

土壤侵蚀定量研究中的统计分析

江忠善 刘 志

(中国科学院西北水土保持研究所, 陕西杨陵)

摘 要

本文主要根据我们在土壤侵蚀定量研究中使用统计分析方法的体会, 就土壤侵蚀定量研究中的统计分析和应注意的问题作一初步论述, 旨在引起土壤侵蚀学界的讨论和重视。

关键词: 土壤侵蚀 统计分析 定量评价

随着科学技术的发展, 特别是数学方法的发展及计算机应用为土壤侵蚀定量研究提供的有利条件, 加之将土壤侵蚀的定性研究与定量研究有机地结合, 推动了土壤侵蚀研究的进展。在定量研究中多数问题是难以通过精确的数学方法进行, 因为影响土壤侵蚀的因素错综复杂, 并具有多元性和变异性, 这在一定程度上限制了精确数学方法的应用, 因此必须借助数学统计分析方法来进行解决。下面就土壤侵蚀定量研究中的统计分析和应注意的几个问题作一介绍和讨论。

1 概率法

在土壤侵蚀定量研究中, 为研究各因子在土壤侵蚀发生事件中出现的机率, 通常以它们的出现的概率来统计分析, 并以此来说明土壤侵蚀量出现的概率与各影响因子出现概率的区间及其分布特性, 进而寻出相互联系的规律。如分析土壤侵蚀量与降雨量在同期内出现的概率情况必须借助概率统计分析方法来进行, 通过长系列的资料统计分析可确定它们出现概率大小、出现区间和分布特征, 这为进一步土壤侵蚀定量研究提供了基本前提。土壤侵蚀定量研究中应用概率统计分析的例子很多, 如面蚀在不同坡度上出现的频率(图1), 不同坡度坡面上的浅沟分布频率(图2), 等等。

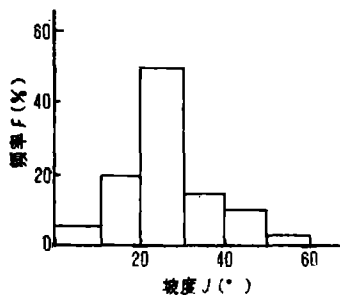


图1 面蚀出现的频率与坡度的关系

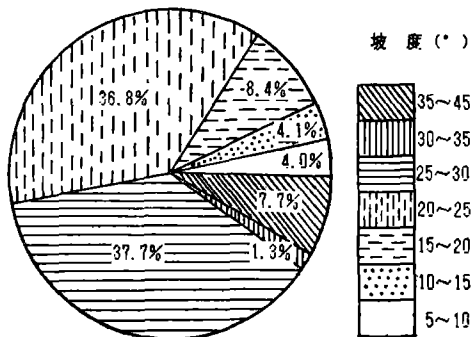


图2 不同坡度坡面上的浅沟分布频率

2 数理统计法

数理统计法在土壤侵蚀定量研究中应用广泛。数理统计是从实际资料出发来研究随机现象的规律性,它侧重于用随机现象本身的规律性来考虑资料的收集、整理和分析,从而找出相应的随机变量的分布规律或它的数字特征。数理统计的内容随着科学的不断发展而不断地扩大,但概括地说可以分为两个问题:(1)试验或调查的设计。即研究如何更合理更有效地获得观察资料的方法;(2)统计推断。即研究如何利用所获得的资料对所研究的问题作出尽可能准确和可靠的结论。这两个问题密切联系,前者是后者的基础。试验设计应与试验资料统计分析相一致,这是保证试验研究臻于完善的两个基本环节。下面就土壤侵蚀定量研究具体问题,将数理统计在研究中的应用做一讨论。

