Solar radiation modeling based on stepwise regression analysis in China

LU Yimin[1, 2], YUE Tianxiang[1], CHEN Chuanfa[1], FAN Zemeng[1], WANG Qinmin[2]

1. *State Key Laboratory of Resource and Environment Information System*, *Institute of Geographic Sciences and Nature Resources Research*, *Chinese Academy of Sciences*, *Beijing* 100101, *China*;
2. *Key Laboratory of Data Mining & Information Sharing of Ministry of Education*, *Spatial Information Research Center of Fujian Province*, *Fuzhou University*, *Fujian Fuzhou* 350002, *China*

Abstract: In this paper, new solar radiation modeling based on stepwise regression analysis are put forward for estimating global solar radiation from common climate variables (such as sunshine duration, cloud cover, vapor pressure) and geographical elements (altitude, latitude), which simplify the simulation process, improve the operational efficiency under the similar precision. Based on these models and the observation data of common meteorological elements at more than 730 stations in China, the resulting 1km×1km resolution(20226531 grids) solar radiation distribution show that in the whole country, the annual solar radiation energy on the land surface is about 52.4×10^{18} kJ, and the average annual solar radiation lies between 2780—7560$\mathrm{MJ\cdot m^{-2}\cdot a^{-1}}$. There are regional distribution characteristics of global solar radiation in China; it declines from northwest to southeast. The highest value (≥6700 $\mathrm{MJ\cdot m^{-2}\cdot a^{-1}}$) areas of solar radiation is in the Tibet Autonomous Region, the northeastern of Qinghai Province and the west border of Gansu Province; their total area is about 1300000$\mathrm{km^2}$. The lowest value (≤4200 $\mathrm{MJ\cdot m^{-2}\cdot a^{-1}}$) area of solar radiation is in the Sichuan Basin and the gorge area of the Yarlung Zangbo Grand Canyon in the south of Tibetan Plateau; their total area is about 750000$\mathrm{km^2}$.

Key words: global solar radiation, multiple stepwise regression, spatial interpolation, China

CLC number: TP702 **Document code:** A

Citation format: Lu Y M, Yue T X, Chen C F, Fan Z M, Wang Q M. 2010. Solar radiation modeling based on stepwise regression analysis in China. *Journal of Remote Sensing*. **14**(5): 852—864

1 INTRODUCTION

Solar radiation, the fundamental source of energy for life on our planet is the primary influencing factor for ecoclimatic environment. On one hand it is an important parameter of the models in carbon cycle, hydrology, meteorology, energy-saving and emission reduction study areas (Robaa, 2009; Xu *et al*., 2008); on the other hand, 90%—95% of the dry matter in plants are synthesized by photosynthesis process utilized solar energy which is the only energy source of organic nutrients (Yue *et al*., 2008; Trnka *et al*., 2007). However, compared with common meteorological elements (such as air temperature, humidity and precipitation); solar radiation observations are too costly to be measured continuously in many climate stations located in the remote areas. Based on a few observation data of the global solar radiation, it is difficult to reveal the spatial distribution characteristics of global solar radiation, so it is the fundamental work to develop models based on multiple regression analysis to estimate solar radiation for data sparse regions in China with extensive weather records such as sunshine duration, cloud cover, vapor pressure.

In this paper, new solar radiation modeling based on stepwise regression analysis are put forward for estimating global solar radiation from common climate variables (such as sunshine duration, cloud cover, vapor pressure) and geographical elements (altitude, latitude). These models are developed using C# and Matlab mixed programming technique (Yue *et al*., 2007); and the AO component from ESRI was used to visualize the simulation results. Based on Microsoft Dot NET platform the whole system is integrated and deployed without Matlab and ArcGIS desktop. Based on these models and the observation data of common meteorological elements at more than 730 stations in China, the resulting 1km×1km resolution solar radiation distribution is generated and analyzed in the end.

