

引用格式: 马小晴, 郑明国. 基于统计检验的降雨侵蚀力简易计算模型比较[J]. 资源科学, 2018, 40(8): 1622-1633. [Ma X Q, Zheng M G. Statistical evaluation of proxies for the R factor of the Universal Soil Loss Equation[J]. *Resources Science*, 2018, 40(8): 1622-1633.] DOI: 10.18402/resci.2018.08.12

基于统计检验的降雨侵蚀力简易计算模型比较

马小晴^{1,2}, 郑明国¹

(1. 中国科学院地理科学与资源研究所, 北京 100101; 2. 中国科学院大学, 北京 100049)

摘要: 降雨侵蚀力因子(R)的准确估算对提高侵蚀预测的精度具有重要作用, 但 R 的计算需要高时间分辨率的降雨过程资料。基于日、月、年降雨资料, 目前已开发了众多的 R 值简易计算模型(PR)。本文通过Meng's检验比较了不同 PR 模型结果与 R 的相关性是否存在显著性差异, 并以此为基础评估了国内外15种 PR 模型在中国东部地区的适用性。结果表明, 15种 PR 模型计算结果与 R 均显著相关($r > 0.59, p < 0.03$), 表明这些方法均适用于中国东部地区, 但由于存在系统偏差, 许多模型需要按文中建立的回归关系式校正后得到更优的估算结果。相关系数的统计学比较表明, 日降雨 PR 模型的计算结果与 R 相关性明显优于年和月降雨模型, 而年和月降雨模型差异不明显。所有模型中, 3种基于中国数据(包括国内最常用的章文波模型和第一次全国水利普查采用的模型)和1种基于澳大利亚数据建立的日降雨 PR 模型的相关性并列处于最高等级($r > 0.99$), 但该澳大利亚模型的计算偏差随纬度增高而变大, 章文波模型也存在约30%的系统误差, 两者校正后使用能获得更好的效果; 第一次全国水利普查采用的模型则可以直接用于 R 的估算, 不需要进行校正处理。论文最终给出了获取不同时间分辨率降雨资料时建议使用的 PR 模型。

关键词: 土壤侵蚀; 降雨侵蚀力; 相关系数; Meng's检验; 统计检验; 简易计算模型

DOI: 10.18402/resci.2018.08.12

1 引言

土壤侵蚀是中国一大环境问题^[1], 其影响因子包括降雨、植被、地形等。与植被、地形等因子不同, 降雨因子很难被人为控制和改变, 是土壤侵蚀的重要地带性限制因素^[2]。降雨侵蚀力反映了降雨引发土壤侵蚀的潜在能力, 目前一般采用USLE (Universal Soil Loss Equation) 和RUSLE (Revised Universal Soil Loss Equation) 方法计算多年平均降雨侵蚀力(即 R 因子), 这两种方法都是基于次降雨的 EI_{30} 指标计算(E 为次降雨总动能, I_{30} 为最大30min降雨强度)^[3-5]。王万忠等^[6]探究了不同降雨指标与土壤侵蚀的相关性, 认为用 EI_{30} 指标计算降雨侵蚀力在中国同样适用, 目前, USLE和RUSLE中 R 因子广泛用于中国降雨侵蚀力相关研究^[7-9]。

R 因子的计算需要较高时间分辨率的降雨过程数据。Wischmeier等认为^[4], 计算 R 因子时, 降雨资料的时间分辨率应小于15min, 且资料的时间序列长度大于20a^[4], 但这类降雨资料往往不易获取, 尤其在发展中国家, 此外, 该计算方法相对复杂, 计算需要花费大量的时间和劳力。因此, 国内外众多研究者建立了简易模型(Proxies for the R factor, 以下简称 PR), 该模型可以基于更容易获取的降雨数据来估算 R 因子, 这些降雨数据包括日降雨、月降雨和年降雨等。表1列举了目前国内外应用较为普遍的15种降雨侵蚀力简易计算模型。这15种模型中, 有2种(PR_1 、 PR_2)基于年降雨数据计算, 6种基于月降雨数据计算(PR_3 — PR_8), 7种基于日降雨量数据计算(PR_9 — PR_{15}), 有5种基于中国数据建立(PR_7 、 PR_8 、

