

水平面上太阳辐射空间模拟对比研究

田永中^{1,2}, 朱莉芬^{3,4}, 岳天祥³

(1 西南师范大学资源环境科学学院, 重庆 400715; 2 重庆市城市气象工程技术研究中心, 重庆 401147;

3 中国科学院地理科学与资源研究所, 北京 100101; 4 中国科学院研究生院, 北京 100039)

摘要: 面域的太阳辐射数据是进行其他各项有关辐射研究的基础。由辐射台站的点数据来模拟太阳辐射的空间分布时, 传统的模拟方式都会产生精度与效率之间的矛盾。直接以辐射观测值与其影响因子建立非线性回归模型, 以各因子的面域数据来模拟水平面上太阳辐射的空间分布, 与引入了辐射基数值的模式相比, 模型的误差更小, 复相关系数大于 0.96, 运算效率提高 29~286 倍。当采用模拟的天文辐射经验公式来计算辐射基数时效率仍可提高 2~11 倍。用两种方式的回归模型模拟中国 1km 栅格的水平面太阳辐射, 平均差值仅 2.41%, 而本文模式模拟结果与实测值更为接近。

关键词: 太阳辐射; 天文辐射; 空间模拟; 栅格

中图分类号: P422.11; TP3

1 问题的提出

人类对太阳辐射的研究已有较长的历史, 最初的太阳辐射数据来源于辐射台站的观测值, 而对大范围的面域辐射数据, 则只能通过各种模拟计算取得。

目前计算水平面上太阳辐射的方式, 一是根据某种辐射基数值, 利用影响辐射的相关要素建立模型进行模拟; 二是依据辐射站点的实测资料, 采用一定的空间插值方法直接进行。

第一种方式目前在国内外使用最为普遍, 其基本计算公式为^[1]:

$$Q = Q_0(a + bS) \quad (1)$$

式中, Q_0 为太阳辐射基数值 (或称背景值); S 为某 (几) 种辐射影响因子; a 、 b 是经验系数。利用本方式建立模型, 其基本前提是先计算各测站的辐射基数值。对基数的选取, 朱志辉、翁笃鸣、Liou 等采用天文辐射^[2-4], 王炳忠等主张采用理想大气总辐射^[1], 然而更多的是采用晴天总辐射, 认为它包含了水汽和大气微粒的影响, 最接近于实际总辐射, 反映了地表可能利用的太阳总辐射的极限数量^[1,5,6]。天文辐射可以根据日地之间的相互关系, 推导出理论计算公

式; 理想大气辐射同时还须考虑干洁大气对太阳辐射的消光作用和确定透射比函数值的问题, 计算过程相当复杂; 晴天总辐射考虑的因素更多, 理论计算更为困难, 但由于晴天辐射有实际观测值, 因此有人提出用实际观测的晴天辐射与其影响因素建立相关模型来推算晴天总辐射, 如张炯远采用纬度、高度、水汽含量 3 个要素来建立与观测值的模型^[6]。但是应该看到, 每天的晴天辐射值都不会相同, 而且并非每天都可以测得晴天辐射值, 西南的四川盆地、贵州等地, 一年中晴天本来就少, 因此用晴天辐射作为基数有很大的不确定性。由此可见, 基数值的确定是相当不易的, 而更为重要的是, 当使用模型时, 还必须再次计算各地的相应辐射基数值, 而不同地理位置的点, 其辐射基数是不一样的, 因此运算量很大, 运算效率不高。

第二种方式直接建立在观测值的基础上, 简单易行, 但前提是要有尽可能多的辐射观测站资料, 而且在空间分布上应尽量均匀。我国自 1957 年开始进行辐射观测, 至今也只有 140 多个辐射站的观测记录 (含现已撤销的观测站), 1993 年以前分甲、乙两种站, 最多时有 105 个站; 1993 年后, 全国所有台站开始使用相对误差 $\leq 0.5\%$ 的新型观测仪器, 同

收稿日期: 2004-07-10; 修回日期: 2004-11-16.

