虑资料不满足一致性要求,本文暂未对系列进行还原计算。

(3) 代表性审查:使用流量资料年数为 55 年,远大于要求连续实测数据最小年数 20,且包括了大、中、小等各种洪水年份,并有寸滩建站以来最大洪水(1981 年)、两次考证的历史特大洪水,说明资料具有较好的代表性。

1.2 方法

(1) 根据文献[11]利用 Excel 强大图表功能完成了海森几率格纸的绘制;利用软件内置函数求解 P- \hat{Q} 型频率曲线离均系数 Φ ;通过算例完成了 P- \hat{Q} 型理论频率曲线的绘制;优选统计参数。

(2) 选用武汉大学研发的“频率曲线 2004”软件,按照帮助文件说明,准备数据文件,完成适线,优选统计参数。

(3) 根据文献[15]原理,集对分析在水文频率曲线拟合定量评估的应用步骤如下。

a. 根据实测值利用适线法估计 P- \hat{Q} 型分布的统计参数,获得对应的水文频率曲线,继而可以获得与实测值同一频率上对应的曲线上的估计值。

b. 利用实测值与估计值构造集合 A, B₁, B₂, B₃, ...和集对。

c. 通过一定的分类标准,将 A 与 B₂, B₃, ...集合的元素

进行符号量化处理,计算 A 中各元素与中对应元素的相对误差,相对误差落在(0, 5%]内则对应元素符号为 \bar{N} , 相对误差落在(5%, 10%)内则 B_i 对应元素符号为 $\hat{0}$, 相对误差落在[10%, + ∞)内则 B_i 对应元素符号为 $\hat{0}$ 。相对误差计算公式为

$$e_{i,j} = \left| \frac{x_{0,j} - x_{i,j}}{x_{0,j}} \right| \quad (1)$$

式中: $e_{i,j}$ 表示 A 中第 j 个元素 $x_{0,j}$ 和 B_i 中第 j 个元素 $x_{i,j}$ 的相对误差。

为了描述拟合优度,将集合中各元素符号量化为 I, 则变为 A = (I, I, I, ..., I)。

d. 统计集对 H(B₁, A), H(B₂, A), ...中符号的相同、相异、相反个数,利用式(2)计算各集合联系度。

$$\mu_{X-Y} = \frac{S}{n} + \frac{F}{n}I + \frac{P}{n}J \quad (2)$$

记 $a = \frac{S}{n}$, $b = \frac{F}{n}$, $c = \frac{P}{n}$, 则式(2)可以写成

$$\mu_{X-Y} = a + bI + cJ \quad (3)$$

e. 根据求得的联系数及其 a, b, c 的数量综合分析水文频率曲线与实测点的拟合程度。

2 算例分析

本算例中调查考证期 N = 2008 - 1870 + 1 = 139; 实测期 n = 2008 - 1954 + 1 = 55, 特大洪水项数 a = 3, 实测期特大洪水项数 l = 1。考虑特大洪水以后,系列成为不连续系列,频率采用用《水利水电工程设计洪水计算规范》中不连续系列的统一处理法进行计算;统计参数 \bar{x} , C_v 的计算采用《水利水电工程设计洪水计算规范》中,不连续系列的矩法公式进行计算,得 $\bar{x} = 47767.12$, C_v = 0.238。

2.1 Excel 应用

海森几率格纸的绘制、离均系数的计算、理论频率曲线的绘制,按照文献[11]的介绍处理。目估适线法按照文献[1]步骤完成,为避免适线时修改参数的盲目性,参照“统计参数对频率的影响”进行。调整原则:主要调整 C_s(也即 C_s 和 C_v 的倍数), \bar{x} , C_v 也可做微小调整,频率曲线选配计算过程见表 1, 适线图见图 1。经过调整参数,适线,根据理论频率曲线与经验点的配合情况,认为第三次配线配合效果最好,故而选择 $\bar{x} = 47770$, C_v = 0.25, C_s = 6C_v = 1.5 为最终所选参数。