2.1 土壤侵蚀试验设计

在土壤侵蚀试验中,首先要根据研究问题进行必不可少的试验设计。目前在试验设计中(野外、实验室内的模拟人工降雨试验),经常采用多元回归正交组合试验设计方法;除此之外,随机互比法试验设计也应用较多。如西北水土保持研究所蒋定生在陕北安塞水土保持试验站野外进行的水土保持措施优化配置和效益试验中,采用二次正交旋转组合试验设计方法,试验中考虑了影响土壤侵蚀的降雨强度、降雨类型、地面坡度、地面植被覆盖度和地面处理措施等因素,取5因素、5水平进行试验。通过利用二次回归正交旋转组合试验设计,使试验次数减少并提高信息资料的收集量,使试验设计和研究目的有机地结合,加快了研究的进展。如按传统方法进行5因素5水平需3125个试验小区,采用二次正交旋转组合试验设计仅需36个试验小区,然后进行统计分析,用计算机同样可模拟出全面实施方案中的试验结果。

2.2 土壤侵蚀定量研究中的统计推断

2.2.1 一元统计分析法。在土壤侵蚀试验研究中,为了揭示影响土壤侵蚀量的各因子中单因子的作用,经常采用一元统计回归分析方法来进行。这种方法简便明了,并且根据试验资料很易进行各种分析及配线。

(1)线型的选择。在配线分析中,选择线型是非常重要的,在选择线型时要考虑到经过统计分析得到的结果要与定性分析认识基本一致,并且具有物理意义。同时,也要考虑试验资料使用的区间及适用范围。在配线统计相关分析时,不能片面追求高的相关系数,否则会得出与物理概念相违背的错误结论。

土壤侵蚀的许多问题在变化过程中常出现临界极值,拟合线型时应考虑此现象。例如,在研究雨滴溅蚀时,对在各级坡度上的溅蚀量与坡度关系的选取线型的问题,应首先考虑到引起溅蚀现象发生的物理机制,其次是随坡度变化的区间。根据野外天然降雨雨滴溅蚀试验资料分析,得到以下结论:在一定坡度范围内,开始溅蚀量随着坡度的增大而增大,达到一定坡度后溅蚀量反而随着坡度的增加而减少。在选择线型时就要考虑所采取的线型应选择具有极大值存在的曲线形式,一元二次抛物线方程具有这种性质之一,所以选择一元二次抛物线形式来配线较为适宜。通过对试验资料(53次均值)的统计分析,得到向坡下溅散量 $Ed(g/m)$ 与试验坡度 $J(^{\circ})$ ($0^{\circ} \leq J \leq 40^{\circ}$)关系的回归方程:

$$Ed = 15.4 + 1.1884J - 0.02258J^2$$

由上式得出 Ed 的极大值对应的临界坡度 $J_c = 26.32^\circ$ 。

(2) 化繁为简, 化难为易。为使问题更明了地表现, 常将多因素化为一元统计分析形式来分析, 在土壤侵蚀定量中此法经常应用。这种方法要求将试验数据预先处理(各因子乘、除、加、减等), 而后进行一元统计分析。如对所用的降雨侵蚀力指标, 通常是将一次降雨动能与其30min最大雨强相乘后(即 EI_{30}), 再利用一元统计分析方法寻求在一定坡度条件下土壤侵蚀量与 EI_{30} 的关系。有时对单因子进行一元统计分析后, 先将其结果(看作常数)与另外因子进行数学处理, 再利用一元统计方法进行相关分析。

2.2.2 多元统计分析方法。多元分析方法有线性及非线性的分析方法, 是土壤侵蚀定量研究中经常采用的数学方法之一。例如: 逐步回归分析、趋势面分析、多组逐步判别分析、聚类分析、主成分分析、多项式正交旋转组合回归试验设计等。此外, 近年来其他一些数学方法也用到土壤侵蚀定量研究中, 例如灰色系统理论、模糊聚类分析, 等等。我们知道, 影响土壤侵蚀发生发展的因子很多, 为了正确地反映各影响因子对土壤侵蚀量(或流域的产沙量)的影响程度与各因子间相互关系, 在不能借助精确数学物理方法描述的前提下, 应用多元统计分析方法和其他数学方法是可取的。因此对不同的研究目的或内容, 就应借助不同的数学方法来解决。

