

文章编号:1671-6833(2003)02-0067-03

Excel 及其 VBA 在水文统计参数优选中的应用

马细霞¹, 王 丰¹, 马巧花²

(1. 郑州大学环境与水利学院, 河南 郑州 450002; 2. 河南宇疆边贸科技应用总公司, 河南 郑州 450003)

摘 要: 水文统计参数优选方法中, 最优化方法要求使用者具备一定的系统工程基础和编程技巧, 而图解适线法计算工作量大, 计算结果因人而异. 在分析以上两类方法的基础上, 针对以上问题, 基于 Excel 的内置函数、表格及图形处理功能, 运用 Excel 的 VBA 编写了简单的优选程序, 实现了参数优选的简单化和图形输出的规范化. 实例分析结果证明了该方法的可行性.

关键词: 统计参数; P-Ⅲ型分布曲线; 优选; Excel 函数

中图分类号: TV 122 **文献标识码:** A

0 引言

水文频率计算是水文分析与计算的主要方法, 在生产实践中应用广泛. 若频率曲线的线型一定(通常采用皮尔逊Ⅲ型分布曲线), 则水文频率计算实质上就是优选频率曲线的统计参数. 目前, 水文统计参数优选方法可以分为两大类: 最优化方法和图解适线法. 最优化方法如单纯形法^[1]、遗传算法^[2]等不需试算, 计算精度较高, 但要求使用者具备一定的系统工程基础和编程技巧, 不便于初学者和基层工作者掌握和使用; 图解适线法过程明了, 便于掌握, 但适线过程繁琐, 计算工作量大, 适线结果因人而异. 本文基于 Excel 的内置函数、表格及图形处理功能, 运用 Excel 的 VBA 编写简单的优选程序, 实现参数优选的简单化和图形输出的规范化, 以补充和完善现有的水文统计参数优选方法.

1 水文样本系列统计参数优选模型

设某一随机变量 X 服从皮尔逊Ⅲ型分布, 它的 3 个统计参数分别为均值 \bar{x} 、离势系数 c_v 和偏态系数 c_s . 因总体未知, 故 3 个参数需要根据样本来估计, 通常采用与经验频率点配合最佳的皮尔逊Ⅲ型分布曲线所对应的 3 个参数的值. 因此, “配合最佳”常作为确定统计参数的适线准则. 在量值上, 一般取经验频率点距与相同频率的皮尔

逊Ⅲ型分布曲线的纵坐标之差——离差的组合值最小作为准则. 本文采用离差平方和最小准则, 其目标函数为

$$\min \left[\sum_{i=1}^n (x_i - x_{p_i})^2 \right] \quad (1)$$

式中: x_i 为样本系列中第 i 项的取值; p_i 为 x_i 的经验频率, 可根据样本的特性, 按连序系列或不连序系列公式计算^[3]; x_{p_i} 为皮尔逊Ⅲ型分布曲线上与经验频率 p_i 相应的纵坐标值.

2 基于 Excel 的水文统计参数优选

2.1 水文统计参数优选方法

本文依据图解适线法的思路, 采用逐步迭代法, 通过增加步长不断改变统计参数各种取值的组合, 并计算各种取值组合下的离差平方和, 从中取出最小值, 则与该最小值相应的统计参数即为所求.

2.2 水文统计参数优选步骤

2.2.1 计算经验频率及其相应的皮尔逊Ⅲ型分布曲线上的 x_p 值

打开 Excel 工作表, 按以下步骤操作:

(1) 计算经验频率. 在 Excel 工作表的第 1 列, 输入样本系列值 x_i , 并按由大到小排序; 在第 2 列计算样本各项的经验频率 $p_i, i = 1, 2, \dots, n$. 为便于在海森几率格纸上绘图, 在第 3 列计算经验频率 p_i 在海森几率格纸上的横坐标值 l_i , 其计算公式为