表 1 频率曲线计算选配表

Tab. 1 Fitting of frequency curve

频率 p (%)	第一次配线 $\bar{x} = 47767$, C _v = 0.238 C _s = 4.35 C _v = 1.02			第二次配线 $\bar{x} = 47770$, C _v = 0.24 C _s = 5 C _v = 1.2			第三次配线 $\bar{x} = 47770$, C _v = 0.25 C _s = 6 C _v = 1.5		
	Φ	K _p	x _p	Φ	K _p	x _p	Φ	K _p	x _p
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
0.1	4.59	2.09	99927	4.81	2.16	102966	5.23	2.31	110264
1	3.05	1.73	82422	3.15	1.76	83872	3.33	1.83	87537
10	1.34	1.32	63009	1.34	1.32	63134	1.33	1.33	63689
20	0.75	1.18	56325	0.73	1.18	56165	0.69	1.17	56013
30	0.37	1.09	52025	0.35	1.08	51753	0.30	1.07	51294
40	0.08	1.02	48685	0.05	1.01	48374	0.00	1.00	47778
50	-0.17	0.96	45831	-0.20	0.95	45530	-0.24	0.94	44901
60	-0.40	0.90	43229	-0.42	0.90	42976	-0.45	0.89	42400
70	-0.62	0.85	40711	-0.63	0.85	40551	-0.64	0.84	40115
80	-0.85	0.80	38100	-0.84	0.80	38095	-0.83	0.79	37913
90	-1.12	0.73	35039	-1.09	0.74	35316	-1.02	0.75	35609
95	-1.30	0.69	32962	-1.24	0.70	33516	-1.13	0.72	34264
99	-1.56	0.63	30031	-1.45	0.65	31151	-1.26	0.69	32767
99.9	-1.74	0.59	27963	-1.58	0.62	29689	-1.31	0.67	32090

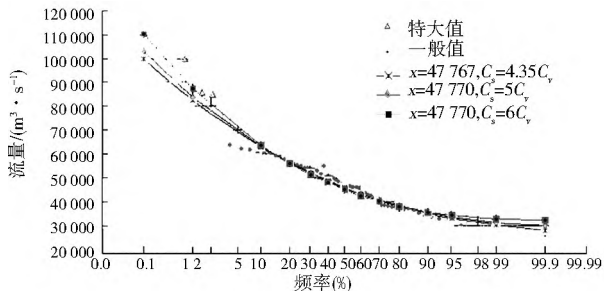


图 1 频率曲线

Fig. 1 Fitting of frequency curve

2.2 频率曲线分析软件应用

按照帮助文件说明,准备好数据文件,“选择项目文件”,“频率曲线视图”,就会出现如图 2 的频率曲线适线图,可以手动修改参数,也可点击“参数优化”等,选择计算机优化。点击频率成果选项,则可以产生频率成果记事本文件,成果包含统计参数 \bar{x} 、 C_v 、 C_s 及各频率对应的变量值。此频率计算软件“参数一”采用《水利水电工程设计洪水计算规范》中不连续系列相应频率、统计参数的计算公式。优选原则:离差平方和最小。经过几次调整参数,根据离差平方和最小原则,最终选定参数为: $x = 47\ 770$, $C_v = 0.25$, $C_s = 6C_v = 1.5$ 。