资助项目: 国家自然科学基金资助项目 (40371094)。

作者简介: 田永中 (1971-), 男, 重庆人, 博士, 主要研究方向 GIS 与资源环境模型。E-mail: tianyz@lreis.ac.cn

时,测站由甲、乙两级调整为一、二、三级,目前全国共有 98 个辐射站^[7]。对此我们进行了实验性分析,选取了 117 个数据序列相对稳定的测站,对总辐射采用 Kriging 方法进行插值。分析表明,插值结果能很好地反映我国太阳辐射分布的总体格局,但总体上较为粗糙,许多地方与实际情况不符(如青藏高原、塔里木盆地等)。因此,尽管第二种方法以实测的辐射值为依据,方法固然简单,但结果并不理想,“不能合理地揭示太阳辐射的空间分布特征”^[8]。在全国辐射观测站并不多的情况下,它仅可以用来了解全国辐射分布的概貌,不宜用作精确分析。

由此可见,两种方式在效率与精度上存在矛盾。本文将在保证足够精度的情况下,直接建立太阳辐射观测值与其影响要素之间的关系模型,提高模型运算效率。为检验该模型,我们将它与第一种方式建立的模型进行对比研究。

2 模型对比研究

2.1 模型建立

为体现辐射影响因子的单独作用和联合作用,根据式(1),我们建立如下两个回归模型^[1]:

$$Q/Q_0 = b_0 + \sum_{i=1}^m \left(\sum_{j=1}^5 b_{ij} x_i^j + \sum_{k=i+1}^m c_{ik} x_i x_k \right) \quad (2)$$

$$Q = b_0 + \sum_{i=1}^m \left(\sum_{j=1}^5 b_{ij} x_i^j + \sum_{k=i+1}^m c_{ik} x_i x_k \right) \quad (3)$$

式(2)为式(1)的改进模型,式(3)是直接利用辐射观测值与相关要素建立的回归模型。式中, x_i 和 x_k 为第*i*个和第*k*个基本因子(即根据物理分析和统计筛选得到的对因变量有重要贡献的因子),*m*为基本因子总数, b_0 和 b_{ij} 以及 c_{ik} 为待定系数。从式(2)和式(3)中可知,模型中除了基本因子外,还包括它们的非线性组合(每一因子的 2~5 次项和一次交叉乘积,在此称为辅助因子^[1])。

2.2 辐射基数值模拟

为便于研究,式(2)中的 Q_0 采用天文辐射。根据朗伯余弦定律,可以推导出天文辐射日总量的计算公式为^[9]:

$$Q_0 = \frac{I_0}{\rho^2} \int_{-\omega_0}^{\omega_0} \sinh dw$$

$$= \frac{TI_0}{\pi \rho^2} (\omega_0 \sin \varphi \sin \delta + \cos \varphi \cos \delta \sin \omega_0) \quad (4)$$

式中, I_0 是太阳常数, T 是时间周期, φ 为纬度, ρ 是以日地平均距离, R_0 为单位距离时的日地距离, h 为太阳高度角, δ 为赤纬, ω_0 为太阳日落时角。 δ 与 ρ 的计算方法有多种,这里我们利用 Bourges 采用傅立叶级数逼近算法所得到的经验公式来计算^[10,11]。

利用式(4)计算的中国 117 个站的年天文辐射总量,与纬度的相关系数为 1,所作的散点图(略)显示,所有的点几乎在一条曲线上,用三次曲线进行模拟,模型为:

$$Q_0 = 0.0109\varphi^3 - 2.4612\varphi^2 + 15.173\varphi + 13041 \quad (5)$$

式中, φ 为纬度(度)。为检验该模型,我们分别用上述两种方法计算了全国 1km 栅格的年天文辐射总量,统计结果表明,模型平均误差为 0.006%,最大误差仅为 0.08%。且由于这 117 个测站在全国分布比较均匀,包括最北端的漠河和最南端的西沙,因此完全有理由相信可以用这种相关模型来计算全国范围内任意地方的水平面天文辐射年总量,更为重要的是,该模型具有很高的运行效率,是理论模型的 27 倍。该模型是否可以应用于我国周边地区或北半球或全球,有待于进一步考证。