收稿日期: 2017-11-15; 修订日期: 2018-02-27

基金项目: 国家自然科学基金资助项目(41671278); 水利部黄土高原水土流失过程与控制重点实验室开放课题基金项目(2016002)。

作者简介: 马小晴, 女, 吉林白山人, 硕士生, 主要从事降雨侵蚀力研究。E-mail: mxq.15s@igsnrr.ac.cn

通讯作者: 郑明国, E-mail: zhengmg.04b@igsnrr.ac.cn

表1 降雨侵蚀力简易计算模型

Table 1 The simple models of the R factor

方程	建模地区	参考文献编号	变量说明
<p>年降雨量资料</p> $PR_1=0.048\ 30MAP^{1.610}\ (P<850\text{mm})$ $PR_1=587.8-1.219\ MAP+0.004105\ MAP^2\ (P\geq 850\text{mm})$ $PR_2=699.3+7.000\ 1\ MAP-2.719\ 0H$	<p>美国</p> <p>美国</p>	<p>[10]</p> <p>[11]</p>	<p>PR_i表示第<i>i</i>个降雨侵蚀力简易计算方法；</p> <p>AR_i、MR_j和HMR_k ($\text{MJ}\cdot\text{mm}/(\text{hm}^2\cdot\text{h}\cdot\text{a})$)分别表示第<i>i</i>年、第<i>j</i>月和第<i>k</i>个半月的降雨侵蚀力值；</p>
<p>月降雨量资料</p> $PR_3=9.8\times\sum_{i=1}^n AR_i=\sum_{i=1}^n\sum_{j=1}^{12}1.735\ 10^{\left[1.5\log\left(\frac{p_j^2}{p_a}\right)-0.818\ 8\right]}$ $PR_4=0.073\ 97\ MFI^{1.847}\ (MFI<55\text{mm})$ $PR_4=95.77-6.081MFI+0.477\ 0MFI^2\ (MFI\geq 55\text{mm})$ $MFI=\sum_{i=1}^{12}\frac{p_i^2}{MAP}$ $PR_5=21.56MFI^{0.927}$ $PR_6=25.1F$ $F=\frac{p^2}{MAP}$ $PR_7=0.358\ 9MFI^{1.9462}$ $PR_8=10\times\sum_{j=1}^{12}0.012\ 5p_j^{1.6295}$	<p>美国</p> <p>美国</p> <p>西班牙</p> <p>西班牙</p> <p>中国</p> <p>中国</p>	<p>[12]</p> <p>[10]</p> <p>[2]</p> <p>[13]</p> <p>[14]</p> <p>[15]</p>	<p>p_i、p_a、p_i、p和$MAP(\text{mm})$分别表示第<i>j</i>月的降雨量,年降雨量,第<i>i</i>月的年平均降雨量,最高的月平均降雨量以及年平均雨量；</p> <p>$rain_{10}(\text{mm})$表示一个月中日降雨量超过10mm的月降雨量,day_{10}表示一个月中日降雨量超过10mm的天数；</p> <p>$P_{d>12.7(12,0)}(\text{mm})$表示大于12.7mm(或12mm或0mm)的日降雨量；</p> <p>$P_{amax}(\text{mm})$表示多年平均的最大日降雨量；</p> <p>$H(\text{m})$和L分别表示站点的高程和以十进制为单位的纬度；</p> <p>$S(\text{mm})$表示多年平均的雨季降雨量(对于北半球区域为5—9月的降雨量,对于南半球区域为11—4月的降雨量)；</p> <p>α、β和η表示模型计算的参数,这些参数针对某一固定的站点为常量；</p> <p>F、MFI分别表示 Fournier 指数和修正的 Fournier 指数。</p>
<p>日降雨量资料</p> $PR_9=\sum_{i=1}^n\sum_{j=1}^{12}MR_j=\sum_{i=1}^n\sum_{j=1}^{12}6.97(rain_{10})_j-11.23(day_{10})_j$ $PR_{10}=\sum_{i=1}^n\sum_{j=1}^{12}M\sum_{k=1}^{12}0.266(rain_{10})_j^{2.071}(day_{10})_j^{-1.367}$ $PR_{11}=\sum_{i=1}^n\sum_{k=1}^{24}(HMR)_k=\sum_{a=1}^n\sum_{k=1}^{24}\alpha\sum_{m=1}^n(P_{d\geq 12})_m^\beta$ $\alpha=21.586\beta^{-7.1891}$ $\beta=0.836\ 3+\frac{18.144}{P_{d12}}+\frac{24.455}{P_{y12}}$ $PR_{12}=\sum_{i=1}^n\sum_{k=1}^{24}(HMR)_k=\sum_{a=1}^n\sum_{k=1}^{24}\alpha\sum_{m=1}^n(P_{d\geq 12})_m^\beta$ $\alpha=10^{2.124-1.495\beta+0.00214P_{dmax}}$ $\beta=0.836\ 3+\frac{18.144}{P_{d12}}+\frac{24.455}{P_{y12}}$ $PR_{13}=\sum_{i=1}^n\sum_{k=1}^{24}(HMR)_k=\sum_{a=1}^n\sum_{k=1}^{24}\alpha\sum_{m=1}^n(P_{d\geq 12})_m^{1.7265}$ $\alpha=\begin{cases} 0.393\ 7 & \text{暖季(5—9月)} \\ 0.310\ 1 & \text{冷季(10—12月,1—4月)} \end{cases}$ $PR_{14}=\sum_{i=1}^n\sum_{j=1}^{12}MR_j=\sum_{i=1}^n\sum_{j=1}^{12}\alpha[1+\eta\cos(2\pi fi-\omega)]\sum_{m=1}^n(P_{d\geq 12.7})_m^\beta$ $f=\frac{1}{12};\omega=\frac{\pi}{6};\beta=1.49;\eta=0.29$ $\alpha=0.359[1+0.098\ 9\exp(3.26S/MAP)]$ $PR_{15}=\sum_{i=1}^n\sum_{j=1}^{12}MR_j=\sum_{i=1}^n\sum_{j=1}^{12}\alpha[1+\eta\cos(2\pi fi-\omega)]\sum_{m=1}^n(P_{d\geq 0})_m^\beta$ $f=\frac{1}{12};\omega=\frac{\pi}{6};\eta=0.3$ $\beta=1.02-0.020\ 9L$ $\alpha_0=1.05\times 10^{(2.08-1.58\beta)}$ $\alpha/\alpha_0=2.349+0.040\ 40L-0.000\ 268\ 4H$	<p>马来西亚</p> <p>马来西亚</p> <p>中国</p> <p>中国</p> <p>中国</p> <p>澳大利亚</p> <p>澳大利亚</p>	<p>[16]</p> <p>[16]</p> <p>[17]</p> <p>[18]</p> <p>[19]</p> <p>[20]</p> <p>[21]</p>	

PR_{11} 、 PR_{12} 、 PR_{13}),其中 PR_{11} 在国内应用最为普遍(一般称为章文波模型), PR_{13} 为中国第一次水利普查时所采用的模型。

如上所述,长时间序列的高分辨率降雨过程数据很难获取,甚至表1中的许多研究也并未获取到按USLE或RUSLE要求计算的 R 值,并以此为基准来建立或校正 PR 模型。如国内运用最为普遍的章文波模型^[17],在其建立过程中根据 $P_d I_{10d}$ (P_d 表示日降雨量, I_{10d} 表示日最大10min降雨强度)指标估算 R 值^[22],并以此为基准建立模型。显然,以 R 值的估计值作为基准建立降雨侵蚀力模型,必然会影响模型的精度,此外,大多数研究者是根据相对偏差(relative error)和纳什效率系数(Nash-Sutcliffe efficiency coefficient)评价 PR 对 R 的逼近程度^[2, 23-25],而相对偏差和纳什效率系数都是样本数据的函数,属于随机变量,随样本数据的变化而变化,因此必须采用统计检验的方法以排除随机性因素的影响,但目前的相关研究中基本没有对相对偏差和纳什效率系数进行统计学比较,因此难以排除所观测到的相对偏差和纳什效率系数的提高仅仅是随机因素造成的结果这一可能性。

本研究获取了中国东部地区14个站点的基于1min分辨率降雨资料计算的 R 值数据,以此为基础,本文以相关系数为评价指标,对表1中15种 PR

模型在中国东部地区的适用性进行了统计学评估。基于该评估结果,本文也对不同时间分辨率降雨资料情况下的最优 PR 模型给出了建议。

2 数据来源与研究方法

2.1 数据来源

中国东部地区主要受水力侵蚀的影响,西部地区主要受风蚀和冻融侵蚀的影响^[9]。本文以中国东部地区的14个站点为研究区,所采用的 R 值引用自参考文献[8],为根据1min时间分辨率的降雨过程资料计算结果。用于计算 PR 值的日降雨数据源于国家气象科学数据共享服务平台(<http://data.cma.cn/>)^[26],并根据这些数据计算获得月和年雨量数据。计算 PR 的降雨资料的年限与 R 值保持一致,除五寨和阳城为1971—2000年,其余站点为1961—2000年,个别年份数据缺测较多没有参与计算。各站点基本情况及其 R 值见表2。

