

基于地理因子的现代降水空间模拟方法探讨

王芳芳¹, 许长军², 陈占寿³, 侯小青¹, 王青波¹

(1.青海师范大学地理科学学院,青海 西宁 810008;

2.青海省地理空间信息技术与应用重点实验室,青海 西宁 810001;

3.青海师范大学数学与统计学院,青海 西宁 810008)

摘要:将经度、纬度、海拔、坡度、坡向、分区(干燥度)和植被等七个因子作为影响现代降水的地理因子,结合 R 语言程序、Arcgis 空间分析和 SPSS 统计分析工具,探讨对现代降水进行空间模拟时,其影响因子与重建方法的选择。基于广义可加模型(GAM)用 R 语言实现了对现代降水影响因子的非线性分析,得到除坡度外其它各因子均与多年平均降水量之间大体呈线性相关;继而对各地理因子进行共线性检验,通过检验的因子使用最小二乘法建立与多年平均降水之间的多元线性回归模型。计算结果表明:逐次叠加地理因子进行回归分析时,其方程的 R^2 均有提升,显示地理因子的类型对降水空间模拟具有一定的影响;非线性和线性建模中坡向因子的 P 值都大于 0.05,说明在本研究中其对降水的影响不显著;线性回归建模的模型拟合优度 R^2 为 0.849,非线性模型的解释达到 89.6%,两种建模结果都具有一定的可信度,对分析今后中国大范围区域中现代降水的影响因子及空间模拟方法具有重要的参考意义。

关键词:地理因子;现代降水;空间模拟;重建方法

中图分类号:P426.615;TP319

文献标识码:A

文章编号:1000-0852(2018)04-0055-07

全球变化(Global Change)与荒漠化(Desertification)是当前世界上重大的环境问题,不仅对气候、植被、土壤、水文及动物等自然环境造成了破坏,还困扰着世界各地人类的生存与发展^[1]。大气降水是“三圈”(大气圈、生物圈和水圈)中至关重要的因子^[2-3]。另外,随着现代化的发展,工业中的各种污染元素伴随现代降水可能波及到地球最偏远地区及整个大气对流层^[4],进而对自然环境造成破坏。准确把握现今降水的规律与特征显得尤为重要,尤其是建立全国尺度的降水空间模拟方程。这不仅有助于了解中国气候变化的规律和机理,而且能为水资源综合规划、国家和地区气候变化适应策略和措施的制定,提供一定的科学依据^[5]。有学者利用 Arcgis 中的主要插值算法对短时段小区域的年均降水进行数据分析和空间建模,同时分析其时空变化及成因^[6-7]。另有研究者在插值方法的基础上再

运用最小二乘法(OLS)对大别山区东段雨季降水进行模拟,以更好地提高山区降水模拟的精度^[8]。而针对那些自变量并非相互独立,而是存在显著多重相关性的影响因子,有研究者改用偏最小二乘法(PLS)构建模型以进行改进^[9]。使用遥感数据也可对区域降水的时空变化规律及空间模拟进行研究,例如,使用 TRMM 卫星降水数据并结合气象站的观测数据,可探索现代年降水量的时空分布规律^[10-11]。

综上,所有对现代降水进行的各种研究内容(如:空间建模及探索时空变化规律),其研究区的选择大都集中在山区、河流等小区域,缺乏对更大地理空间范围的研究;且方法大体集中在使用空间插值、多元统计回归和遥感定量反演三种。这些研究虽然得出较可靠的结论并取得一定的效果,但综合考虑地理研究的空间属性而各有利弊^[12]。卢毅敏等^[13]使用回归拟合、

收稿日期:2017-05-26

基金项目:青海省科技厅自然科学基金项目(2017-ZJ-903);青海省地理空间信息技术与应用重点实验室基金项目(2018-006)

作者简介:王芳芳(1989-),女,河南鹿邑人,在读硕士研究生,主要从事全球变化方面研究。E-mail:wff0207@163.com

通讯作者:许长军(1976-),男,青海平安人,高级工程师,主要从事地理信息与遥感数据分析与应用研究。E-mail:39950625@qq.com

结合 HASM 算法插值并以虚拟站点辅助对全国现代降水进行高精度曲面建模,其所构建模型的结果既突出降水空间分布的总体趋势,也能反映局地细节变化。模拟结果整体比较理想,但其对影响降水的因子选择比较有限。

本文在前人研究的基础上增加影响因子、调整模拟方法,初步选择了经度、纬度、海拔、坡度、坡向、分区和植被共七个地理因子,结合 Arcgis、SPSS 和 R 语言等工具综合运用空间分析、线性和非线性方法对数据进行分析和建立模型。将通过检验的地理因子使用最小二乘法并逐次添加变量来建立回归模型,同时分析模型的 R^2 判断其可信度。以此探讨对现代降水进行空间模拟时,其影响因子与重建方法的选择。