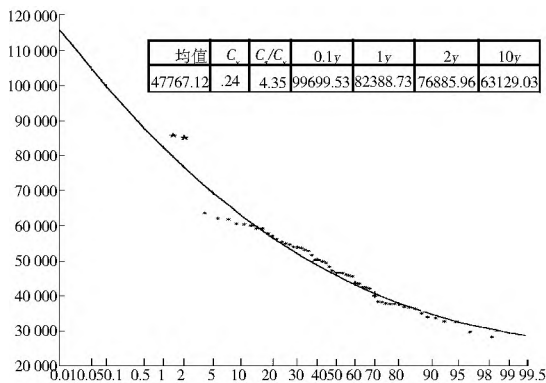


图 2 频率曲线

Fig. 2 Frequency curve

2.3 集对分析应用

矩法公式已得统计参数初值,用适线法调整参数 \bar{x} 、 C_v 、 C_s ,得到 3 组参数,见表 2。分别拟合年最大洪峰频率曲线,得到 3 个对应的集合 B_1 、 B_2 、 B_3 (根据调整的统计参数 \bar{x} 、 C_v 、 C_s ,根据实测值对应频率,利用 Excel 内置函数计算 P 0 型频率曲线离均系数 φ ,根据公式 $x = \bar{x}(1 + C_v\varphi)$ 得到估计值集合 B_1 、 B_2 、 B_3)。取分类数 $K = 3$,计算集对 $H(B_1, A)$ 、 $H(B_2, A)$ 、 $H(B_3, A)$ 的联系度,结果如下:

$$\begin{cases} \mu_{B_1-A} = 0.807 + 0.140I + 0.053J \\ \mu_{B_2-A} = 0.789 + 0.105I + 0.105J \\ \mu_{B_3-A} = 0.596 + 0.316I + 0.008J \end{cases}$$

取 $I = 0.5$, $J = -1$ 分别计算各集对的联系度,见表 2。为了对比起见,用同样洪水资料和各组参数拟合的估计值(也即集合 A 和相应集合 B_1 、 B_2 、 B_3 相应数据),计算了相应相关系数和离差平方和。结果同列于表 2。

以联系度最大为原则,则选取参数组 1 的结果($x = 47767$, $C_v = 0.238$, $C_s = 4.35C_v = 1.035$)为最终统计参数的取值。

表 2 统计参数及其拟合度指标

Tab. 2 The statistical parameters and indexes of fitting degree

组别	参数组 1	参数组 2	参数组 3	
$\bar{x}/(m^3 \cdot s^{-1})$	47767	47770	47770	
C_v	0.238	0.24	0.25	
C_s/C_v	4.35	5	6	
相关系数	0.9758	0.9758	0.9734	
离差平方和	581625212	560596055	557547427	
联系度	a	0.807	0.789	0.596
	b	0.14	0.105	0.316
	c	0.053	0.105	0.088
联系数	0.8246	0.7368	0.6667	

3 结果分析

用相关系数、离差平方和、联系数来评判均属于量化评估方式(准则适线)。以相关系数最大为原则,则参数组 1、参数组 2 为最佳;以离差平方和最小为原则,则参数组 3 为最佳。但是各方案相关系数,离差平方和变化均很小,作为推断指标均不灵敏。联系数在各方案中变化相对较大,该指标灵敏度较好。Excel 法、频率计算软件倾向于选参数组 3、集对分析法倾向于选取参数组 1。

(1) Excel 法。优点在于几条理论线与经验点绘制在同一张坐标纸上,目估理论线与经验点配合情况,得出哪条线拟合较好结论(也即统计参数 \bar{x} 、 C_v 、 C_s 取值)。缺点:目估适线具有因人而异性、任意性(有了 Excel 配合,也计算出其离差平方和,但是各组离差平方和较为接近,该指标不敏感)。

(2) 频率计算软件法。优点在于省时省力,快速得到拟合较好结论(统计参数取值),缺点:以离差平方和最小为优选原则,但是该指标不敏感。

(3) 集对分析法。该方法属于量化评估,优点在于以联系度式中 a 、 b 、 c 数值为基础分析和推断,可以考量拟合结构,更加合理。对于参数组 1,拟合程度较好的点据(同一性)份额,由 $a = 0.807$ 表示,清楚得说明:大约 80% 的点据拟合较好;拟合程度一般(差异性)份额,由 $b = 0.14$ 表示,表明大约 14% 的点据拟合一般;拟合较差点据(对立性),由 $c = 0.053$ 表示,表明大约 5% 点据拟合较差。无需绘制频率曲线适线图,亦可得到统计参数估计值。(如需要适线图时,将最终选定统计参数组绘制图形,仍可获得频率曲线图。)