2.2 R 和 PR 的计算

R 值按RUSLE方法计算,公式为:

$$R = \frac{1}{n} \sum_{j=1}^n \sum_{k=1}^m (EI_{30})_k \quad (1)$$

$$E = \left(\sum_{r=1}^n (e_r \cdot P_r) \right) \quad (2)$$

$$e_r = 0.29(1 - 0.72 \exp(-0.082i_r)) \quad (3)$$

式中 R 为多年平均降雨侵蚀力; n 为降雨资料的年

表2 研究区气象站点基本信息

Table 2 Information of the selected meteorological stations

省份	名称	经度/°E	纬度/°N	高程/m	降雨资料年数/a	$R/(MJ \cdot mm/(hm^2 \cdot h \cdot a))$
黑龙江	嫩江	125.13	49.10	222.3	30	1 368.7
	通河	128.44	45.58	108.6	38	1 632.5
山西	五寨	111.49	38.55	1 401.0	30	781.9
	阳城	112.24	35.29	659.5	30	1 503.3
陕西	绥德	110.13	37.30	929.7	29	992.8
	延安	109.30	36.36	958.5	39	1 233.7
四川	西昌	102.16	27.53	1 590.9	40	3 021.0
	遂宁	105.35	30.30	278.2	33	4 091.3
湖北	房县	110.46	32.02	426.9	31	2 298.4
	黄石	115.03	30.15	19.6	32	6 049.4
云南	腾冲	98.30	25.01	1 654.6	36	3 648.9
	昆明	102.41	25.01	1 892.4	33	3 479.0
福建	福州	119.17	26.05	84.0	39	5 871.1
	长汀	116.22	25.51	310.0	31	8 258.5

2018年8月

限; m 为一年中侵蚀性降雨的次数; EL_{30} 为一次降雨事件的降雨侵蚀力; E 为一次降雨过程总的降雨动能; I_{30} 为一次降雨中最大30分钟降雨强度; e_r 、 P_r 和 i_r ,分别为单位降雨动能、第 r 个时段的降雨量和第 r 个时段的降雨强度。

本文选择了表1中15种PR模型与RUSLE模型中R因子的计算方法进行比较,所有PR的单位统一为MJ·mm/(hm²·h·a),PR模型的计算公式详见表1。

2.3 相关系数差异的显著性检验

本文通过相关系数的统计学比较来评价不同的PR模型。若某一PR与R的相关系数显著大于另一PR,则认为该PR更适合用于估算R,如未通过“显著大”或“显著小”的检验,则认为两者与R的相关系数相等,用于估算R时效果相同。

一般情况下,由于总体的所有样本数据不可能全部获取,故获取的相关系数随样本的变化而变化,因此相关系数的比较必须基于统计学检验。相关分为独立性相关(independent correlation)和非独立性相关(dependent correlation)两大类,需采用不同类别的检验方法。本文比较表1中PR₁—PR₁₅与R之间的相关系数(即 r_1 、 r_2 、 \dots 、 r_{15}),这些相关系数建立时均采用了相同的R变量,故应采用非独立性相关检验方法。本文采用Meng's检验方法^[27],该方法作为一种比较非独立性相关的方法,目前广泛应用于社会科学研究。Zheng等^[28]用此方法分析了黄土高原最佳次降雨侵蚀力指标。陈晓安等^[29]也从统计学角度评估了黄土高原最佳次降雨侵蚀力指标,但采用的方法为独立性相关系数的比较方法,不适合用于非独立性相关系数的比较。

Meng's检验对两个和两个以上相关系数的比较,采用不同的计算公式。设有 k 个自变量($x_1, x_2, \dots, x_i, \dots, x_k$)和一个因变量 y , r_i 表示 x_i 与 y 的相关系数。当比较两个相关系数时(假定为 r_1 和 r_2),Meng's检验的计算公式为:

$$Z = (z_{r1} - z_{r2}) \sqrt{\frac{N-3}{2(1-r_x)h}} \quad (4)$$

$$h = \frac{1 - f r_i^2}{1 - r_i^2} \quad (5)$$

$$f = \frac{1 - r_x}{2(1 - r_i^2)} \quad (6)$$

式中 Z 值服从标准正态分布; z_{ri} 为 r_i 的Fisher z变换结果, $z_{ri} = \frac{1}{2} \ln \frac{1+r_i}{1-r_i}$,通过该变换将 r_i 转换为正态分布形式; N 为样本数量; r_x 为 x_1 和 x_2 间的相关系数, $\overline{r_i^2} = (r_1^2 + r_2^2)/2$;当公式(6)的计算结果大于1时, f 取1。

当比较多个相关系数($r_1, r_2, \dots, r_i, \dots, r_k; k > 2$)时,Meng's检验的计算公式为:

$$\chi^2(k-1) = \frac{(N-3) \sum_i (z_{ri} - \overline{z_r})^2}{(1-r_x)h} \quad (7)$$

式中 $\chi^2(k-1)$ 服从 $k-1$ 个自由度的卡方分布; $\overline{z_r}$ 为 z_{ri} 的平均值; r_x 为所有 k 个非独立自变量组间相关系数的中位数; $\overline{r_i^2}$ 为 r_i^2 的平均值。

公式(7)的原假设为待检验的 k 个相关系数全相等,该公式只能判断 k 个相关系数全相等或全不相等,不能具体判定哪一个为异常值(显著大或显著小)。以下公式(8)通过检验某一相关系数 r_i 与其它相关系数的平均值是否有显著差异,以此确定 r_i 在待检验的 k 个相关系数中是否为异常值:

$$Z = r_{\lambda_i} \sqrt{\chi^2(k-1)} \quad (8)$$

式中 r_{λ_i} 为 z_{ri} 和 λ_i 间的相关系数; λ_i 为 z_{ri} 对应的权重,其和等于0。当对 r_i 进行判定时,除 λ_i 外的其余权重值均设置为1,例如,当判定 r_i 在(r_1, r_2, r_3, r_4)4个相关系数中是否为异常值时, λ_i 的取值依次为-3、1、1、1; Z 值服从标准正态分布。

进行相关系数比较时,所进行的检验均为单侧检验,按一般规范,进行统计检验时显著性水平定为0.05。如对公式(8),其显著性水平的单侧检验阈值为 ± 1.65 ,如计算的 Z 值小于-1.65,则认为待检验的 r_i 为“显著大”的异常值,如 Z 值大于1.65,则 r_i 为“显著小”的异常值,否则认为 r_i 与其它相关系数的均值并无显著差异, r_i 为非异常值。

本研究设计了一种逐步检验的方法来进行 $r_1, r_2, \dots, r_i, \dots, r_k$ 的排序。首先以所有待检验的相关系数为输入,通过公式(8),逐个判断每个相关系数是否为异常值,如对 r_i 的检验结果为显著小,则将 r_i 划归为L组(Less-correlated Class),否则划归为O组(Other Class);然后,分别以L组和O组为输入,重复进行以上步骤,直至每个组中均未检测到异常值或仅包含一个相关系数为止。

除进行相关系数的比较外,另选择平均相对误差($MRE, \%$)、均方根误差($RMSE, MJ \cdot mm/(hm^2 \cdot h \cdot a)$)和纳什效率系数(NSE)来评估 PR 对 R 的估算偏差,计算公式为:

$$MRE = 100 \times \frac{\sum_{i=1}^n |R_j - PR_j|}{R_j} \quad (9)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (PR_j - R_j)^2}{n}} \quad (10)$$

$$NSE = 1 - \frac{\sum_{i=1}^n (R_j - PR_j)^2}{\sum_{i=1}^n (R_j - \bar{R})^2} \quad (11)$$