1 数据来源与自变量选择

本文研究的现代降水是指有气象测量资料以来的降水数据变化,文中将影响降水的因子分为基础地理因子经度、纬度、海拔、坡度和坡向;并叠加分区(干燥度)和植被。

(1)地理因子:选取 1951~2009 年全国(不含台湾)722 个气象站点近 60 年的年均降水数据(<http://data.cma.cn/site/index.html>,如图 1),首先计算出多年平均降水量。然后将气象站点的基础数据经度、纬度和海拔等与中国 90×90m 分辨率的数字高程模型(Digital Elevation Model,简称 DEM)相结合(<http://www.gscloud.cn/>),使用 Arcgis10.2 空间分析工具建立中国的经度、纬度、海拔、坡度及坡向矢量和栅格结构的数据库。

(2)植被赋值:本文选择了 1:400 万中国植被图

(数字化版本),图中不仅统计了全国的自然植被,还收录了农业植被,并对每种植被的类型进行详细分类。虽然改革开放以来,种植结构和建设用地,已大大改变了下垫面和植被,但植物的生长和更替具有时段性,故本文植被因子的年份需选择处在降水因子中间阶段的数据,此阶段的植被信息对反映降水具有更好的代表性,获取的数据更能准确的反映环境信息,作为重建降水的因子也更具说服力。该图是 1979 年由中国科学院植物研究所编辑,中国地图社出版,范围覆盖全国主要区域(如图 2)。

将植被数据在 Arcgis 软件中进行投影转换、统一坐标系统并进行空间信息分析,最终提取出数字化的植被信息。然后在对照编码说明的基础上,依据植被与降水之间的自然地理特征对植被进行赋值,将其分为 20 个级别,用数字 10~0 对其进行划分,需水量最高的植被类型赋值为 9.5,每隔 0.5 划分下个级别,需水量最少的为 0(如图 2)。

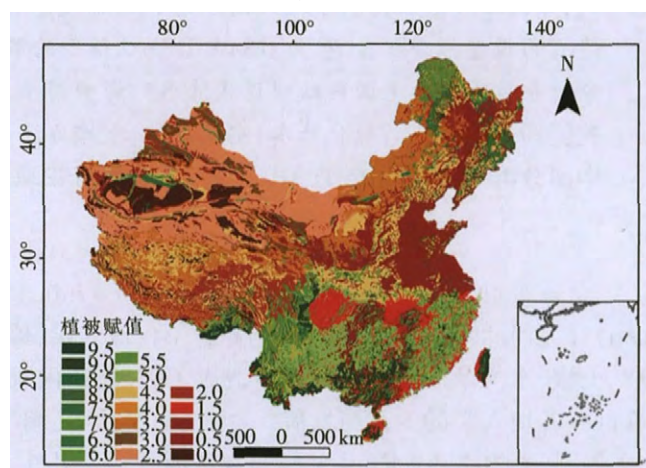


图2 植被分布图

Fig.2 The distribution of the vegetation

(3)区域划分:中国具有特殊的地理位置并拥有各种地形地貌,其自然环境具有一定复杂性,因而影响现代降水的因子也较多。其中季风对降水的影响尤为重要,但因暂时没有合适的量化方法,故文中引用“分区”的方法,对其进行处理。综合考虑时空和分区依据等因素,本文采用陈建伟和张煜星^[14]修正后的气候区划图,该气候区划是在我国传统干燥度分区的基础上

再把前苏联 R.R 谢良尼诺夫经验公式 $K=0.10 \sum t/r$ (式中 K 为干燥度; t 为气温/ $^{\circ}\text{C}$; r 为降水/ mm) 做出适当改进,将式中的经验系数由 0.10 改为 0.16,并使用 16.0 等值线作为极干旱与干旱区的分界。其中干燥

