

# 夏季室外高温皮尔逊—Ⅲ型分布模型的分析与确定\*

刘广海<sup>1,2</sup>, 谢如鹤<sup>2</sup>, 屈睿瑰<sup>3</sup>

(1. 中南大学 交通运输学院, 湖南长沙 410075; 2. 广州大学 物流与运输研究中心, 广东广州 510006; 3. 华东交通大学 土木建筑学院, 江西南昌 330013)

**摘要:**在分析、统计夏季室外高温数据的基础上,引入频率分析法,采用皮尔逊—Ⅲ型分布对夏季室外高温参数的极值分布进行了模拟。并以北京地区 1951~1981 之间 30 年气象参数为例,得到其不同保证率条件下的夏季高温分布图。研究使用分布检验了模型拟合效果,结果显示,上述模型能较好的反映真实气象状况。研究为不同标准层次 HVAC 系统设计气象参数的选取提供了有益补充,并可供相关研究参考。

**关键词:**室外高温参数;频率分析法;皮尔逊—Ⅲ型分布

**中图分类号:**TU831.2 **文献标志码:**A **文章编号:**1006-7329(2007)04-0107-04

## Study on Outdoor High Temperature Parameter by Pearson-Ⅲ Distribution Model

LIU Guang-hai<sup>1,2</sup>, XIE Ru-he<sup>2</sup>, QU Rui-gui<sup>3</sup>

(1. School of Traffic & Transportation Engineering, Central South University, Changsha 410075, P. R. China; 2. Research Center for Logistics & Transportation, Guangzhou University, Guangzhou 510006, P. R. China; 3. School of Civil Engineering and Architecture, East China Jiaotong University, Nanchang 330013, P. R. China)

**Abstract:**Based on analysis of the outdoor high temperature data from 1951 to 1981 in summer in Beijing, using frequency analytical method to imitate the high temperature data by Pearson-Ⅲ distribution, the high temperature distribution of Beijing has been gained, then the simulated effect is tested by distribution. This paper provides available complementarities for the selection of weather parameter for different HVAC design standard, and can be consulted by correlated research.

**Keywords:** outdoor high temperature parameter; frequency analytical method; Pearson-Ⅲ distribution

长期以来,我国空调建筑设计、能耗分析所使用的室外气象参数均根据各地的气象资料分析统计并考虑平均不保证时间来确定的<sup>[1,2]</sup>,对极端气象数据研究较少。然而,气象研究表明,因大气环流异常、厄尔尼诺现象、拉尼娜现象等因素的影响,约有 60% 的年份与标准年的气象参数相差 20% 甚至 50% 以上,持续高温频频出现,持续的全球变暖更加大了理论气象参数与实际气候条件的差异。2003 年,在全球极端高温一再创造纪录的同时,亦夺去了数以万计的生命;2006 年夏季,重庆地区 35℃ 以上高温天数超过 39 d, 37℃ 以上高温天数超过 25 d, 8 月 15 日最高气温达到 44.5℃,均居有气象记录以来同期第一位,给人民生活带来严重影响。随着人民生活水平的不断提高,建筑格局和功能的日趋多样化,国内规范只给出了 1 个不保证时间的设计气象参数已不能完全满足不同标准层

次 HVAC 系统设计的需要(如医院、住宅、恒温室等)。为此,ASHRAE 在新的设计气象参数构成上,对极端气象参数进行了详细的统计和分析,并给出了若干不保证率的设计条件<sup>[3~5]</sup>,从而较好的解决了上述问题。针对我国,研究者拟以北京地区夏季室外高温数据为例,通过频率分析法,最终获得该地区各种不保证率条件下的相关计算温度,从而为不同类型建筑空调设计提供参考。

### 1 频率分析法的引入

#### 1.1 频率分析法

所谓频率分析法是指在对事物规划设计和管理工作用时,通过分析已有事物特征值的相关规律,给出该事物具有机率含义的特征值。在本研究中,频率分析的目的在于提供具有不同机率含义的夏季极高温值。为

\* 收稿日期:2007-01-22

基金项目: EASTS 国际研究团队合作项目(IRG-08-2005); 国家自然科学基金资助项目(50378091)

作者简介: 刘广海(1978-), 男, 湖南邵阳人, 博士生, 主要从事建筑能耗分析, 冷藏运输等相关领域的研究。

使频率分析的成果有较为可靠的基础,必先做好气温系列样本的审查工作:详细审查它们的一致性、代表性、可靠性和独立性,在有了审定后的高温系列样本后,即可以开展频率分析和计算工作。