(1) 土壤侵蚀定量评价及建模的研究。在多因子对土壤侵蚀量(或流域产沙量)的影响定量研究中, 许多学者都应用了数学分析方法, 为土壤侵蚀多因子影响定量研究作出了贡献。目前已有的坡面侵蚀、流域产沙数学模型表达形式有各种各样, 但多基于多元分析方法基础上提出的。在多因子对土壤侵蚀的定量研究中, 常用的数理统计方法为多元逐步回归分析、主成分分析、多项式正交旋转组合回归, 等等。土壤侵蚀定量研究提出的各种数学模型大都是基于数理统计提出的, 在一定程度上, 数学模型越来越趋向于简单明了, 所以在多元统计中, 也存在化繁为简、化难为易的问题。我们在进行天然降雨溅蚀研究中, 利用逐步回归分析方法进行降雨侵蚀力指标 EI_{30} 及地面坡度 J 对总溅蚀量 E_s 影响相关分析, 得到的二次多项式回归方程为:

$$E_s = 6.12 + 3.36 \times 10^{-1}(EI_{30}) + 6.94 \times 10^{-1}J - 1.509 \times 10^{-2}(EI_{30})^2 - 2.075 \times 10^{-2}J^2 + 2.522 \times 10^{-3}(EI_{30})J$$

在坡面土壤侵蚀和流域产沙的定量研究中, 为了建立各主导因素与土壤侵蚀量的定量关系, 利用上述的数理统计方法进行回归统计分析, 得到许多包括多因素的幂函数回归方程。例如, 根据安塞水土保持试验站土壤侵蚀试验场观测的降雨侵蚀资料, 进行坡面侵蚀量多因素影响分析, 研究因素有降雨量 $P(\text{mm})$ 、最大30min雨强 $I_{30}(\text{mm/min})$ 、坡度 $J(^{\circ})$ 及坡长 $L(\text{m})$ 。通过资料相关分析, 得到适用于最大30min雨强大于 0.25mm/min 时的一次降雨坡面土壤侵蚀模数 $M_s(\text{t/km}^2)$ 预测方程:

$$M_s = 13.607S^{0.712}L^{0.195}P^{0.868}I_{30}^{2.293}$$

上述多元幂函数回归方程和多元多项式回归方程是目前土壤侵蚀统计分析中常用的两种基本数学模型。不同建模者往往是从各自不同的研究目的, 依据不同的试验调查数据, 采取不同的参数和指标, 而建立实用于一定条件下的侵蚀产沙统计模型, 并采用逐

步回归分析方法,剔除回归模型中的不显著变量。

土壤侵蚀定量研究中应用新发展的一些数学方法,如上述的灰色系统理论、模糊数学,等等。例如:中国科学院地理研究所陈永宗等利用模糊数学方法综合评价黄河中游各主要支流侵蚀强度(输沙模数)与各因子(定量因子:沟壑密度、沙比系数、降雨不均匀系数;定性因子:植被状况和新构造运动)的关系,建立了模糊方程。通过模糊方程分析表明,侵蚀的产生是各种因素相互作用、相互制约的结果。

(2) 土壤侵蚀类型分区、强度分区及分级的研究。土壤侵蚀类型、强度分区属于模糊数学范畴,对于分区界线来讲,在小尺度范围内是模糊的,在大尺度范围上是分明的。因此,在土壤侵蚀类型分区、强度分区及分级时,应用数理统计中多组判别分析、聚类分析及灰色系统理论等方法较为普遍。中国科学院遥感应用研究所对陕北安塞侯沟门流域为例,应用聚类分析方法建立了土壤侵蚀强度分析模型。我们也用多组判别分析对黄土高原土壤侵蚀影响因子区域特征分区和模糊聚类分析进行了土壤侵蚀强度分级等等。