式中 n 为站点的数量; R_j 和 PR_j 为第 j 个站点的 R 值和 PR 值; \bar{R} 为所有站点 R 的均值。

3 结果及分析

如表2所示,14个站点的 R 值变化在(781.9~8258.5) $MJ \cdot mm/(hm^2 \cdot h \cdot a)$ 之间,平均值为3159.3 $MJ \cdot mm/(hm^2 \cdot h \cdot a)$,最大值出现在长汀站,为8258.5 $MJ \cdot mm/(hm^2 \cdot h \cdot a)$,最小值出现在五寨站,为781.9 $MJ \cdot mm/(hm^2 \cdot h \cdot a)$,从南到北逐渐递减,表现出明显纬度地带性(R 与站点纬度的相关系数

$r = -0.69, p = 0.007$),经度地带性不明显($r = -0.008, p = 0.98$)。各 PR 值也表现出类似的空间趋势。

PR 与 R 的相关系数及其显著性检验结果见图1。进行比较的15种 PR 均与 R 显著相关($r > 0.59, p < 0.03$),表明这些方法都可用于估算中国东部地区的 R 值。图1同时给出了 PR 和 R 间的线性回归方程,这些方程可用于对各 PR 进行修正。表3给出了直接用 PR 和按图1中方程修正后的 PR 估算 R 值的 $RMSE$ 、 MRE 和 NSE 。成对 t 检验表明,修正情况下的 $RMSE$ ($p = 0.008$)、 MRE ($p = 0.002$)均显著变小, NSE ($p = 0.015$)显著变大,表明修正后的估算误差显著降低。但并非所有的 PR 均如此,如 PR_8 和 PR_{13} 修正前后的 $RMSE$ 、 MRE 和 NSE 变化很小。

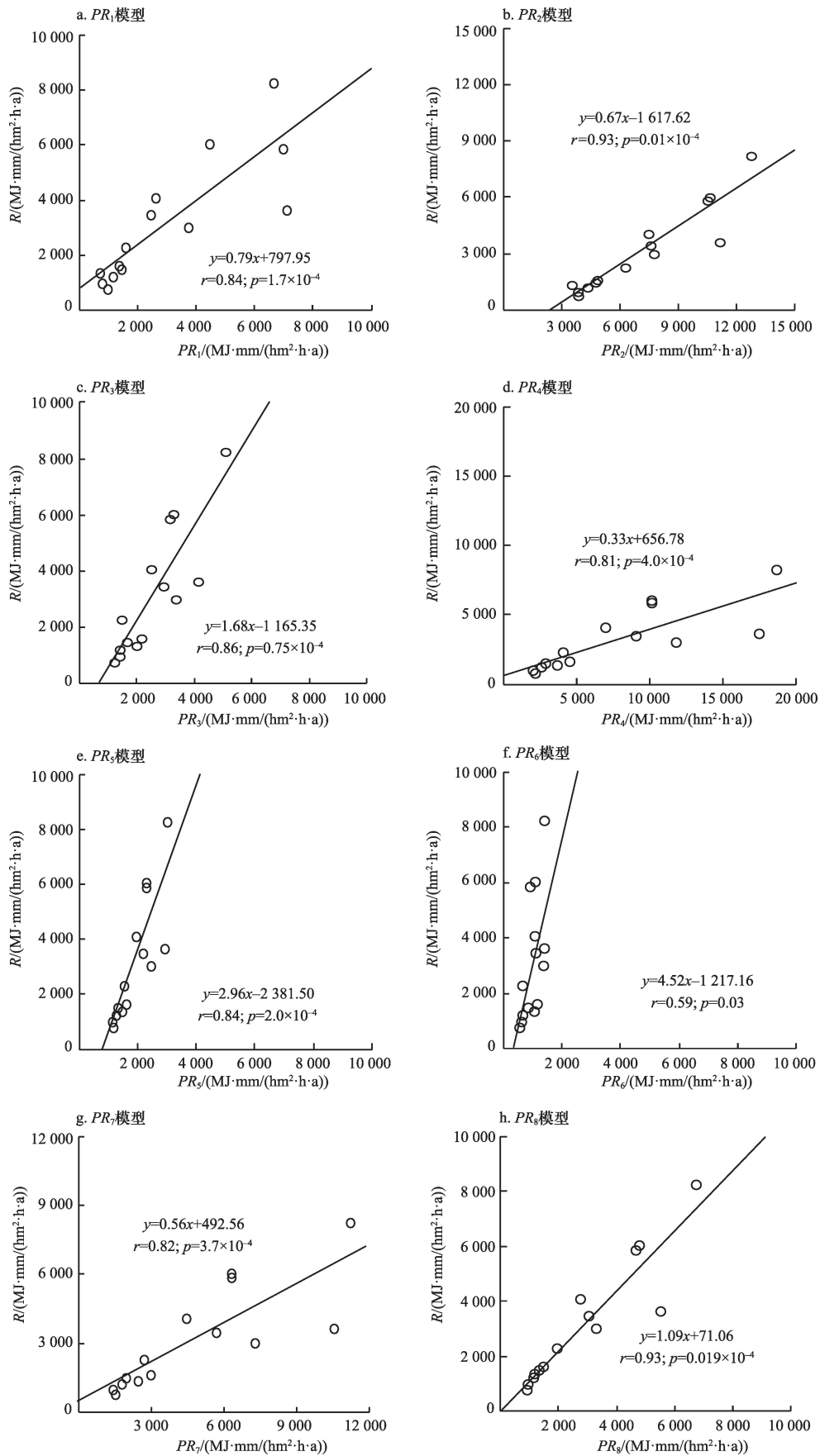
按本文设计的逐步检验方法(流程见图2,见第1629页),方程(8)首先将15种 PR 与 R 的相关系数分为了 L_1 组($r_1, r_3, r_4, r_5, r_6, r_7, r_{14}$)和 O_1 组($r_2, r_8, r_9, r_{10}, r_{11}, r_{12}, r_{13}, r_{15}$)。 L_1 组中的相关系数经检验又可分为 L_2 组(r_6)和 O_2 组($r_1, r_3, r_4, r_5, r_7, r_{14}$),经检验 O_2 组内相关系数大小无显著性差异($p > 0.16$)。 O_1 组内系数经检验分为 L_3 组(r_2, r_8, r_9)和 O_3 组(r_{10}, r_{11}, r_{12} 、