2.3 因子选择

到达地面的太阳总辐射通量,要受很多因子的影响,主要有以下几类:

(1) 天文因子:太阳常数、日地距离、太阳赤纬、时角;(2) 地理因子:纬度、海拔高度、坡度、坡向、周围地形;(3) 气象因子:天空云量、大气中的水汽含量、臭氧含量。

根据数据的易得性和有效性,我们选取了以下三类共 6 个太阳辐射因子:

(1) 年日照时数(x_1 , 0.1 小时)、日照百分率(x_2 , %)和年平均云量(x_3 , 0.1 成),它们反映天空的遮蔽特征;(2) 年均绝对湿度(x_4 , 百帕),它反映空气中的水汽含量;(3) 海拔高度(x_5 , 0.1 米)和纬度(x_6 , 度),它们是对辐射有影响的地理因子。

2.4 数据源与模型变量的确定

(1) 日照时数、云量、绝对湿度来源于国家气象局提供的全国 730 多个地面站的多年观测资料;(2) 年日照百分率的计算可参见文献[12];(3) 太阳辐射来源于国家气象局提供的多年日射台站观测资料,经整理后保留了 117 个测站的数据;(4) 海拔

高度为全国范围内的 1km DEM, 来源于美国地质调查局的 GTOPO30 数据集; (5) 观测站的纬度数据来源于台站信息文档, 1km 栅格上的纬度数据源于对该点地理坐标的提取。

由于式(2)和式(3)中因变量的不同, 上述因子对它们的影响也各不相同, 为此我们先对它们进行相关分析, 结果见表 1;

表 1 辐射因子间的 Pearson 相关系数矩阵

Tab.1 Pearson correlation coefficient matrix of factors affecting solar radiation

	x_1	x_2	x_3	x_4	x_5	x_6	Q/Q_0	Q
x_1	1.00000							
x_2	0.99997	1.00000						
x_3	-0.91955	-0.91835	1.00000					
x_4	-0.69255	-0.69004	0.72323	1.00000				
x_5	0.32574	0.32791	-0.15976	-0.58853	1.00000			
x_6	0.61714	0.61182	-0.74506	-0.77440	-0.00881	1.00000		
Q/Q_0	0.95138	0.95017	-0.83853	-0.75108	0.48170	0.58910	1.00000	
Q	0.84801	0.85055	-0.66017	-0.54922	0.60535	0.24596	0.91714	1.00000

由表 1 可知, 除了高度因子外, 其余各因子与 Q/Q_0 的相关系数都明显比 Q 的相关系数高, 因此若从单因子建模的角度出发, 宜采用式(2)。日照百分率和日照时数与其他各要素之间的相关系数都非常接近, 它们本身之间的相关系数几乎为 1, 这主要是由于我国最南端的测站(三亚)与最北测站(漠河)之间, 年可能日照时数之差仅 59 小时, 约为其平均值的 1.3%。所以在选择因素时, 两者可只取其一, 但若采用日照百分率, 还必须计算各地的可能日照时数, 该值可以通过查表得到, 也可以用 Goudriaan 等人提出的公式进行计算^[12]。为简化运

算, 本文只取日照时数建模。在自由度为 116, 置信概率为 0.99 时, 各因子与 Q/Q_0 或 Q 之间的相关性都是显著的。因此, 对式(2)和式(3), 我们选取除日照百分率外的 5 个因子作为基本因子。

2.5 模型误差对比分析

对式(2)和式(3)两种模式, 分别采用逐步回归和普通最小二乘法两种方式建模, 普通最小二乘法又分别采用基本因子和所有因子建模。各模型的误差与效率见表 2。

从表 2 可以看出, 采用不同的建模方法, 其模

表 2 太阳辐射不同建模方法的误差分析与效率比较

Tab.2 Comparison of errors and efficiency between different models of solar radiation