### 1.2 重现期和保证率的确定

重现期指许多次试验里某一事件重复出现的时间间隔的平均数,也就是平均的重现间隔期。在对夏季高温的研究中,频率  $P$  与重现期  $T$  的关系可表示为:

$$T = 100/P \quad (1)$$

式中:  $T$  为重现期,年;  $P$  为频率, %。

因此,如对重现期分别为十年一遇,二十年一遇,五十年一遇和百年一遇的高温情况进行分析,则相应的  $T$  值分别为:  $T=10(P=10)$ 、 $T=20(P=5)$ 、 $T=50(P=2)$  和  $T=100(P=1)$ 。值得特别说明的是,所谓重现期为百年一遇并不是说正好一百年出现一次,事实上也许一百年出现数次,或是一次也未出现。它的含义仅表示每个年份出现的可能程度为 1%,只有在较长时间内,才能体现出其规律性。

保证率则指的是保证安全的概率。在本研究中,设计频率  $P$  表示每年会出现超过夏季室外干球设计温度的机率,按保证率的计算公式,在年内均出现不超过该值的机率为  $(1-P)^n$ ,因此保证率为:

$$P_b = (1-P)^n \quad (2)$$

式中:  $P_b$  为保证率, %;  $n$  为考虑保证率的年数,年。

对于不同的空调系统而言,其使用期一般间于 5~30 年之间。对于不同重现期的极端高温值,由上式可计算其相应的保证率,如图 1 所示。由图中数据亦可看出,即便是百年一遇的极端高温,空调设备仍有 4.9%~26% 的可能会遇到,这也进一步验证了本研究的必要性。

### 1.3 选样方法的确定

选样指从现有记录中合理地选出若干个数值以组成一个样本来作为频率分析的依据。因此,在选样时要求样本中每个单元具有一致性和独立性,所组成的样本还必须富有代表性和足够的可靠性。频率分析中常用的选样法有年最大值法、年超大值法和超定量法。

**1.3.1 年最大值法** 每年中仅选取一个最大值的方法称为年最大值法。这种方法非常简单且易于选取,在我国暴雨和洪水的统计分析中多采用此法。同时,许多学者的研究也指出相邻年份中年最大值的关系十分微弱,可以认为是独立的。用年最大值法选样,在  $n$  年的观测资料中能选出  $n$  个最大值。

**1.3.2 年超大值法** 在  $n$  年观测资料中不逐年选取最大值,而是按数值的大小依次选取  $n$  个最大值,这种方法称为年超大值法。由该法选取的样本往往有些年

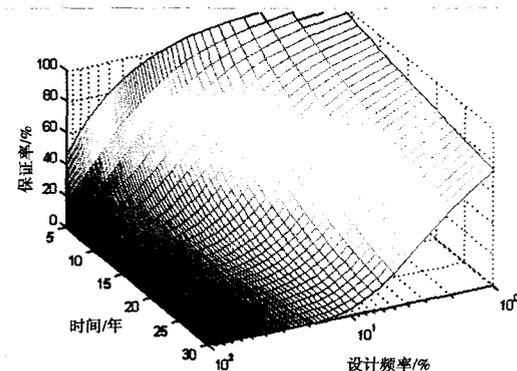


图 1 保证率  $P_b$  / %

份没有选到,而个别年份中则有两个或两个以上。这样,可能在一年中连取数个最大值,其独立性较差。同时,采用该方法在选用时比年最大法费时,且在成果的主要部分两法一般都比较接近。但是在大型水利建设中此法较为合适。

### 1.3.3 超定量法

凡每年中超过一定量的数值均选取的方法称为超定量法。超定量法的选择非常麻烦且相互独立性较差,特别是在观测年份短时就缺乏代表性。同时,选取定量的最小限值亦有主观成分。一般,超定量法主要应用于洪水的分析上。

选样工作是频率分析中重要的一环,如果选择不合理就会影响到成果的代表性和可靠性,而具体的选样方法则应根据具体研究目的来定,对于此类研究,使用年最大法最为合适<sup>[6]</sup>,表 1 列出了通过此方法获取的北京市 1951~1980 年夏季高温的样本值。

表 1 北京市逐年夏季高温值(1951~1980)