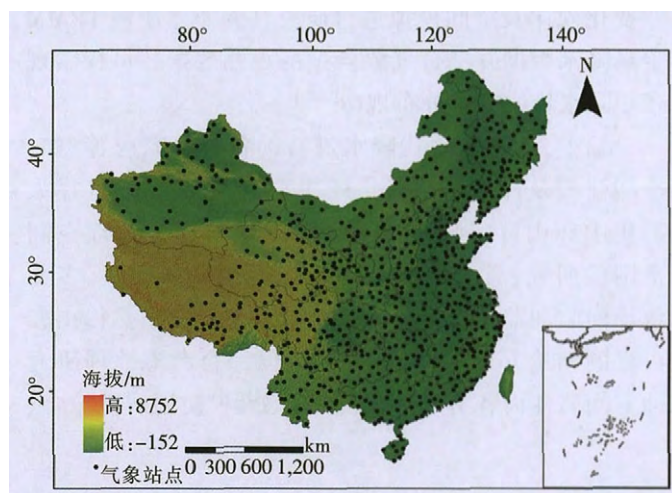


图1 气象站点分布图

Fig.1 The spatial distribution of the meteorological stations

度<1.0的湿润区为一区、干燥度在1.0~1.49之间的干旱亚湿润区为二区、干燥度在1.5~3.99之间的半干旱区为三区、干燥度在4.0~15.99之间的干旱区为四区、干燥度>16.0的极干旱区为五区,根据中国传统的氣候分区特点最终将中国分为5个区域^[14]。

2 研究方法

2.1 广义可加模型

广义可加模型(Generalized Additive Models, GAM)为广义线性模型的非参数化扩展,是基于非参数回归和平滑技术发展起来的一种模型,能将参数和非参数函数加入到模型中,具有较高的灵活性,适用范围非常广泛,其在整体上表现为线性形式,而每一个加项又采用非参数形式^[15-17]。将两种模型结合起来,其表达形式为:

$$g(\mu Y) = \beta_0 + f_1(X_1) + \dots + f_n(X_n) + \varepsilon \quad (1)$$

式中:函数 $g(\dots)$ 为联系函数(link function),在 $g(\mu Y)$ 项中, μY 是 Y 的期望值; β_0 为常数截距项;在 $f_n(X_n)$ 项中, $f_n(\dots)$ 是解释变量 X_n 的单变量函数; ε 为随机变量。为了检验模型中每个因子的可信度及清楚模型的解释度,本研究采用逐步法分析数据,因此GAM模型的具体表达形式如下:

$$P \sim s(X) + s(Y) + s(Z) + s(Pd) + s(Px) + s(Zb) + s(Fq), Fq=1, 2, \dots, 5 \quad (2)$$

式中: s 为自然样条平滑; P 为降水; X 为经度; Y 为纬度; Z 为海拔; Pd 为坡度; Px 为坡向; Zb 为植被; Fq 为分区。

2.2 多重共线性检验

多重共线性检验指示自变量间存在近似的线性关系,即某个自变量能近似的用其他自变量的线性函数来描述,其被广泛应用于多元回归分析的检验中。为了保证所得回归模型有较好的应用效果,本文在利用回归分析时充分考虑了变量与变量之间的关系,在建模之前对各因子进行多重共线性检验,主要方法有以下两种:

①容忍值,亦称容忍度(Tolerance 简记 TOL),在多元线性回归分析中设有 p 个自变量 x_1, x_2, \dots, x_p ,容忍值定义为:

$$TOL = 1 - R_i^2 \quad (3)$$

式中: R_i 为自变量 x_i 与其他 $p-1$ 个自变量间的复相关系数,反映了它们之间的线性相关程度。 TOL 的值在

0~1之间,小于0.1时,就认为这一变量与其他自变量间的多重共线性超过了容许的界限;越接近于1,说明变量间的线性相关性越弱,即能作为解释变量,假若 TOL 很小的变量进入方程,将会导致回归方程的不稳定^[18]。而不同的统计软件定义的 TOL 缺省值也不同,SPSS 中为 0.001,通过这个检验,容忍值小于指定标准的变量将不能进入方程,进而保证进入方程的变量的相关矩阵为非奇异阵,计算结果具有稳定性。

②方差膨胀因子(Variance Inflation Factor简称 VIF),它是诊断多元共线性严重程度的常用指标之一,能检验出模型是否存在共线性,还能判定共线性的强弱。

$$VIF = 1/TOL \quad (4)$$

式中:如果 x_i 与其他 $p-1$ 个自变量间无线性相关(即 $R_i^2=0$),则 $VIF=1$;反之 VIF 大于 1。通常情况下,当 $VIF \geq 5$ 或 $VIF \geq 10$ 时,可认为自变量间存在严重共线性^[19]。

2.3 回归分析

SPSS 是目前非统计专业人员应用最多的统计软件^[20-21]。但直接用 SPSS 系统内部给出的方法进行回归分析时有时难以得到所需要的结果。本文采用“逐步法”进行回归,即首先分别求 y 与 $x_1 \sim x_7$ 的回归方程,得到 7 个基本方程,然后结合地理知识选出一个方程作为基础方程,再逐次叠加其他变量,与此同时根据地理学和统计学两个层面来分析引入的变量是否保存在方程中^[22-23]。