建模方式	Q			Q/Q_0		
回归方法	逐步回归	普通最小二乘法		逐步回归	普通最小二乘法	
所采用因子	$x_1, x_2, x_3, x_4, x_5, x_6$	基本因子	基本因子与辅助因子	$x_1, x_2, x_3, x_4, x_5, x_6$	基本因子	基本因子与辅助因子
相对误差平均值(%)	3.43	3.44	3.37	3.63	3.55	3.39
相对误差标准差(%)	4.57	4.65	4.31	4.70	4.71	4.34
复相关系数*	0.9614	0.9610	0.9630	0.9726	0.9712	0.9740
运算效率	1	0.8	0.1	0.0037	0.0036	0.0035
				0.09**	0.088**	0.047**

* 为便于比较具有不同因子个数的复相关系数, 表中的相应数据为校正复相关系数。

** 该值是使用天文辐射的模拟公式计算辐射基数值时的运算效率。

型都具有很高的显著性, 相对误差的平均值介于 3.39%~3.63%之间, 标准差介于 4.31%~4.71%之间, 差值并不显著, 但运算效率却差别很大, 最大效率为最低值的 289 倍, 这一方面是由于因子个数的不同, 但更主要的是由于式(2)的建模方式需要计算辐射基数值的缘故。特别值得注意的是, 当我们建立了天文辐射的经验公式, 并用它来计算辐射基数值时, 运算效率提高了 27 倍。从表 2 还可看出, 引入了辅助因子的模型, 其复相关系数都最大, 相对误差平均值及标准差都最小; 而只采用基本因子的模型却几乎相反。

尽管采用式(2)的方式建模时, 其复相关系数都比式(3)的高, 但由于因变量不同, 两者之间缺乏可比性, 因此不能以此认为采用式(2)要优于式(3)。

从相关系数表中可以看到, 部分因子之间的相关系数也非常高, 分析表明, 这些要素之间存在多重共线性(multi-collinearity)。多重共线性的存在,

不但降低回归系数的准确性, 而且使结果有较大的标准差^[13]。并且, 一个合理的回归方程应该只包括显著的因子, 逐步回归可以建立起最优回归方程。表 2 中, 为保证逐步回归模型中因子的可靠性, 显著性水平取 0.01。

式(6)和式(7)是根据式(2)和式(3), 采用逐步回归方法建立的回归模型:

$$Q/Q_0 = 0.23147 + 1.134 \times 10^{-5} x_1 + 4.347 \times 10^{-8} x_1 x_3 - 2.172 \times 10^{-5} x_4 x_6 \quad (6)$$

$$Q = 170292 + 20.73189 x_1 - 0.19171 x_1 x_6 + 0.07212 x_5 x_6 \quad (7)$$

式中, Q 为年太阳总辐射 (0.01 MJ/m^2)。式中的各气象因子通过对全国 730 多个地面气象站的观测值采用普通 Kriging 方法插值生成。图 1 是分别采用式(6)和式(7)进行计算所得的预测值与观测值对比图, 可以看出, 式(7)所得的点略比式(6)的点更接近于对角线。

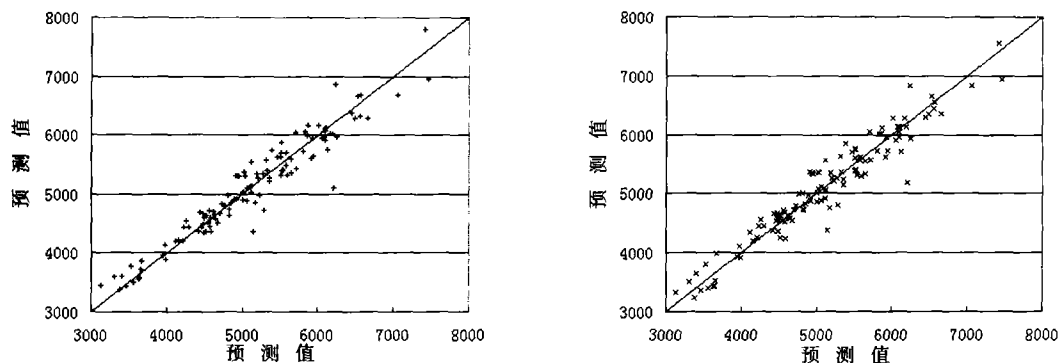


图 1 两种逐步回归模型误差对比(左图源于式(6), 右图源于式(7), 单位: MJ/m^2)