表3 PR 模型估算 R 值的平均相对误差、均方根误差和纳什效率系数

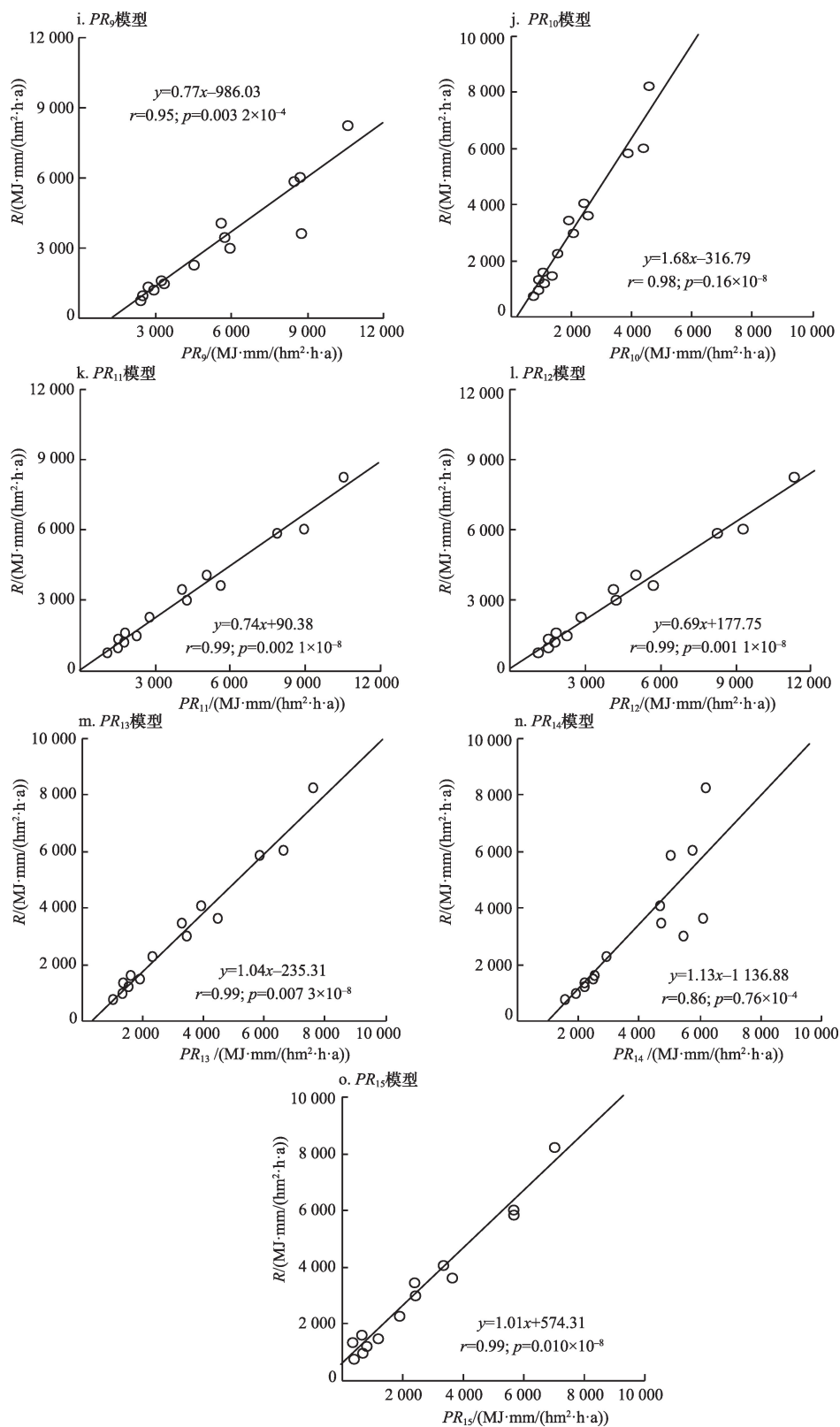
Table 3 $MRE, RMSE$ and NSE of the selected PRs in estimating R value

PR_i	直接估算			修正后估算		
	MRE /%	$RMSE$ /($MJ \cdot mm/(hm^2 \cdot h \cdot a)$)	NSE	MRE /%	$RMSE$ /($MJ \cdot mm/(hm^2 \cdot h \cdot a)$)	NSE
PR_1	29.0	1 287	0.65	31.8	1 183	0.70
PR_2	174.3	4 122	-2.60	15.9	775	0.87
PR_3	32.5	1 459	0.55	29.9	1 102	0.74
PR_4	147.0	5 819	-6.18	32.7	1 262	0.66
PR_5	31.0	2 128	0.04	26.6	1 195	0.70
PR_6	57.5	2 978	-0.88	54.9	1 755	0.35
PR_7	58.8	2 438	-0.26	31.7	1 253	0.67
PR_8	17.7	895	0.83	15.2	816	0.86
PR_9	96.8	2 396	-0.22	13.1	703	0.90
PR_{10}	29.2	1 457	0.55	12.9	454	0.96
PR_{11}	31.7	1 314	0.63	10.6	316	0.98
PR_{12}	35.5	1 550	0.49	11.0	299	0.98
PR_{13}	12.5	376	0.97	9.7	351	0.97
PR_{14}	49.9	1 294	0.65	18.7	1 103	0.74
PR_{15}	28.2	698	0.90	14.1	361	0.97
均值	55.4	2 014	-0.26	21.9	862	0.80

注:直接估算是指直接以 PR 值作为 R 的估计值,修正后估算指以图1中的回归方程的计算结果为 R 的估计值。



(续图1)

图1 R 和 PR 的相关关系及回归方程Figure 1 Relationships between R and PR

2018年8月

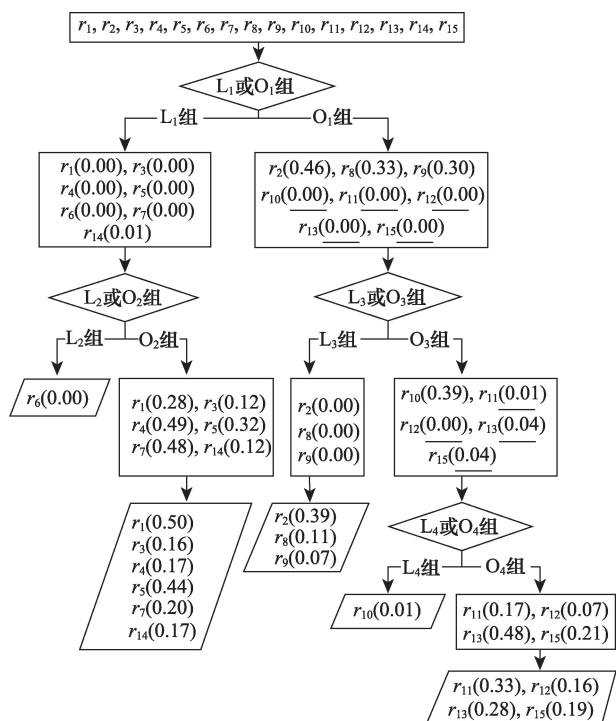


图2 R与15个PR间的相关系数差异性检验流程

Figure 2 Flow chart for comparing correlations between R and PR

注:括号中的数字为按公式(8)检验的 p 值,L组表示显著低于平均水平的相关系数,O组包括显著高于平均水平(加下划线)以及与平均水平无显著差异的相关系数。每一个平行四边形表示一个终端节点,其中的相关系数无显著性差异或仅有一个相关系数。

r_{13} 、 r_{15}),其中L3组内的相关系数无显著性差异($p > 0.07$)。O3组统计检验进一步分为L4组(r_{10})和O4组(r_{11} 、 r_{12} 、 r_{13} 、 r_{15}),O4组内相关系数的Meng's检验表明无显著性差异($p > 0.16$)。经过四轮检验后,得到了15种PR与R的相关系数的排序:

$$r_6 < r_1 = r_3 = r_4 = r_5 = r_7 = r_{14} < r_2 = r_8 = r_9 < r_{10} < r_{11} = r_{12} = r_{13} = r_{15} \quad (12)$$