年份	气温/℃	年份	气温/℃	年份	气温/℃
1951	38.3	1961	40.6	1971	35.2
1952	39.6	1962	37.1	1972	39.5
1953	37.0	1963	39.0	1973	37.5
1954	33.4	1964	37.5	1974	35.8
1955	39.6	1965	38.5	1975	38.4
1956	34.6	1966	37.5	1976	35.0
1957	36.2	1967	35.8	1977	33.9
1958	37.6	1968	40.1	1978	37.5
1959	36.8	1969	35.9	1979	35.9
1960	38.1	1970	35.3	1980	35.1

## 2 皮尔逊-Ⅲ型分布模型的引入

就目前而言,国内外在极端气象参数的分析中,公认较为出色的有皮尔逊-Ⅲ型分布模型<sup>[7]</sup>,并在暴雨、冰雹等多种气象灾害预测中得到应用。其概率密度分布曲线通用式可表示为:

$$dy/dx = (x+d)y/(b_0 + b_1x + b_2x^2) \quad (3)$$

式中:  $d$  为偏差半径;  $b_0$ 、 $b_1$ 、 $b_2$  等为待定参数,可以用资料中的各阶矩来代替。

III型曲线的判别准则为  $b_2=0$ ,就数学上解释,在  $b_0+b_1x+b_2x^2=0$  中,若  $b_2 \rightarrow 0$ ,则其中一根趋于无穷大,而另一根为  $-b_0/b_1$ 。此时,(3)式可变为:

$$dy/dx = (x+d)y/(b_0+b_1x) \quad (4)$$

将坐标轴由均值移至众值处,以  $x-d$  代替上式中的  $x$ ,则有:

$$dy/y = xdx/b_1(x+b_0/b_1-d) \quad (5)$$

令  $a=b_0/b_1-d$  且由 III 型分布曲线性质积分可得:

$$y = y_0(1+x/a)^{a/d}e^{-x/d} \quad (6)$$

在此将曲线坐标原点由众值处移至曲线起点处,同时令  $\alpha = \frac{a}{d} + 1, \beta = \frac{1}{d}$ ,则:

$$y = \beta^\alpha x^{\alpha-1} e^{-\beta x} / \Gamma(\alpha) \quad (7)$$

如果再将(7)式的坐标原点移至极端气温系列的实际零点,则在本研究中皮尔逊-III型密度分布曲线最终可表示为:

$$y = \beta^\alpha (x-x_0)^{\alpha-1} e^{-\beta(x-x_0)} / \Gamma(\alpha) \quad (8)$$

式中: $x_0$  为 III 型分布曲线的起点与系列零点的距离,即随机变量  $X$  所能取到的最小值; $\alpha$  为形状参数; $\beta$  为尺度参数; $\Gamma(\alpha)$  为  $\alpha$  的伽马函数。

由矩法原则,三参数可分别用下式计算<sup>[8]</sup>:

$$\alpha = 4/c_s^2 \quad (9)$$

$$\beta = 2/\sigma \quad (10)$$

$$x_0 = m(1-2c_v/c_s) \quad (11)$$

式中: $m$  为数学期望; $\sigma$  为均方差; $c_s$  为偏态系数; $c_v$  为变差系数。

这些数字特征的估计量分别为:

$$\hat{m} = \bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (12)$$

$$\hat{c}_s = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3 / \left( \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^{3/2} \quad (13)$$

$$\hat{\sigma} = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \quad (14)$$

$$\hat{c}_v = \hat{\sigma} / \hat{m} \quad (15)$$

以上各统计量中,偏态系数  $\hat{c}_s$  含有三阶样本矩,故抽样误差较大,样本实测值  $\hat{c}_s$  与真值  $c_s$  之间可能有比较大的差异,常需要对拟合的线型进行验证,并对  $\hat{c}_s, \hat{c}_v$  进行适当的调整,以获得理想的分布曲线。其中一个主要原则为:

$$\hat{c}_s \geq 2\hat{c}_v \quad (16)$$

但是也应看到,虽然(16)式为  $\hat{c}_s$  的调整提供了一定的指导,但其具体值的确定仍在较大程度上依赖于试根和主观评价,就目前而言并无其他好的办法,这也是该方法的一大遗憾<sup>[9,10]</sup>。