2 MATERIALS AND METHODS

2.1 Data

Meteorological data about monthly solar radiation covering

Received: 2009-08-25; **Accepted**: 2009-12-24

Foundation: China National Science Fund for Distinguished Young Scholars (No.40825003), National Key Technologies R&D Program of China (No.2006BAC08B04, No. 2007BAH16B01), National High-tech R&D Program of China (No. 2006AA12Z219).

First author biography: LU Yimin (1973—), male, doctoral candidates. He is now doing his PhD research in Chinese Academy of Sciences, focusing on resource & environmental modelling. E-mail: luym@lreis.ac.cn

the period from 1957 to 2001 observed in 122 stations and common meteorological elements covering the period from 1951 to 2002 observed in 735 stations were obtained from National Meteorological Information Center in China. The data about sunshine percentage observed in 603 stations were obtained from the Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences. The 1km DEM data of study area were obtained from GTOPO30 dataset provided by United States Geological Survey.

There were some discontinuous periods in meteorological data records, such as the period of monthly solar radiation data records ranging from 10 to 44 years and common meteorological element records ranging from 10 to 50 years; especially some individual new stations having only 1 to 2 years observed records. After rejecting the abnormal value, there were monthly solar radiation data observed in 122 stations, cloud cover data observed in 733 stations and vapor pressure, sunshine duration data observed in 735 stations. The focus of this paper is modelling the annual amount of solar radiation in China, so the key to processing data is how to obtain the most representative annual amount of solar radiation in stations. Monthly mean meteorological data were calculated from every month record in the past years. Then we accumulated monthly mean solar radiation and sunshine duration value and calculated the weighted mean value of cloud cover and vapor pressure in 12 months a year.

2.2 Methodology

2.2.1 Traditional methods

At present, there are mainly three kinds of methods on the global solar radiation estimation. The first one is to estimate the global solar radiation based on the spatial interpolation methods. It is one of the simplest ways, however, the simulation results are difficult to satisfy for practical application for solar radiation observation data sparse regions in China. The second is the remote sensing retrieval of solar radiation at the surface (Wei *et al.*, 2003; He *et al.*, 2004). The third is a numerical climatological method. Although it is mature, the method is rather complicated to apply in large-scale study areas, such as the whole country of China. It is difficult to determine the cardinal number of global solar radiation, empirical coefficients *a* and *b* for their spatiotemporal variation in national scale (Ju *et al.*, 2005). Based on stepwise regression analysis, we put forward a new solar radiation modeling for estimating global solar radiation from common climate variables and geographical elements, which simplify simulation process, improve the operational efficiency under the similar precision.

2.2.2 Method based on stepwise regression

There are many factors that influence annul global solar radiation (Q). Generally, the more the influencing factors are used in regression equation the higher the value of explained sum of squares and the lower the value of residual sum of squares will be. As multicollinearity exist among these factors, the parameters estimation based on traditional methods are not satisfactory. Models containing more correlating parameters may suffer from the defect of collinearity, and its reliability and accuracy decrease accordingly. Likewise, containing inferior variables and omitting important ones also lead the parameters estimation to be biased and inconsistent. We therefore carried out multiple stepwise regression analysis with an aim of obtaining most appropriate model for modelling global solar radiation under present data collation and compilation.