式中 r_i 表示第 i 个PR与R之间的相关系数。方程(12)将表1中15个PR按与R的相关性大小分成了5组,从高到低依次为:(r_{11} 、 r_{12} 、 r_{13} 、 r_{15}),(r_{10}),(r_2 、 r_8 、 r_9),(r_1 、 r_3 、 r_4 、 r_5 、 r_7 、 r_{14}),(r_6)。公式(7)的检验结果也表明每一个组内(只有一个相关系数值的除外)中的相关系数全相等($p > 0.27$)。从趋势上看,基于日降雨简易计算模型(PR_9 — PR_{15})要显著优于年降雨(PR_1 、 PR_2)和月降雨(PR_3 — PR_8)模型,但年降雨和月降雨模型差异不明显,这和表3中MRE、RMSE和NSE的结果基本一致。章文波等也得到了类似结果^[14]。

PR_1 和 PR_2 是基于年降雨量资料的简易计算模

型。方程(12)表明, r_2 显著大于 r_1 ,表3中数据也表明,根据修正后的 PR_2 估算R值要明显优于 PR_1 , PR_2 与R的相关性甚至显著优于许多月(PR_3 — PR_7)和日(PR_{14})模型,因此当仅有年降雨量资料时,可采用修正后的 PR_2 估算R值。

PR_3 — PR_8 为所选择的6种基于月降雨数据的简易计算模型。方程(12)表明, r_8 显著大于其它月降雨模型。因此当仅有月降雨量资料时,应选择 PR_8 进行R值的估算。此外,注意到表3中对 PR_8 是否作修正处理对估算结果影响不大,图1中 PR_8 的回归方程的斜率项也很接近于1,进一步的t检验结果表明,可认为其斜率项显著等于1($p = 0.49$),常数项显著等于0($p = 0.87$),因此可以直接用 PR_8 进行R值的估算,不必按图1中的方程进行修正。 PR_8 是吴素业在大别山区建立的月降雨量模型,田刚等人对比分析了国内外常用的6种降雨侵蚀力简易计算模型在江西省潦水河流域的适用性,也发现 PR_8 的结果最理想,具有最高的纳什效率系数和最低的相对偏差^[24]。表1中6种月降雨侵蚀力模型中, PR_8 与其它模型的建模方式不同, PR_3 为复合型函数,包含了对数函数和幂函数形式, PR_4 — PR_7 是基于MFI指数或F指数建立的模型,而 PR_8 为幂函数模型。Richardson^[30]的研究结果也表明降雨量与降雨侵蚀力一般表现为幂函数关系,本文比较的日降雨模型中,相关性最高的PR模型(PR_{11} 、 PR_{12} 、 PR_{13} 、 PR_{15})也均为幂函数模型。

表1中包括7种基于日降雨的简易侵蚀力计算模型(PR_9 — PR_{15})。按方程(12), PR_{11} 、 PR_{12} 、 PR_{13} 、 PR_{15} 与R的相关性没有显著差异,且显著优于其它模型,这和表3中MRE、RMSE和NSE的计算结果一致。 PR_{15} 模型公式包含了纬度信息,该模型基于澳大利亚新南威尔士州的数据建立,该地区纬度在21°S—37°S之间。图3d表明,直接用 PR_{15} 估算中国东部地区的R值时,计算结果全部偏小,且偏小程度随纬度增加而增加,纬度超过40°的地区相对误差可达到60%~80%。表3计算中数据表明,用图1中的公式修正后, PR_{15} 的MRE由31.7%降低到了10.6%,因此,尤其当应用于中国东部高纬度地区时, PR_{15} 按图1中的公式进行校正能够得到更为优化的结果。

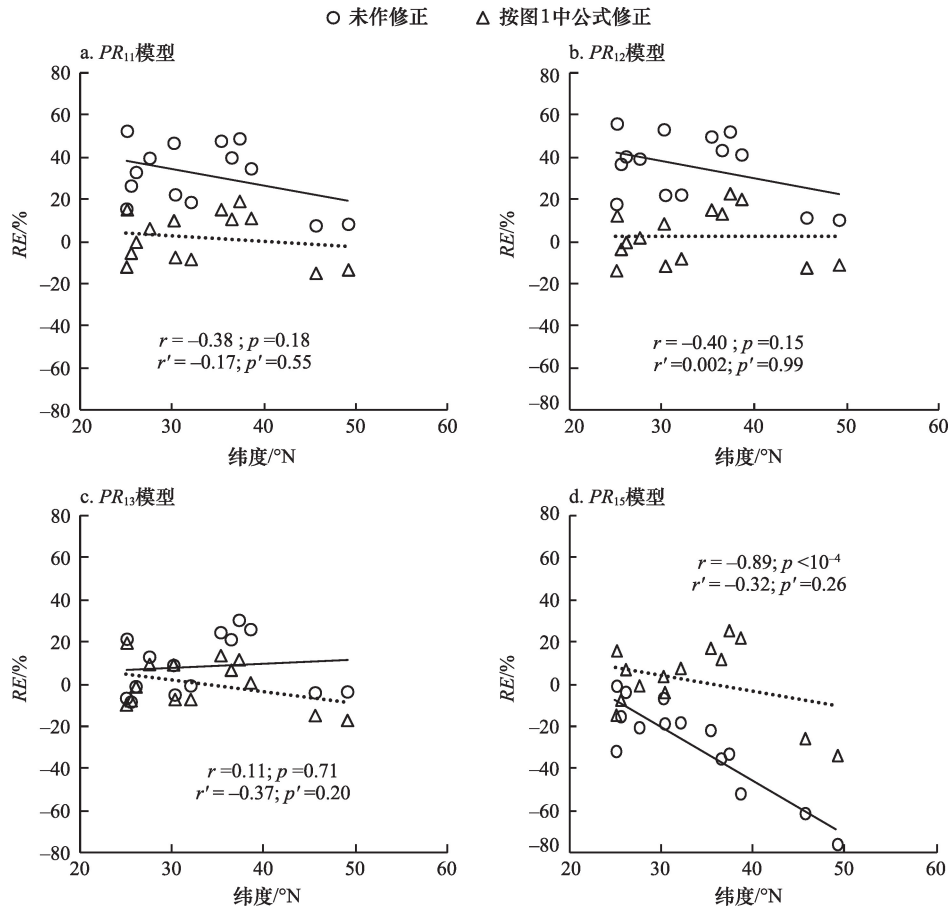


图3 纬度与相对误差的关系

Figure 3 Relationship between latitude and RE

注:图中实线、 r 、 p 分别表示直接估算的相对误差的趋势线、与纬度的相关系数以及 p 值;图中虚线、 r' 、 p' 分别表示修正后的相对误差的趋势线、与纬度的相关系数以及 p 值。