收稿日期: 2003-01-02; 修订日期: 2003-02-28

基金项目: 河南省自然科学基金资助项目(0211061700); 河南省高校青年骨干教师资助项目

作者简介: 马细霞(1963-), 女, 河南省叶县人, 郑州大学副教授, 硕士, 主要从事工程水文及水资源系统分析方面的

$$l_i = \text{NORMSINV}(p_i) - \text{NORMSINV}(0.01\%) \quad (2)$$

(2) 初估统计参数. 在 Excel 工作表的第 4、5 列, 分别计算样本系列的模比系数 K_i 和 $(K_i - 1)^2$; 选取两个单元格(比如 C35 和 C36), 利用 Excel 的统计功能, 根据样本系列资料情况, 按连序系列或不连序系列的矩法估计值公式^[4] 分别计算样本的均值 \bar{x} , 离势系数 c_v , 并根据研究对象的特点, 选取 c_s 值, 可取 $c_s = 2c_v$, 填入另一单元格中(比如 C37)。

(3) 计算皮尔逊Ⅲ型曲线上与经验频率相应的 x_p 值. 在工作表的第 6、7 列按以下公式计算离均系数 Φ_p 及 x_p 的值:

$$\Phi_{p_i} = \frac{c_s}{2} \text{GAMMAINV}\left(1 - p_i, \frac{4}{c_s^2}, 1\right) - \frac{2}{c_s} \quad (3)$$

$$x_{p_i} = \bar{x} (1 + \Phi_{p_i} c_v), (i = 1, 2, \dots, n) \quad (4)$$

(4) 计算离差平方和. 在工作表的第 8 列, 计算同一经验频率 p_i 下, 经验频率点纵坐标与皮尔逊Ⅲ型曲线上纵坐标差值的平方, 即 $(x_i - x_{p_i})^2$, 并统计离差平方和 $\sum_{i=1}^n (x_i - x_{p_i})^2$, 其结果放在单元格中(比如 I34)。

2.2.2 利用 Excel 的 VBA, 实现统计参数优选

离差平方和 $\sum_{i=1}^n (x_i - x_{p_i})^2$ 是统计参数 \bar{x}, c_v, c_s 的函数, 因此可通过步长逐步调整 \bar{x}, c_v, c_s 的值(主要调整 c_v, c_s 值), 来寻求目标函数的最小值。

在工作表中, 插入命令按钮, 如 CommandButton1, 并编写如下简单脚本:

Private Sub CommandButton1_Click()

Cv0=Range("C36")

Cv=Cv0

Cs=Range("C37")

Cs0=Cs

X=Range("I34")

Do While Cv<Cvmax

Do While Cs<Csmax

Cs=Cs+0.05

Range("C37")=Cs

f0=Range("I34")

If x>f0 Then

x=f0

Cvmin=Cv

Csmin=Cs

End If

Loop

Cv=Cv+0.01

Range("C36")=Cv

Cs=Cs0

Range("C37")=Cs

f0=Range("I34")

If x>f0 Then

x=f0

Cvmin=Cv

Csmin=Cs

End If

Loop

Range("C36")=Cvmin

Range("C37")=Csmin

Text Box1.Text="Cv="&Str(Cvmin)+"

"&Cs="&Str(Csmin)+" "&x="&

Str(Range("C35"))

End Sub

单击该命令按钮, 可以自动优选出随机变量 X 的三个统计参数。

3 水文统计参数优选结果的输出

水文频率计算结果是给出优选的统计参数, 并绘出该组参数相应的皮尔逊Ⅲ型分布曲线. 在 Excel 工作表中, 具体操作如下:

(1) 绘制海森几率格纸. 海森几率格纸的绘制方法参见文献^[4]。

(2) 点绘经验频率点. 选取海森几率格纸, 以表中第 1 列 x_i 、第 3 列 l_i 为添加数据的数据源, 选择图表类型“散点图”中的散点子图, 即可在海森几率格纸上点绘出经验频率点。

(3) 绘制皮尔逊Ⅲ型分布曲线. 选取海森几率格纸, 以表中第 7 列 x_p 、第 3 列 l_i 为添加系列的数据源, 选择图表类型“散点图”中的折线子图, 即可绘出最优统计参数 \bar{x}, c_v 和 c_s 相应的皮尔逊Ⅲ型分布曲线。

根据步骤 1~3 所绘出的水文频率计算结果(频率曲线图及相应的统计参数), 可根据需要在 A3、A4 或 B5 等不同规格打印纸上, 以横向或纵向方式, 任选比例地打印输出。

4 实例分析

某水库坝址处有 1950~1979 年 30 年实测洪水资料^[3]. 历史上调查到 1880 年曾发生一次特大洪水, 洪峰流量为 $9\,200\text{ m}^3/\text{s}$, 为 1880~1979 年的最大一次洪水. 此外, 1962 年实测大洪水的洪峰