Fig.1 Error comparison of two step-regression models

2.6 模型结果对比分析

用以上两种模型, 我们分别计算了全国 1km 栅格的水平面年太阳总辐射(图 2), 对比分析两种结果可以看到, 虽然两图所反映的全国太阳辐射基本格局是一致的, 但图 2b 更为细腻, 局部的差异反映得很清楚, 特别是对于地表复杂程度差异明显的地区。图 2b 的极值、平均值、标准差都比图 2a 大, 平均值比图 2a 大 1.8%。将图 2b 与图 2a 作差值运算发现, 尽管最大差值达到 28%, 但平均仅 1.06%, 绝对差值百分比的平均值也仅为 2.41%。从差值的地理分布上看, 正值主要出现在大兴安岭、太行山、巫山一线及其以西的山地、高原, 最大差值出现在青

藏高原西北的昆仑山以南地区、天山南北及阿尔泰山等地; 负值主要出现在全国地势低平的地区, 最小值表现为点状分布, 如塔里木盆地东部、西藏的墨脱地区、长江中下游的部分地区等。通过对模型及参与模型运算的原始数据的分析发现, 造成这种差异的原因主要有两个方面: ①式(7)中引入了 DEM 数据, 使得计算结果表现出海拔高度上的差异, 而式(6)中太阳辐射随高度变化的这一特点并未得到体现, 因而它们的差值表现为山地、高原为正值, 平原区为负值的情况; ②式(6)中包括有湿度因子, 湿度数据在插值后表现出了这种点状的低值区。与前人的研究成果及实地观测数据比较, 式(6)

Comparative Study on Spatial Simulation of Solar Radiation on Level Surfaces

TIAN Yongzhong³, ZHU Lifan^{1,2}, YUE Tianxiang³

(1 *Institute of Resource and Environment Science, Southwest Normal University, Chongqing 400715, China;*

2 *Chongqing urban meteorology engineering technology research center 401147;*

3 *Institute of Geographic Sciences and Natural Resources Research, CAS, Beijing 100101, China;*

4 *Graduate School of Chinese Academy of Sciences, Beijing 100039, China)*

Abstract: The traditional modes usually cause the conflict between precision and efficiency when simulating the spatial distribution of solar radiation on level surface. A new mode was brought forward which builds non-linear regression model directly between observed solar radiation and its influencing factors. Comparing the traditional mode with base radiation, its errors are minor and calculation is 29~289 times faster. Stepwise regression model of the two modes are used to simulate the solar radiation in 1km grid-cells of China, the results show that their average difference is 2.41%, and the outcome of the new model is more approximate to the actual data. An experiential model was also built to simulate extraterrestrial radiation on level surface, which made the calculation 27 times faster than that of the theoretic model with almost the same outcome.

Key words: solar radiation; extraterrestrial radiation; spatial simulation; grid

上接 P94

Effects of Information Technology Development on GIS Industry

LIU Li^{1,2,3}

(1 *Institute of Geographic Sciences and Natural Resource Research, CAS, Beijing 100101, China;*

2 *Center for GIS Industry Development, CAS, Beijing 100101, China;*

3 *Graduate School of the Chinese Academy of Sciences, Beijing 100039, China)*

Abstract: This paper firstly analyzes the mechanism of the effects of IT on the GIS industry. Then based on the analysis of IT development trends, it tries to analyze the great effects of the IT development on the GIS industry. The information infrastructure has higher speed and capacity. The IT dynamically and intelligently processes larger quantity of data in real-time. The network and integration have been the emerging characteristics of information technology and applications. The trends of IT development have two different directions, both being more and more complex for professional use and more and more simple for public use. The middle-ware component technology is also the important trend. The IT development exerts great influence on the GIS industry. IT has been the most important innovation diffusion source of the GIS technology. It is also predicted to have deep and multi-aspect effects on the GIS industry in industrial standard, service mode, WebGIS architecture and business mode. The GIS application and industry will be more and more integrated and varied with the IT development.

Key words: IT; GIS industry; trend