3 结果与讨论

3.1 GAM 模型结果

GAM 模型的分析结果(见图 3)可知:经度与降水之间大体呈正相关,而纬度因子则与降水之间呈负相关,即随着经度的增大降水呈现增加,随着纬度的增大而呈减小的趋势;海拔与降水之间呈不太明显的线性关系,可以看到在 0~1 000m 之间分布的样点较多,降水多集中于此海拔高度;坡度因子与降水之间的关系较为复杂,首先平坦地区降水较为稳定,坡度在 15~20 度之间时降水呈增加趋势,随后降低,在 25~30 度之间时又出现增加趋势;从图上看坡向与降水之间的关系不太明显;植被数值随着降水较多也呈增加的趋势。由此可见,该非线性关系分析结果基本与自然地理状况相吻合,换言之这些地理因子能作为表征降水的变量。其中,坡度与降水之间关系的复杂性,可能是由于我国海拔相差较大,气候条件较为复杂的原因。

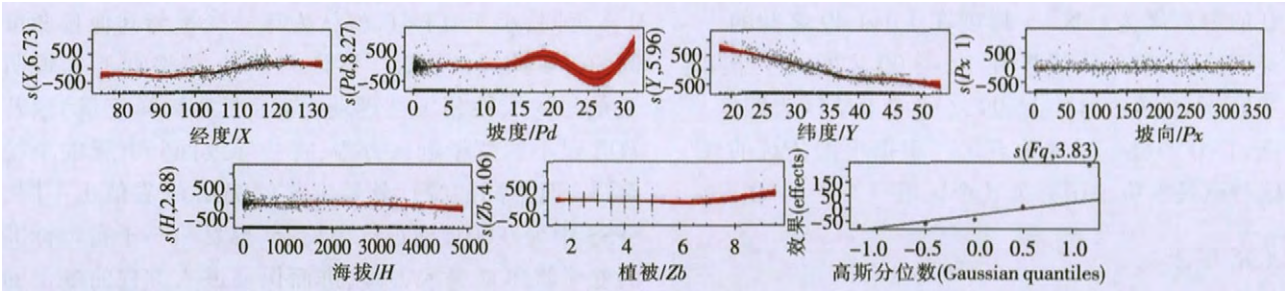


图3 各影响因子与现代降水之间的 GAM 模型分析结果线性关系

Fig.3 The GAM model analysis results between influencing factors and modern precipitation

在地形较为复杂的地区或山谷中其降水要比平缓地区较多^[24],这可能是本文坡度在 25~30 度之间时出现了一段降水增加趋势的原因。文中全国分区的基础是以干燥度为主导的,映射的主要地理信息是季风和蒸发量与降水之间的关系。任一地理区域在不同坡度和降雨强度下的干燥度也会存在不同^[25],将分区作为随机变量分析其它因子时,可能会出现坡度与降水之间这种复杂的关系。

本文主要目的是探讨对现代降水进行重建时,其影响因子有哪些及重建方法的选择,表 1 是广义可加模型非参数的部分估计结果,可以看到该模型对整个函数的解释达到了 89%。七个因子对降水具有不同的显著性,响应有所差异,说明了各地理因子对降水的具有不同的敏感性。其中,除坡向外其它各因子的 P 值均 <0.01 ,呈极显著状态,即能作为降水的解释变量。GAM 模型具有很好的效果,能将多个环境因子综合起来,对影响的重要程度进行定量的评估,得出各地理因子对降水影响的效应图、显著性以及整个模型的效果^[26-27]。本文在 GAM 模型初步分析的基础上,再将各个具有线性相关的地理因子做进一步分析,以期得到具体的模拟方程。

表1 GAM模型的拟合结果

Table1 The results for the generalized additive models (GAMs) fitting to the geographical data

	Ref.df	F	p-value	估值	标准 误差	调整 R^2	解释 偏差
(intercept)			2E-16***	629.07	45.22		
纬度	7.184	89.924	<2e-16***				
经度	7.894	29.551	<2e-16***				
海拔	3.622	4.465	0.00239**				
坡度	8.844	8.288	3.18E-10***			0.891	89.6%
坡向	1.005	0.976	0.32354				
分区	4	28.349	2E-16***				
植被	4.902	6.073	0.0000269***				

注:Ref.df 为模拟参数;intercept 为截距;p-value 表示差异的显著性水平,其中 *** 表示 <0.001 ; ** <0.01 ; * <0.05 ;调整 R^2 表示模型的拟合优度。

3.2 因子分析

文中共有 7 个影响降水的地理因子,因此在进行多元线性回归建模之前,有必要对各因子之间的多重共线性进行检验。首先做出各自变量间的相关系数矩阵(见表 2),以对共线性作初步的判断,任意两因子之间的相关系数 >0.9 ,则表明该变量在后期模型分析时会存在共线性的问题,在 0.8 以上可能会有问题。分析