### 3 北京地区夏季室外高温值的确定

求对应于指定频率  $P$  的数值  $x_p$ ,可从(17)式中求解  $x_p$  值:

$$P = P(X \geq x_p) = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_{x_p}^{\infty} (x-x_0)^{\alpha-1} e^{-\beta(x-x_0)} dx \quad (17)$$

令  $t = \beta(x-x_0)$ ,得:

$$P = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_{x_p}^{\infty} t^{\alpha-1} e^{-t} dt \quad (18)$$

式中: $t_p = \beta(x_p - x_0)$ 。

进一步推得:

$$x_p = \frac{\bar{x}c_s c_v}{2} t_p + \bar{x} - \frac{2\bar{x}c_v}{c_s} \quad (19)$$

$$\text{则: } \frac{x_p - \bar{x}}{\bar{x}} = \frac{c_s c_v}{2} t_p - \frac{2c_v}{c_s} \quad (20)$$

上式  $c_s$  将移项后,即化为标准化形式:

$$\Phi = \frac{c_s}{2} t_p - \frac{2}{c_s} \quad (21)$$

式中: $\Phi$  为离差系数。

当已知  $P$  及  $c_s$  后,  $t_p$  即可获得,则  $x_p$  可从(22)式求得:

$$x_p = (\Phi c_s + 1)\bar{x} \quad (22)$$

因此,在已知  $\bar{x}, c_v$  及  $c_s$  的条件下,频率分布曲线即确定下来。

在本研究中,以表 1 中北京市逐年夏季高温值为样本,按式(12)、(14)和(15)分别求得:

$$\bar{x} = 37.077; \hat{c}_v = 0.0499; \hat{c}_s \approx 0$$

显然,上述结果不满足  $\hat{c}_s \geq 2\hat{c}_v$  的条件,因此,由  $\hat{c}_v = 0.0499$  将  $\hat{c}_s$  的值调整为 0.1 以上,再以此为根据,查皮尔逊-III型曲线的离差系数表,进而求得  $x_p$  的对应值。为获得较好的频率曲线,本文分别对  $\hat{c}_s = 0, \hat{c}_s = 0.1, \hat{c}_s = 0.5, \hat{c}_s = 1.0$  和  $\hat{c}_s = 1.5$  的情况进行了研究,作出北京市极端高温的频率分布曲线,如图 2 所示。从直观上,由图可粗略地判断,将  $\hat{c}_s$  的值调整为 0.5 时拟合效果较好。由图亦可算出不同重现期的极

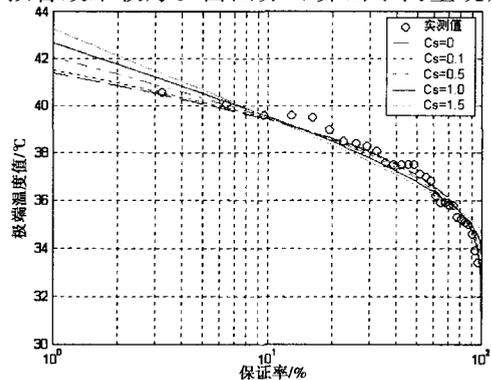


图 2 北京市夏季高温的频率分布曲线图

端高温值,如表 2 所示。可见,北京地区十年一遇,二十年一遇,五十年一遇和百年一遇的高温值分别为 39.52 °C、40.35 °C、41.35 °C 和 42.04 °C,其他地区的高温值亦可采用同样方法来确定。

·表 2 皮尔逊-Ⅲ型分布频率 P 与  $x_p$  的对应值( $c_3=0.5$ )

P/%	0.001	0.01	0.1	0.2	0.333	0.5	1	2	3	5	10	20	25
$x_p$	47.77	46.01	44.13	43.52	43.09	42.70	42.04	41.35	40.93	40.35	39.52	38.58	38.22
P/%	30	40	50	60	70	75	80	85	90	95	97	99	99.9
$x_p$	37.93	37.39	36.93	36.47	36.00	35.76	35.50	35.19	34.82	34.32	34.01	33.45	32.64