PR_{11} 目前在国内运用最为广泛, PR_{12} 仅对 PR_{11} 作了轻微调整,两者的计算结果非常接近,使得图1中两者的校正公式也很接近。值得关注的是图1中 PR_{11} 和 PR_{12} 与 R 回归方程的斜率项分别为0.74和0.69,与此对应,直接使用 PR_{11} 和 PR_{12} 估算 R 值时大部分情况下误差大于30%(图3a和图3b),因此可认为 PR_{11} 和 PR_{12} 估算 R 值时存在系统偏差,校正后得到更优的估算结果。钟莉娜等^[31]在黄土高原的研究结果也表明, PR_{11} 和 R 之间存在显著的线性关系: $R = 0.849PR_{11} - 29.651$,该公式同样表明 PR_{11} 系统地高估了 R 值,文献[9]中图8也清楚表明这一趋势。 PR_{11} 是以指标 $P_{dI_{10d}}$ 作为基准值建立的, $P_{dI_{10d}}$ 与 EI_{30} 的转换关系是依据全国8个气象站共约104个年度的降雨资料建立,在这104个年度中北方地区约占80%^[22]。这导致主要基于北方数据建立的 $P_{dI_{10d}}$ 和

EI_{30} 的转换关系可能并不能很好地适用于中国东部,因此建议用修正后的 PR_{11} 和 PR_{12} 估算 R 。

PR_{13} 是中国第一次水利普查估算 R 的方法,图1中 PR_{13} 的回归方程的斜率项非常接近于1,常数项也较小,t检验结果表明,其斜率项可认为显著等于1($p = 0.45$),常数项可认为显著等于0($p = 0.24$),因此当能够获取到日降雨量资料时,可直接用 PR_{13} 估算 R 值,表3中数据也表明修正后的 PR_{13} 估算 R 值的精度并没有显著提高。

4 结论

(1)本文分析比较了国内外运用较为广泛的15种 PR 模型与 R 之间的相关性,并建立了各 PR 与 R 的回归关系式。参与比较的15种 PR 都与 R 显著相关($r > 0.59$, $p < 0.03$),相关系数统计检验的结果表明,日降雨模型的计算结果与 R 的相关系数显著高

2018年8月

于年降雨和月降雨模型,而年降雨和月降雨模型差异不明显。由于大多数 PR 直接估算 R 时存有系统偏差,运用建立的回归关系式校正后能获得更好的估计精度。

(2)依据相关系数统计学比较结果,本文对不同时间分辨率降雨资料情况下的最优 PR 模型给出了建议。当仅有年降雨量资料时,建议采用 PR_2 模型;当仅有月降雨量资料时,建议采用 PR_8 模型;以日降雨量数据估算 R 值时,使用 PR_{11} 、 PR_{12} 、 PR_{13} 和 PR_{15} 均可。

(3)相关系数分析结果表明,其它国家建立的 PR 模型也可以较好地运用于中国,尤其是澳大利亚的模型 PR_{15} ,在与 R 的相关性排序中,和中国的3个模型 PR_{11} 、 PR_{12} 、 PR_{13} 并列处于最高等级,但该模型的预测误差随纬度增大而增大,按文中给出的公式校正后能获得更为优化的结果。 PR_{11} 为目前国内应用最为广泛的降雨侵蚀力模型,但估算 R 存在约30%系统偏差,实际运用时也最好作校正处理。 PR_{13} 为第一次全国水利普查采用的降雨侵蚀力模型,该模型估算 R 时不存在系统偏差,可直接用于估算 R 值。

致谢 本文在数据搜集过程中得到了华中农业大学蔡崇法教授以及北京师范大学符素华教授、殷水清副教授等的鼎力协助,特此致谢。

参考文献(References):

- [1] 景可,王万忠,郑粉莉. 中国土壤侵蚀与环境[M]. 北京: 科学出版社, 2005. [Jing K, Wang W Z, Zheng F L. Soil Erosion and Environment in China[M]. Beijing: Science Press, 2005.]
- [2] Angulomartínez M, Beguería S. Estimating rainfall erosivity from daily precipitation records: a comparison among methods using data from the Ebro Basin (NE Spain)[J]. *Journal of Hydrology*, 2009, 379(1): 111–121.
- [3] Wischmeier W H, Smith D D. Predicting Rainfall–Erosion Losses from Cropland East of the Rocky Mountains: Guide for Selection of Practices for Soil and Water Conservation [M]. Washington: USDA Agriculture Handbook No. 282, 1965.
- [4] Wischmeier W H, Smith D D. Predicting Rainfall Erosion Losses—a Guide to Conservation Planning[M]. Washington: USDA Agriculture Handbook 537, 1978.
- [5] Renard K G, Foster G R, Weesies G, et al. Predicting Soil Erosion by Water: A Guide to Conservation Planning with the Revised Universal Soil Loss Equation (RUSLE)[M]. Washington: USDA Agriculture Handbook 703, 1997.
- [6] 王万忠,焦菊英,郝小品,等. 中国降雨侵蚀力 R 值的计算与分布(I)[J]. 水土保持学报, 1995, (4): 7–18. [Wang W Z, Jiao J Y, Hao X P, et al. Study on rainfall erosivity in China[J]. *Journal of Soil and Water Conservation*, 1995, (4): 7–18.]
- [7] 徐丽,谢云,符素华,等. 北京地区降雨侵蚀力简易计算方法研究[J]. 水土保持研究, 2007, 14(6): 433–437. [Xu L, Xie Y, Fu S H, et al. Simple method of estimating rainfall erosivity under different rainfall amount of Beijing[J]. *Research of Soil and Water Conservation*, 2007, 14(6): 433–437.]
- [8] Yin S, Xie Y, Liu B, et al. Rainfall erosivity estimation based on rainfall data collected over a range of temporal resolutions[J]. *Hydrology & Earth System Sciences Discussions*, 2015, 19(10): 4965–4996.
- [9] Xie Y, Yin S Q, Liu B Y, et al. Models for estimating daily rainfall erosivity in China[J]. *Journal of Hydrology*, 2016, 535: 547–558.
- [10] Renard K G, Freimund J R. Using monthly precipitation data to estimate the R -factor in the revised USLE[J]. *Journal of Hydrology*, 1994, 157(1–4): 287–306.
- [11] Mikhailova E A, Bryant R B, Schwager S J, et al. Predicting rainfall erosivity in Honduras[J]. *Soil Science Society of America Journal*, 1997, 61(1): 273–279.
- [12] Wischmeier W H, Smith D D. Rainfall energy and its relationship to soil loss[J]. *Eos Transactions American Geophysical Union*, 1958, 39(2): 285–291.
- [13] Hernando D, Romana M G. Estimating the rainfall erosivity factor from monthly precipitation data in the Madrid Region (Spain)[J]. *Journal of Hydrology & Hydromechanics*, 2015, 63(1): 55–62.
- [14] 章文波,付金生. 不同类型雨量资料估算降雨侵蚀力[J]. 资源科学, 2003, 25(1): 35–41. [Zhang W B, Fu J S. Rainfall erosivity estimation under different rainfall amount[J]. *Resources Science*, 2003, 25(1): 35–41.]
- [15] 吴素业. 安徽大别山区降雨侵蚀力简化算法与时空分布规律[J]. 中国水土保持, 1994, (4): 12–13. [Wu S Y. The simplified rainfall erosivity model and spatial and temporal distribution in Dabie Mountains in Anhui[J]. *Soil and Water Conservation of China*, 1994, (4): 12–13.]
- [16] Shamsad A, Azhari M N, Isa M H, et al. Development of an appropriate procedure for estimation of RUSLE EI30 index and preparation of erosivity maps for Pulau Penang in Peninsular Malaysia[J]. *Catena*, 2008, 72(3): 423–432.
- [17] 章文波,谢云,刘宝元. 利用日雨量计算降雨侵蚀力的方法研究[J]. 地理科学, 2002, 22(6): 705–711. [Zhang W B, Xie Y, Liu B Y. Rainfall erosivity estimation using daily rainfall amounts[J]. *Scientia Geographica Sinica*, 2002, 22(6): 705–711.]
- [18] Men M X, Yu Z R, Xu H. Study on the spatial pattern of rainfall