表2 各影响因子的共线性诊断

Table2 The collinearity diagnosis of the influencing factors

相关矩阵								共线性统计量	
	纬度	经度	海拔	坡度	坡向	分区	植被	TOL	VIF
纬度	1							0.442	2.263
经度	0.129	1						0.438	2.282
海拔	-0.056	-0.639	1					0.525	1.905
坡度	-0.195	-0.136	0.298	1				0.855	1.170
坡向	-0.028	-0.002	0.094	0.092	1			0.981	1.019
分区	0.573	-0.405	0.306	-0.120	-0.019	1		0.405	2.467
植被	-0.590	-0.045	-0.064	0.081	0.025	-0.473	1	0.609	1.642

结果中海拔和经度两个因子之间的相关系数最高(−0.639)但低于 0.8,初步判断 7 个因子之间没有共线性的可能。进而对各因子使用容忍度和 VIF 两种方法进行共线性检验,共线性检验的结果显示:所有因子的 TOL 都小于 1, VIF 均小于 5,整体上通过了检验标准。表明 7 个因子均可以作为因变量的解释变量,该组数据可以使用最小二乘法来建立多元线性回归分析。多重共线性变量的检验还有相应的特征值、条件指数和方差分解比等数据的处理和检验方法,因其所有的方法基本原理较为一致,文中选用的两种方法,是进行偏最小二乘回归常用的检验方法,具有很好的代表性^[28]。

3.3 模型及其解释能力

通过共线性检验的各地理因子分别与降水进行回

归分析得到 7 个基本方程(见表 3),其中降水与纬度和分区之间的关系最高($R=0.749;0.787$),因此在降水和分区基本回归方程的基础上依次再叠加新的解释变量进行运算分析。结果显示:逐次叠加其它地理因子后,不仅提升了方程的拟合优度 R^2 同时也改变了其它因子的回归系数,并且在统计上依旧是显著的。模拟方程可信度的增加,说明整个模型对降水的解释也更为准确。而最后叠加坡向因子时,其方程的 R^2 没有变化,其它回归系数也只有部分存在微小变化,且 sig. 值为 0.936 远大于 0.05。结合 R 语言中 GAM 模型分析,本文综合判定坡向因子对降水的影响不大,将其剔除后模型的 R^2 为 0.849,基本能反映出其它各地理因子对现代降水的影响,故降水空间模拟的最终方程如下(见表 4):

表3 降水回归分析结果
Table3 The results of precipitation regression analysis

因子	基本方程	R	R^2	累加因子	模型	R	R^2
Y	$P=2224.121-45.211Y$	0.749	0.561	$Fq、Y$	$P=2027.675-195.530Fq-26.681Y$	0.865	0.749
X	$P=-789.390+13.44X$	0.353	0.125	$Fq、Y、Zb$	$P=1812.388-189.346Fq-23.783Y+30.454Zb$	0.869	0.754
Z	$P=820.047-0.014Z$	0.384	0.147	$Fq、Y、Zb、Z$	$P=2026.738-140.659Fq-29.668Y+23.758Zb-0.01Z$	0.903	0.816
Pd	$P=632.125+8.319Pd$	0.141	0.02	$Fq、Y、Zb、Z、X$	$P=1212.642-104.583Fq-33.522Y+31.046Zb-0.007Z+7.344X$	0.912	0.832
Px	$P=660.764+0.041Px$	0.009	0	$Fq、Y、Zb、Z、X、Pd$	$P=1204.608-93.536Fq-34.294Y+32.380Zb-0.008Z+7.410X+5.158Pd$	0.921	0.849
Fq	$P=1325.814-289.409Fq$	0.787	0.619	$Fq、Y、Zb、Z、X、Pd、Px$	$P=1204.310-93.535Fq-34.294Y+32.368Zb-0.008Z+7.406X+5.154Pd+0.005Px$	0.921	0.849
Zb	$P=57.001+184.981Zb$	0.576	0.332				

注: P 为降水; Y 为纬度; X 为经度; Z 为海拔; Pd 为坡度; Px 为坡向; Fq 为分区; Zb 为植被; $R、R^2$ 为模型拟合优度。

表4 最小二乘法建模结果
Table4 The fitting results by using least square method

模型	非标准化系数		标准系数		t	Sig.	R	R^2
	B	标准误差	试用版					
(常量)	1204.608	106.051			11.359	0.000		
纬度	-34.294	1.324	-0.567	-25.897	0.000			
经度	7.41	0.824	0.197	8.998	0.000			
海拔	-0.008	0.001	-0.227	-11.351	0.000	0.921 ^a	0.849	
坡度	5.158	0.925	0.088	5.574	0.000			
分区	-93.536	8.315	-0.257	-11.25	0.000			
植被	32.38	6.334	0.095	5.112	0.000			