表 3 皮尔逊-Ⅲ型分布  $\chi^2$  检验

$A_i$	$f_i$	$\hat{p}_i$	$n\hat{p}_i$	$f_i - n\hat{p}_i$	$(f_i - n\hat{p}_i)^2 / n\hat{p}_i$
$A_1: x < 34$	2	0.029 9	0.895 7		
$A_2: 34 \leq x < 35$	1	0.092 4	2.770 9		
$A_3: 35 \leq x < 36$	8	0.176 5	5.294 5		
$A_4: 36 \leq x < 37$	2	0.217 8	6.535 1	-4.535 1	3.147 1
$A_5: 37 \leq x < 38$	7	0.196 4	5.891 6	1.108 4	0.208 5
$A_6: 38 \leq x < 39$	4	0.138 6	4.157 0	-0.157 0	0.005 9
$A_7: 39 \leq x < 40$	4	0.080 7	2.420 6		
$A_8: 40 \leq x < 41$	2	0.040 0	1.200 3		
$A_9: 41 \leq x < 42$	0	0.017 3	0.518 3	1.544 9	0.535 7
$A_{10}: x \geq 42$	0	0.010 5	0.316 0		
$\Sigma$					4.361 2

#### 4 模型的验证

在上文中,目测经验分布函数与理论曲线差别的方法比较快捷直观,但人为因素较大,尤其当样本容量不够大时,一般仅用在适线调整上。因此,用样本经验分布拟合出北京市夏季高温理论曲线后,该随机变量实际总体分布是否符合所选理论分布模型,需要以客观方法判定。在此引入  $\chi^2$  检验法,用以检验上述分布对夏季高温的拟合效果。

在检验极端气温分布情况前,首先需对样本数据分区。由表 1 可知,在原始数据中,最小值、最大值分别为 33.4 °C 和 40.6 °C,即所有样本数据均落在区间 [33.4, 40.6] 之间,现取  $\Delta=1$  °C,将样本数据等分为  $x < 34, 34 \leq x < 35, \dots, 41 \leq x < 42, x \geq 42$  等 10 个区间并确定各区间数据的频数,进而在确定  $\hat{p}_i, n\hat{p}_i$  等值得到其  $\chi^2$  值。

对于皮尔逊-Ⅲ型分布,按上述方法,如取  $n\hat{p}$ ,由于相关被估计参数为 3,此种做法将使其自由度  $k-r-1=0$ ,从而无法检验其优越性。为此,研究者在合并  $A_i$  时,适当的放宽要求,使最小  $n\hat{p}_i \approx 5$  的,从而使其自由度  $k-r-1=1$ 。此时,因为:

$$\begin{aligned} \chi_{0.01}^2(k-r-1) &= \chi_{0.01}^2(5-3-1) \\ &= \chi_{0.01}^2(1) = 6.635 > 4.361 2 \end{aligned}$$

故认为在水平 0.01 下接受关于总体分布的假设(如表 3),即夏季高温数据符合皮尔逊-Ⅲ型分布。

#### 5 结语

气象参数的合理选择是 HVAC 设计的前提与基

础。本文以北京地区为例,采用皮尔逊-Ⅲ型分布获得其不同保证率条件下夏季室外高温分布状况,其他地区亦可同理确定。研究为不同标准层次 HVAC 系统设计气象参数的选取提供了有益补充,并可供相关研究参考。

#### 参考文献:

- [1] GB 50019-2003. 采暖通风与空气调节设计规范[S].
- [2] 苏芬仙,张从军,田胜元. BIN 建筑能耗计算方法的改进[J]. 重庆建筑大学学报,2006,28(1):88-91.
- [3] ASHRAE. Climatic design information[S]. U. S. A. :1997 ASHRAE Handbook Fundamentals,Chapter 26. 1997.
- [4] HARRIMAN III L G, COLLIVER D G, HART K Q. New weather data for energy calculations[J]. ASHRAE Journal,1999,41(3):31-38.
- [5] COLLIVER D G, GATES R S. Development of the design climatic data for the 1997 ASHRAE handbook - fundamentals[J]. ASHRAE Trans. 2000,106(1):3-14.
- [6] 马开玉. 气候统计原理与方法[M]. 北京:气象出版社,1993.
- [7] 么枕生. 气候统计[M]. 北京:气象出版社,1991.
- [8] 马秀峰,阮本清. 对权函数法求概率分布参数讨论[J]. 水利学报. 2001,46(11):16-19.
- [9] 邱兆富,曾晓岚,郝以琼,等. 重庆主城暴雨强度公式推求[J]. 重庆建筑大学学报,1999,21(6):82-84.
- [10] 刘广海. 气候异常条件下空调建筑热性能及能耗状况模拟与分析[D]. 长沙:中南大学,2004.

(编辑 陈 蓉)