(3) 土壤侵蚀制图的研究。近几年来,随着地理信息系统(GIS)的深入发展,通过建立数字高程模型(DEM)作为格网空间位置和数据框架结构,以此利用微机进行数字侵蚀地形要素制图。在此基础上与气候、土壤、植被和土地利用等信息相复合,通过统计分析处理,可进行土壤侵蚀强度分区及其侵蚀危险评价机助制图等。

3 研究中应注意的几个问题

在土壤侵蚀定量研究中,所采用的数学方法应与成因分析紧密结合,亦即以专业分析为基础,并运用数学方法去探索科学规律,促进土壤侵蚀学科的发展。在研究中,我们采取的变量类型及数学方法一定要基于研究问题的成因概念,不能片面、盲目地选择变量或数学方法。在研究中应注意以下几个问题:(1)对所研究问题的本质属性应进行认真分析。实际上要求以成因分析和理论基础为前提,确定不同研究对象描述数学方法,这样才会取得准确的结论;(2)变量选择要适当。例如:在采用逐步回归分析方法时,一定要注意交互项的选择,如果选择不当将会剔除对研究问题有显著影响的项,这主要是由于不显著项与显著项交互的结果;(3)对问题要正确地抽象,选择合理的数学方法。例如在利用多组判别分析或用模糊聚类分析方法进行土壤侵蚀类型分区、强度分区及分级时,要求对研究问题的影响变量是非常显著。为了片面地追求变量数量,有时反而事倍功半。近几年来应用主成分分析方法很多,试图用此法分析决定各因素的权重和发现各因素的相关关系等。此方法的数学运算仅保证在所给定的数据范围内所获得的各主成分之间互不相关(这是经过严格数学推导的),在实际应用中由于影响因素的选择不当,就会产生使用不当的问题。例如次要因素远多于主要因素,通过此方法分析,将主要因素排出在外的机率很大;(4)所采用的方法应使问题化繁为简、化难为易。总而言之,在研究问题时,一定要将试验设计、理论基础和数学应用有机地结合起来,这样必将收到好的统计分析效果。

在本文写作中参考了若干已刊和未刊文章,作者和题目均不一列举。

(英文摘要下转第90页)

Effects of Different Controlling Measurements on Runoff and Quantity of Silt in Small Watersheds

Zhang Zhenzhong Zhang Ming

(Shaanxi Institute of Loess Plateau Controlling Research, Shaanxi Mizhi)

Abstract

In order to study the effects of different controlling ways on runoff and quantity of silt in small watershed with the same treatments on the slope lands of gullies and hills, some certain gullies were tested. The treatments are as follows: (1) the gully was controlled by engineer; (2) the gully was controlled by plants cover and simple engineer; (3) the gully was controlled by the natural vegetation; and (4) the traditional tillage as the check. The annual erosion by rainfall was determined in the died-dam. The data for three years was proved, the ratio of amount of silt for treatment (1), (2) and (3) to (4) was 63.1%, 74.3% and 43.6% respectively. The preliminary conclusion was that the treatment (1) was the best one, the treatment (2) was the second and the treatment (3) was the last.

Key words: Check sample gully, Controlling forms, Sediment reduction benefits

(Abstract presented article from page 91 to 94)

Statistical Analysis in Quantitative Study of Soil Erosion

Jiang Zhongshan Liu Zhi

(Northwestern Institute of Soil and Water Conservation,
Under the Academia Sinica and the Ministry of Water Conservancy)

Abstract

This paper gives a tentative discuss on the statistical analysis in quantitative study of soil erosion and its relevant problems depending on our studying experience, in order to make this study to be highly considered in the soil erosion field.

Key words: Soil erosion, Statistical analysis, Quantitative estimation