There are mainly three types of methods used to overcome multicollinearity under present study. One is "difference method"; the other two are decreasing variance of parameter estimator and excluding the variables caused by collinearity. Deferent from the other two methods, the third method omitting out all variable that contribute the least to the model therefore has exerted a greater effect. Usually, there are four kinds of methods to establish an optimal regression equation by excluding variables lead to multicollinearity. The first one is to fit every possible model, accordingly to certain criteria (such as Cp, AIC). If K potential predictor variables, there are 2^K-1 models. So, it is difficult to select the model with the fewest predictors from the model sets (including 2^K-1 models). The second is top down approach (Backward Elimination). Begin with a model that includes all of the predictors; remove the variable that contributes the least to the model and is insignificant. If the variable is highly correlated, the regression equation may not get the correct results. The third is bottom up approach (Forward Selection). Begin with no predictor variables; add predictors one at a time according to which one will result in the largest increase in R^2; stop when R^2 will not be significantly increased. The last is Combines Forward/Backward (Stepwise Regression). Select two thresholds F_{SLS} and F_{SLE}. Starting like Forward Selection, add new variable, if it has $F \geqslant F_{SLS}$. Re-test all "old variables" that have already been added, and retain the old variable only if $F \leqslant F_{SLE}$. Continue until no new variables can be entered and no old variables need to be removed. Complicated as it is, the Stepwise Regression is an ideal method to overcome multicollinearity. Thus, it is widely used in geology (Tian *et al.*, 2005; Fu *et al.*, 2009), ecology (Petersen & Stringham, 2008), meteorology (Ertekin & Evrendilek, 2007), materials (Wang *et al.*, 2007), medicine (Ng, 2003) and other research fields.

Stepwise regression procedures work in an alternating order. Begin with no variables; add variables one at a time according to which one will result in the largest increase in R^2. At each step remove any variable that does not explain a significant portion of variance. Stop when R^2 will not be significantly increased. No matter what variable of regression equation is added or dropped, it needs to check after each step if the significance of other variables has changed. Process repeated (Fig. 1) enter the variable with the highest correlation with y-variable first; remove variables that become insignificant due to other variables being added. And the formula of F statistics is shown as follows.