注: B 表示样本回归系数; t 表示用 T 检验法对方程进行假设检验以说明其有无统计学意义的值;Sig. 表示差异的显著性水平; $R、R^2$ 表示模型拟合优度; a 表示预测变量。

$$P=1204.608+7.410X-34.294Y-0.008Z+5.158Pd-93.536Fq+32.380Zb$$

(5)

4 结论与展望

本文选定经度、纬度、海拔、坡度、坡向、分区和植被七个地理因子,通过 R 语言、Arcgis 和 SPSS 等工具分析出除坡向因子外其它均与降水之间存在较高的相关性,可以作为现代降水重建模型的解释变量,通过多重共线性检验后的各因子,使用最小二乘法逐次添加变量的方法建立现代降水的多元回归模型,最终得到方程 ($P=1204.608+7.410X-34.294Y-0.008Z+5.158Pd-93.536Fq+32.380Zb$, P 为降水; X 为经度; Y 为纬度; Z 为海拔; Pd 为坡度; Zb 为植被),基本结论如下:

(1)剔除坡向因子,建立地理因子与现代降水之间的方程,其模型的 R^2 达到了 0.849,能基本反映因变量现代降水的分布状况。逐次添加解释变量,其模拟

方程的拟合优度存在不同程度的提升, 启示我们影响降水的因子有很多, 需要不断的探索。

(2)因数据和分析方法的有限, 在后续的研究中需对无降水站点的区域进行合理的处理, 文中使用非线性进行分析时, 其模拟的效果比线性的要高, 故需要后期继续进行深入探讨。

参考文献:

- [1] 慈龙骏. 全球变化对我国荒漠化的影响[J]. 自然资源学报, 1994,9(4):289-303. (CI Longjun. The impact of global change on desertification in China [J]. Journal of Natural Resources, 1994,9(4):289-303. (in Chinese))
- [2] 刘春霖. 气候变化对陆地水循环影响研究的问题 [J]. 地球科学进展, 2004,19(1):115-119. (LIU Chunzhen. The issues in the impact study of climate change on the terrestrial hydrological cycle [J]. Advance in Earth Sciences, 2004,19(1):115-119. (in Chinese))
- [3] 符淙斌, 马柱国. 全球变化与区域干旱化[J]. 大气科学, 2008,32(4):752-760. (FU Congbin, MA Zhuguo. Global change and regional aridification [J]. Chinese Journal of Atmospheric Sciences, 2008,32(4):752-760. (in Chinese))
- [4] 效存德, 秦大河, 姚檀栋, 等. 南、北极和青藏高原现代降水中 Pb, Cd 反映的全球大气污染[J]. 科学通报, 1999,44(23):2558-2563. (XIAO Cunde, QIN Dahe, YAO Tandong, et al. The Pb, Cd of modern precipitation reflect global atmospheric pollution in the South and North Pole and the Qinghai-Tibet plateau [J]. Chinese Science Bulletin, 1999,44(23):2558-2563. (in Chinese))
- [5] 任国玉, 任玉玉, 战云健, 等. 中国大陆降水时空变异规律-II: 现代变化趋势[J]. 水科学进展, 2015,26(4):451-465. (REN Guoyu, REN Yuyu, ZHAN Yunjian, et al. Spatial and temporal patterns of precipitation variability over mainland China-II: recent trends [J]. Advances in Water Science, 2015,26(4): 451-465. (in Chinese))
- [6] 韩兰英, 陈全功, 张正偶. 甘肃省近 30 年平均降水空间模拟方法的研究[J]. 干旱区资源与环境, 2008,22(3):108-112. (HAN Lanying, CHEN Quangong, ZHANG Zhengcai. Simulation methods on annual rainfall in recent 30 years in Gansu [J]. Journal of Arid Land Resources and Environment, 2008,22 (3):108-112. (in Chinese))
- [7] 梁天刚, 王兮之, 戴若兰. 多年平均降水资源空间变化模拟方法的研究[J]. 西北植物学报, 2000,20(5):856-862. (LIANG Tiangang, WANG Xizhi, Dai Ruolan. Simulating methods of average annual rainfall on spatial changes [J]. Acta Botanica Boreali-Occidentalia Sinica, 2000,20(5):856-862. (in Chinese))
- [8] 张宏群, 马晓群, 范伟, 等. 基于 GIS 的大别山区东段雨季降水空间分布模拟[J]. 水土保持通报, 2010,30(5):152-157. (ZHANG Hongqun, MA Xiaogun, FAN Wei, et al. The spatial distribution simulation of rainy season precipitation based on GIS in east Dabie mountain area [J]. Bulletin of Soil and Water Conservation, 2010,30(5):152-157. (in Chinese))
- [9] 张正勇, 何新林, 刘琳, 等. 中国天山山区降水空间分布模拟及成因分析[J]. 水科学进展, 2015,26(4):500-508. (ZHANG Zhengyong, HE Xinlin, LIU Lin, et al. Spatial distribution of rainfall simulation and the cause analysis in China's Tianshan mountains area [J]. Advances in Water Science, 2015,26(4):500-508. (in Chinese))
- [10] 刘向军, 黄朝晖, 曾方明, 等. 中国西部及邻区现代年降水时空分布初步研究 [J]. 盐湖研究, 2014,22(2):1-7. (LIU Xiangjun, HUANG Zhaohui, ZENG Fangming, et al. Modern spatial and seasonal distribution of precipitation in western China and its adjacent areas [J]. Journal of Salt Lake Research, 2014,22(2):1-7. (in Chinese))
- [11] 张涛, 李宝林, 何元庆, 等. 基于 TRMM 订正数据的横断山区降水时空分布特征[J]. 自然资源学报, 2015,30(2):260-270. (ZHANG Tao, LI Baolin, HE Yuanqing, et al. Spatial and temporal distribution of precipitation based on corrected TRMM data in Hengduan mountains [J]. Journal of Natural Resources, 2015,30(2):260-270. (in Chinese))
- [12] 孙鹏森, 刘世荣, 李崇巍. 基于地形和主风向效应模拟山区降水空间分布[J]. 生态学报, 2004,24(9):1910-1915. (SUN Pengsen, LIU Shirong, LI Chongwei. Estimation of precipitation using altitude and prevailing wind direction effect index in mountainous region [J]. Acta Ecologica Sinica, 2004,24(9):1910-1915. (in Chinese))
- [13] 卢毅敏, 岳天祥, 陈传法, 等. 中国区域年降水空间分布高精度曲面建模[J]. 自然资源学报, 2010,25(7):1104-1205. (LU Yimin, YUE Tianxiang, CHEN Chuanfa, et al. Surface modelling of annual precipitation in China [J]. Journal of Natural Resources, 2010,25(7):1104-1205. (in Chinese))
- [14] 陈建伟, 张煜星. 湿润指数与干燥度关系的探讨 [J]. 中国沙漠, 1996,16(1):79-82. (CHEN Jianwei, ZHANG Yuxing. Discussion on relation between humidity index and aridity degree [J]. Journal of Desert Research, 1996,16(1):79-82. (in Chinese))
- [15] STEFAN S, MARIA J L. Local polynomial inference for small area statistics: estimation, validation and prediction [J]. Journal of Nonparametric Statistics, 2010,22(5):633-648.
- [16] 欧阳芳, 戈峰. 基于广义可加模型的昆虫种群动态非线性分析及 R 语言实现[J]. 应用昆虫学报, 2013,50(4):1170-1177 (OU YangFang, GE Feng. Nonlinear analysis of insect population dynamics based on generalized additive models and statistical computing using R [J]. Chinese Journal of Applied Entomology, 2013,50(4):1170-1177. (in Chinese))
- [17] W.N. Venables, C.M. Dichmont. GLMs, GAMs and GLMMs: an overview of theory for applications in fisheries research [J]. Fisheries Research, 2004,70:319-337.
- [18] 柳丽, 魏庆珍, 赵树安. 回归分析中多重共线性的诊断与处理[J]. 中国卫生统计, 1994,11(1):5-7. (LIU Li, WEI Qingzheng, ZHAO Shuan. Diagnosis and treatment for collinearity in regression [J]. Chinese Journal of Health Statistics, 1994,11(1):5-7. (in Chinese))
- [19] 丁元林, 孔丹莉, 毛宗福. 多重线性回归分析中的常用共线性诊断方法[J]. 数理医药学杂志, 2004,17(4):299-300. (DING