4 结论

对比了 Excel、频率计算软件、集对分析三种拟合评价应用在水文频率计算中的优缺点,Excel 应用目估适线法理念,属于定性评估;频率计算软件、集对分析应用属于定量评估,然而集对分析法评价指标较频率计算软件评价指标灵敏,故集对分析方式结果更合理一些。实际应用时,可根据不同需求选取不同方式。

参考文献(References):

[1] 詹道江,徐向阳,陈元芳.工程水文学[M].北京:中国水利水电出版社,2011.(ZHAN Daojiang, XU Xiangyang, CHEN Yuanfang. Engineering Hydrology[M]. China Water Power

- Press, 2011. (in Chinese))
- [2] 邢贞相, 付强, 芮孝芳. 两种实用的洪灾损失频率分析方法[J]. 系统工程理论与实践, 2006(2): 127-132. (XING Zhenxiang, FU Qiang, RUI Xiaofang. Two practical methods on the analyses of inundation loss frequency[J]. Systems Engineering Theory & Practice, 2006(2): 127-132. (in Chinese))
- [3] 耿鸿江. Excel 在 P-0 型分布频率计算中的应用研究[J]. 水电能源科学, 2002, 20(3): 41-43. (GENG Hongjiang. Study and application of excel in P-0 distribution frequency calculation[J]. Water Resources and Power, 2002, 20(3): 41-43. (in Chinese))
- [4] 王红芳, 丁晶, 王文圣, 等. 集对分析在频率拟合度定量评价中的应用[J]. 水利水电技术, 2007, 38(4): 1-15. (WANG Hongfang, DING Jing, WANG Wensheng, et al. Application of spa to quantitative evaluation on fitting of frequency curve[J]. Water Resources and Hydropower Engineering, 2007, 38(4): 1-15. (in Chinese))
- [5] 李宏伟, 宋松柏. 蚁群算法在水文频率曲线参数计算中的应用[J]. 人民黄河, 2009, 31(4): 38-40. [LI Hongwei, SONG Songbai. Application of ant colony algorithms in parameter calculation of hydrological frequency calculation[J]. Yellow River, 2009, 31(4): 38-40. (in Chinese))
- [6] 桑燕芳, 王栋, 吴吉春. 水文频率分析中参数估计 SAGA-ML 方法的研究[J]. 水文, 2009, 29(5): 23-29. (SANG Yanfang, WANG Dong, WU Jichun. Research on saga-ml method for parameter optimization in hydrologic frequency analysis[J]. Journal of China Hydrology, 2009, 29(5): 23-29. (in Chinese))
- [7] 刘力, 周建中, 杨俊杰, 等. 粒子群优化适线法在水文频率分析中的应用[J]. 水文, 2009, 29(2): 21-23. (LIU Li, ZHOU Jianzhong, YANG Junjie, et al. Discussion on the basic concept of water resources carrying capacity[J]. Journal of China Hydrology, 2009, 29(2): 21-23. (in Chinese))
- [8] 王柏勇. 三种参数估计方法在降雨频率分析中的比较研究[J]. 中国防汛抗旱, 2010(5): 42-45. (WANG Boyong. Research on three methods of parameter estimation for rainfall frequency analysis[J]. China Flood & Drought Management, 2010(5): 42-45. (in Chinese))
- [9] 许义和, 魏晓妹. 基于 Matlab 的 P-0 型曲线绘制软件的研发与应用[J]. 水电能源科学, 2010, 28(7): 15-17. (XU Yihe, WEI Xiaomei. Development and application of p-0 distribution curve drawing software based on matlab[J]. Water Resources and Power, 2010, 28(7): 15-17. (in Chinese))
- [10] 梁忠民, 李磊, 王军, 等. 考虑参数和线型不确定性的水文设计值估计的贝叶斯方法[J]. 天津大学学报, 2010, 43(5): 379-384. (LIANG Zhongmin, LI Lei, WANG Jun, et al. Bayesian method for hydrological frequency analysis considering uncertainties of parameter and model[J]. Journal of Tianjin University, 2010, 43(5): 379-384. (in Chinese))
- [11] 孟彩侠, 徐芳, 张毅. Excel 在水文频率计算中的应用[C]. MACE 2011, 2922-2925. (MENG Caixia, XU Fang, ZHANG Yi. Application of excel in hydrologic frequency computation[C]. Second International Conference on Mechanic Automation and Control Engineering, 2011, 2922-2925. (in Chinese))
- [12] 冯平, 曾杭, 李新. 混合分布在非一致性洪水频率分析的应用[J]. 天津大学学报: 自然科学与工程技术版, 2013, 46(4): 298-303. (FENG Ping, ZENG Hang, LI Xin. Non-Stationary flood frequency analysis based on mixed distribution[J]. Journal of Tianjin University (Science and Technology) 2013, 46(4): 298-303. (in Chinese))
- [13] 李扬, 宋松柏. 高阶概率权重矩在洪水频率分析中的应用[J]. 水力发电学报, 2013, 32(2): 14-21. (LI Yang, SONG Songbai. Application of higher order probability-weighted moments to flood frequency analysis[J]. Journal of Hydroelectric Engineering, 2013, 32(2): 14-21. (in Chinese))
- [14] 中国水灾年表[EB/OL]. <http://www.chinacitywater.org/bbs/viewthread.php?tid=18719,2007-02-26>. (China flood chronological table[EB/OL]. <http://www.chinacitywater.org/bbs/viewthread.php?tid=18719,2007-02-26>. (in Chinese))
- [15] 王文圣, 张翔, 金菊良, 等. 水文学不确定性分析方法[M]. 北京: 科学出版社, 2011. (WANG Wensheng, ZHANG Xiang, JIN Jiliang, et al. Methods of Uncertainty Analysis for Hydrology[M]. Beijing: Science Press, 2011. (in Chinese))