流量为 $7\,100\text{ m}^3/\text{s}$ ，仅次于 1880 年的大洪水，且 1880~1950 年间再没有发生大于 $7\,000\text{ m}^3/\text{s}$ 的洪水。

现采用本文介绍的方法对该水库坝址处的洪峰流量进行频率计算，并对其统计参数进行优选，计算结果见图 1。

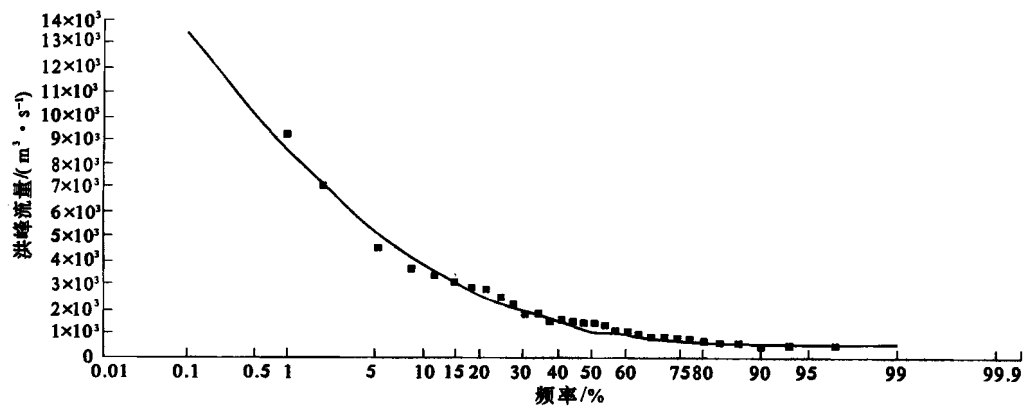


图 1 某水库坝址处洪峰流量频率曲线
Fig. 1 Frequency curve of flood peak

本文方法优选出的统计参数($\bar{x}=1\,718\text{ m}^3/\text{s}$ ， $c_v=1.0$ ， $c_s=2.75$)与文献[4]中图解适线法计算结果($\bar{x}=1\,718\text{ m}^3/\text{s}$ ， $c_v=1.0$ ， $c_s=3$)非常接近，但其精度更高，且计算过程具有通用性。

行，修改经验频率及统计参数矩法计算式等，所以可以大大减少计算工作量。
因此，本文优选水文统计参数的方法具有较好的应用前景。

5 结论

- (1) 本文利用 Excel 的内置函数、VBA 及表格处理和绘图功能，在电子表格中进行了水文统计参数的优选，不需查表，编程简单，易于操作，特别适合初学者和基层工作者使用。
- (2) 利用本文方法优选不同水文随机变量的统计参数时，仅需根据样本情况(样本容量大小，样本中的最大值，样本为连序系列或不连序系列等)，对原 Excel 表稍作修改，如插入或删除某些

参考文献：

[1] 谢平, 郑泽权. 水文频率计算有约束加权适线法[J]. 武汉水利电力大学学报, 2000, 33(1): 49~52.
[2] 姜铁兵, 康玲, 沈同林, 等. GA 适线法软件设计在水文频率计算中的应用[J]. 华中理工大学学报, 1999, 27(5): 52~53.
[3] 王燕生. 工程水文学[M]. 北京: 水利电力出版社, 1995.
[4] 林莺, 李世才. 水文频率曲线简捷计算和绘图技巧[J]. 水利水电技术, 2002, 33(7): 52~53.

Application of VBA in the Optimal Selection of Hydrologic Statistic Parameters

MA Xi-xia¹, WANG Feng¹, MA Qiao-hua²

(1. College of Environmental & Hydraulic Engineering, Zhengzhou University, Zhengzhou 450002, China; 2. Yujiang Frontier Trade Technical Application Company of Henan, Zhengzhou 450003, China)

Abstract : Among the optimum seeking methods of hydrologic statistic parameters ,the optimal methods need system engineering and programming technique . The graphic method needs a lot of time for calculating and the results are always different . Aiming at the disadvantage of those methods ,a simple optimal program is presented by using VBA of Excel , based on the form and graph function of Excel . It realizes the simplification of the optimal method and standardization of the output graph . The result of the example has verified its feasibility .
Key words : statistic parameter ; frequency calculation ; optimal selection ; Excel function