- Yuanlin, KONG Danli, MAO Zongfu. The commonly collinearity diagnosis method in multiple linear regression analysis [J]. Journal of Mathematical and Medicine, 2004,17(4):299–300. (in Chinese))
- [20] 罗凤明,邱劲彪,李明华,等. 如何使用统计软件 SPSS 进行回归分析 [J]. 软件设计开发, 2004,2:293–304. (LUO Fengming, QIU Jinbiao, Li Minghua, et al. How to use the statistics software of SPSS regression analysis [J]. Software Design and Development, 2004,2: 293–304. (in Chinese))
- [21] 张苏江,陈庆波. 数据统计分析软件 SPSS 的应用(五)—相关分析与回归分析[J]. 畜牧与兽医, 2003,35(9):16–18. (ZHANG Sujiang, CHEN Qingbo. The application of statistical analysis software SPSS(5)—correlation analysis and regression analysis [J]. Animal Husbandry & Veterinary Medicine, 2003,35(9):16–18. (in Chinese))
- [22] 石森昌. 使用 SPSS 进行回归分析应注意的一个问题 [J]. 统计教育, 2006,8:16–18. (SHI Senchang. The regression analysis that should be paid attention to a problem using SPSS [J]. Statistical Education, 2006,8:16–18. (in Chinese))
- [23] 张保法. 经济计量学 (第四版)[M]. 北京: 经济科学出版社, 2000. (ZHANG Baofa. Econometrics (4th Edition)[M]. Beijing: Economic and Scientific Press, 2000. (in Chinese))
- [24] 沈永平, 梁红. 高山冰川区大降水带的成因探讨 [J]. 冰川冻土, 2004,26(6):806–809. (SHEN Yongping, LIANG Hong. High precipitation in glacial region of high mountains in high Asia: possible cause [J]. Journal of Glaciology and Geocryology, 2004,26(6):806–809. (in Chinese))
- [25] 代数,蒋光毅,夏清,等. 坡度和雨强对重庆市黄壤旱坡地产生沙特征的影响[J]. 水土保持学报, 2011,25(4):1–5. (DAI Shu, JIANG Guangyi, XIA Qing, et al. The slope and rainfall intensity impact the characteristics of production runoff and sediment at yellow soil and dry slope land [J]. Journal of Soil and Water Conservation, 2011,25(4):1–5. (in Chinese))
- [26] 牛明香,李显森,徐玉成. 基于广义可加模型的时空和环境因子对东南太平洋智利竹筴鱼渔场的影响[J]. 应用生态学报, 2010,21(4): 1049–1055. (NIU Mingxiang, LI Xiansen, XU Yucheng. Effects of spatiotemporal and environmental factors on the fishing ground of Trachurus murphyi in Southeast Pacific Ocean based on generalized additive model [J]. Chinese Journal of Applied Ecology, 2010,21(4):1049–1055. (in Chinese))
- [27] 余黎,雷相东,王雅志,等. 基于广义可加模型的气候对单木胸径生长的影响研究[J]. 北京林业大学学报, 2014,36(5):22–32. (YU Li, LEI Xiangdong, WANG Yazhi, et al. Impact of climate on individual tree radial growth based on generalized additive model [J]. Journal of Beijing Forestry University, 2014,36(5):22–32. (in Chinese))
- [28] 陶然. Logistic 模型多重共线性问题的诊断及改进 [J]. 理论新探, 2008,15:21–25. (TAO Ran. The diagnosis and improvement of multicollinearity problem based on Logistic model [J]. Statistics and Decision, 2008,15:21–25. (in Chinese))

Discussion on Spatial Simulation of Modern Precipitation Based on Geographical Factors

WANG Fangfang¹, XU Changjun², CHEN Zhanshou³, HOU Xiaoqing¹, WANG Qingbo¹

(1. School of Geography Science, Qinghai Normal University, Xining 810008, China;

2. Geomatics Technology and Application Key Laboratory of Qinghai Province, Xining 810001, China;

3. School of Mathematics and Statistics, Qinghai Normal University, Xining 810008, China)

Abstract: The seven geographical factors, such as longitude, latitude, altitude, slope, aspect, regionalization (dryness) and vegetation, were used as influencing factors of modern precipitation. Using R programming language, Arcgis spatial analysis and SPSS (Statistic Package for Social Science) statistical analysis software, this paper discussed on how to select the influence factors and reconstruction method in spatial simulation of modern precipitation. Based on generalized additive model (GAM) by R language to realize the nonlinear analysis between modern precipitation and its influence factors, and the results were obtained that all the geographical factors are roughly linearly related to the mean annual precipitation except the factor of slope. The multivariate linear regression model was established between the geography factors and the multi-year average precipitation, and these factors were tested in multiple collinearity. The calculation shows that R^2 of model was promoted in successive geography factors regression analysis, which means the type of geographical factors have a certain influence on the precipitation space simulation; the P values of the slope orientation factor in nonlinear and linear models are larger than 0.05, indicating that the influence of slope orientation factor on the precipitation is not significant; R -squared of linear regression model is 0.849 and deviance explanation of non-linear model is 89.6%, so as that the results of the two models have a certain credibility.

Key words: geographical factor; modern precipitation; spatial simulation; reconstruction method