- erosivity based on geostatistics in Hebei Province, China[J]. *Frontiers of Agriculture in China*, 2008, 2(3): 281–289.
- [19] 刘宝元, 郭索彦, 李智广, 等. 中国水力侵蚀抽样调查[J]. 中国水土保持, 2013, (10): 26–34. [Liu B Y, Guo S Y, Li Z G, *et al.* The sampling survey of water erosion in China [J]. *Soil and Water Conservation of China*, 2013, (10): 26–34.]
- [20] Yu B F. Rainfall erosivity and its estimation for Australia's tropics [J]. *Australian Journal of Soil Research*, 1998, 36(1): 143–165.
- [21] Yang X H, Yu B F. Modelling and mapping rainfall erosivity in New South Wales[J]. *Australia. Soil Research*, 2015, 53(2): 178–189.
- [22] 谢云, 章文波, 刘宝元. 用日雨量和雨强计算降雨侵蚀力[J]. 水土保持通报, 2001, 21(6): 53–56. [Xie Y, Zhang W B, Liu B Y. Estimation of rainfall erosivity using rainfall amount and rainfall intensity[J]. *Bulletin of Soil and Water Conservation*, 2001, 21(6): 53–56.]
- [23] 伍育鹏, 谢云, 章文波. 国内外降雨侵蚀力简易计算方法的比较[J]. 水土保持学报, 2001, 15(3): 31–34. [Wu Y P, Xie Y, Zhang W B. Comparison of different methods for estimating average annual rainfall erosivity[J]. *Water Conservation*, 2001, 15(3): 31–34.]
- [24] 田刚, 梁音, 陈效民, 等. 几种降雨侵蚀力模型在湫水河流域的对比研究[J]. 土壤学报, 2010, 47(1): 7–12. [Tian G, Liang Y, Chen X M, *et al.* Comparative study on several rainfall erosivity models in Lianshui Basin[J]. *Acta Pedologica*, 2010, 47(1): 7–12.]
- [25] 吴秋菊, 吴发启, 翟艳斌, 等. 陕北黄土高原降雨侵蚀力简易计算模型的比较研究[J]. 灌溉排水学报, 2011, 30(6): 138–140. [Wu Q J, Wu F Q, Zhai Y B, *et al.* Comparative study of rainfall erosion model in loess plateau in the North of Shanxi Province[J]. *Journal of Irrigation and Drainage*, 2011, 30(6): 138–140.]
- [26] 国家气象科学数据共享服务平台. 中国地面气候资料日值数据集(V3.0) [EB/OL]. (2012–08–04)[2017–11–15]. http://data.cma.cn/data/cdcdetail/dataCode/SURF_CLI_CHN_MUL_DAY_V3.0.html. [National Meteorological Science Data Sharing Service Platform. Daily Climate Data Set of China [EB/OL]. (2012–08–04)[2017–11–15]. http://data.cma.cn/data/cdcdetail/dataCode/SURF_CLI_CHN_MUL_DAY_V3.0.html.]
- [27] Meng X, Rosenthal R, Rubin D B. Comparing correlated correlation coefficients [J]. *Psychological Bulletin*, 1992, 111(1): 172–175.
- [28] Zheng M, Chen X. Statistical determination of rainfall–runoff erosivity indices for single storms in the Chinese Loess Plateau[J]. *Plos One*, 2015, 10(3): e0117989.
- [29] 陈晓安, 蔡强国, 郑明国, 等. 黄土丘陵沟壑区的岔巴沟流域次降雨侵蚀力的研究[J]. 泥沙研究, 2010, (1): 5–10. [Chen X A, Cai Q G, Zheng M G, *et al.* Study on rainfall erosivity of Chabagou watershed in a hilly loess region on the loess plateau[J]. *Journal of Sediment Research*, 2010, (1): 5–10.]
- [30] Richardson C W, Foster G R, Wright D A. Estimation of erosion index from daily rainfall amount [J]. *Transactions of the ASAE*, 1983, 26(1): 153–156.
- [31] 钟莉娜, 王军, 赵文武. 基于修正简易模型的陕北黄土丘陵沟壑区降雨侵蚀力分布特征[J]. 中国水土保持科学, 2016, 14(5): 8–14. [Zhong L N, Wang J, Zhao W W. Temporal and spatial distribution characteristics of rainfall erosivity in loess hilly region of Northern Shaanxi based on the modified simplified models[J]. *Science of Soil and Water Conservation*, 2016, 14(5): 8–14.]

Statistical evaluation of proxies for the R factor of the Universal Soil Loss Equation

MA Xiaoqing^{1,2}, ZHENG Mingguo¹

(1. Institute of Geographic Sciences and Nature Resources Research, Chinese Academy of Sciences, Beijing 100101, China;

2. University of Chinese Academy of Sciences, Beijing 100049, China)

Abstract: The accurate estimation of rainfall erosivity (R) is significant for soil erosion prediction. However, the algorithm for the R factor of the universal soil loss equation (USLE), the most widely used algorithm to calculate R , needs pluviograph data with high temporal resolution. Thus, studies have developed a large number of proxies for the R factor (PR_s) using daily, monthly, yearly rainfall data instead of pluviograph data. It is noted that the statistical evolutions of the PR_s is desirable, which is currently lacking. This study aimed to evaluate 15 widely used PR_s , including two yearly, six monthly, and seven daily ones, and further rank them in terms of their correlation with R using a stepwise Meng's test procedure and data collected at 14 rainfall stations in the eastern China. The results indicated that the 15 PR_s under examination were all significantly correlated with R ($r > 0.59$, $p < 0.03$), implying all of them can be reasonably used as an R predictor. The stepwise Meng's test illustrated that the daily PR_s were better correlated with R than the monthly and yearly PR_s , with four daily PR_s , which included three developed in China and one developed in Australia, ranking the first ($r > 0.99$) between the 15 PR_s . Additional examination did show that the Australia PR exhibited a relative error that increased with the latitude and thus should be calibrated by means of the regression equation we developed. One of the three Chinese PR_s did not show a systematic error and can be directly used to estimate R without calibration. The two other PR_s showed approximately 30% overestimation and also be calibrated by the developed regression equations. We also recommended the PR_s to be used if only the monthly or yearly rainfall data is available.

Key words: soil erosion; rainfall erosivity; correlation coefficient; Meng's test; statistical test; simplified estimating model