(上接第 1035 页)

- [16] 国庆喜, 张海燕, 王兴昌, 等. 东北典型森林土壤呼吸的模拟—IBIS 模型的局域化应用[J]. 生态学报, 2010, 30(9): 2295-2303. (GUO Qingxi, ZHANG Haiyan, WANG Xingchang, et al. Simulation of soil respiration for typical forests in Northeast China: localized application of IBIS[J]. Acta Ecologica Sinica, 2010, 30(9): 2295-2303. (in Chinese))
- [17] 刘志理, 金光泽. 小兴安岭三种林型叶面积指数的估测[J]. 应用生态学报, 2012, 23(9): 2437-2444. (LIU Zhi li, JIN Guangze. Estimation of leaf area index of three forest types in Xiaoxing'an Mountains of Northeast China[J]. Chinese Journal of Applied Ecology, 2012, 23(9): 2437-2444. (in Chinese))
- [18] Yu R G, Wen X F, Sun X M, et al. Overview of ChinaFLUX and evaluation of its eddy covariance measurement. Agricultural and Forest Meteorology, 2006, 137: 125-137.
- [19] 李春, 何洪林, 刘敏, 等. ChinaFLUXCO₂ 通量数据处理系统与应用[J]. 地球信息科学, 2008, 10(5): 557-565. (LI Chun, HE Honglin, LIU Min, et al. The design and application of CO₂ flux data processing system at ChinaFLUX[J]. Geoinformation Science, 2008, 10(5): 557-565. (in Chinese))
- [20] 刘敏, 何洪林, 于贵瑞, 等. 中亚热带人工针叶林 CO₂ 通量组分统计不确定性分析[J]. 中国科学 D 辑: 地球科学, 2008, 38(8): 1016-1027. (LIU Min, HE Honglin, YU Guirui, et al. Uncertainty analysis of CO₂ flux components in subtropical evergreen coniferous plantation[J]. Science China D: Earth Sciences, 2008, 38(8): 1016-1027. (in Chinese))