$$F = \frac{U/k}{Q_e/(n-k-1)} \sim F(k, n-k-1) \tag{1}$$

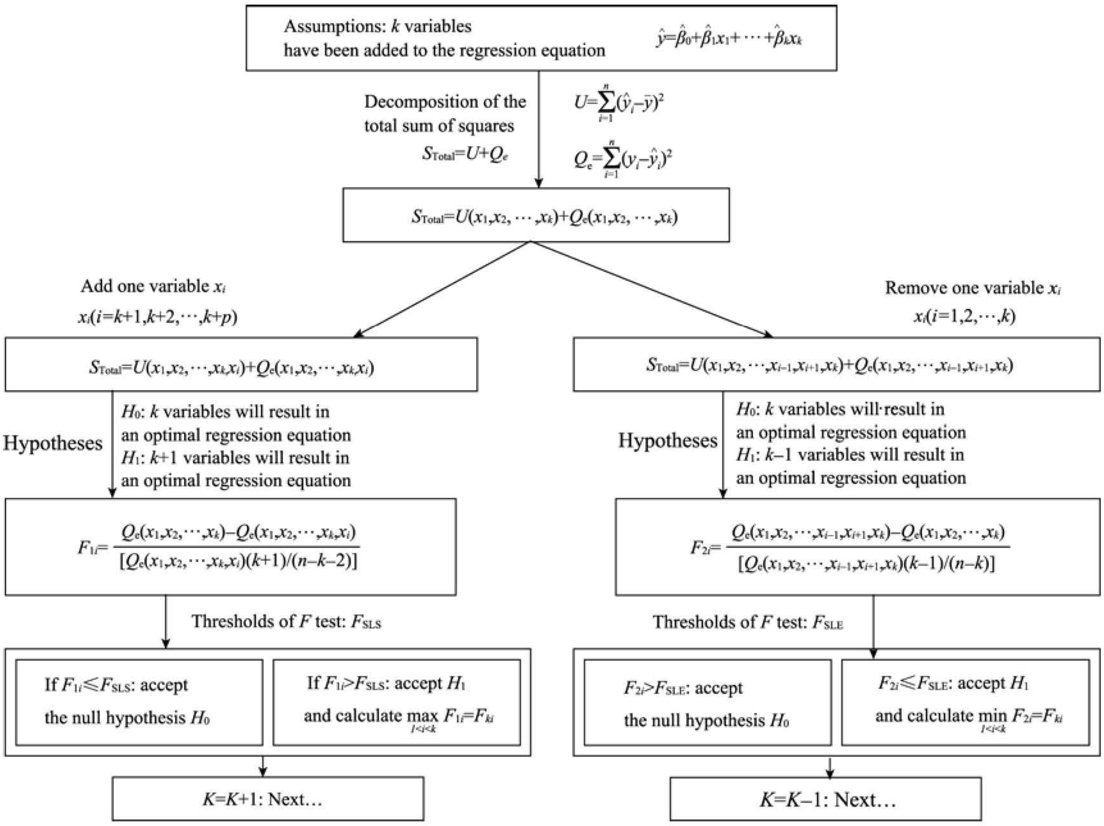

Fig. 1 Variables entered/removed in multiple stepwise regression

3 RESULTS AND DISCUSSION

3.1 Correlation analysis of influencing factors

Scatter Diagram, one of many statistical techniques usually used to measure the strength of the relationship between two variables is visual but not precise enough. The coefficient of correlation, r is a measure of the strength of the linear relationship between two variables. Method of estimating r can portray the correlation between the two variables of interest precisely.

Considering the procurability and integrity of data about influencing factors of solar radiation ($0.01\text{MJ}\cdot\text{m}^{-2}\cdot\text{a}^{-1}$), we select six variables, which are x_1 annual mean sunshine hours (0.1 h), x_2 annual mean sunshine percentage (%), x_3 annual mean total cloud cover (0.01), x_4 annual mean vapor pressure (hPa), x_5 altitude (0.1 m) and x_6 latitude (°). To acquire Pearson correlation coefficient, which measure and interpret the association between variables or solar radiation (Q) and variable, we select 96 stations (from 122 stations solar radiation observed) which have integrated records about six variables aforementioned. The correlation terms involved in the correlation coefficient matrix (Table 1), therefore, indicate that the association between solar radiation (Q) and each variable is significant. The term of x_1 and x_2 is 0.9963 close to one, namely, collinearity exists between x_1 and x_2. Thus, one of them needs to be removed. Here dropping x_2 annual mean sunshine percentage, other five variables will be use to establish the regression equation.

Table 1 Pearson correlation coefficient matrix of influencing factors for solar radiation (N=95)

	x_1	x_2	x_3	x_4	x_5	x_6	Q
x_1	1.0000						
x_2	0.9963	1.0000					
x_3	−0.9158	−0.9189	1.0000				
x_4	−0.6677	−0.6525	0.7214	1.0000			
x_5	0.1824	0.1731	−0.0470	−0.4936	1.0000		
x_6	0.6306	0.6157	−0.7748	−0.7899	−0.0068	1.0000	
Q	0.8456	0.8435	−0.6446	−0.5119	0.4276	0.2347	1.0000

3.2 Solar radiation modeling

3.2.1 Variables of solar radiation modeling

In order to ensure predictors reliable and to promote the model in interpreting the estimation results more precisely, a significance level selected in stepwise regression is 0.01. Moreover, root mean square error (RMSE, $\text{MJ}\cdot\text{m}^{-2}\cdot\text{a}^{-1}$), multiple coefficient of determination (R^2), adjusted R^2 (Adj–R^2) and F statistics are used as evaluation indexes in the regression diagnostics. The root mean square error (RMSE) is an index of discrepancy between the exact and approximate and is a frequently-used measure of the differences between values predicted by a model and the values actually observed from the thing being estimated. The coefficient of determination (R^2, goodness of fit) is the proportion of explained sum of squares to total sum of squares. It provides a measure of how well future outcomes are likely to be predicted by the model. Remember that R^2 may overestimate the true amount of variance explained; thus, adjusted R^2 compensates by reducing the R^2 according to the ratio of subjects per predictor variable.

Data observed in 116 stations which have integrated records about aforementioned five variables, x_1 annual mean sunshine hours, x_3 annual mean total cloud cover, x_4 annual mean vapor pressure, x_5 altitude, x_6 latitude and Q solar radiation, will be used to establish the solar radiation model by the stepwise regression method. According to the coefficients of correlation between solar radiation and each variable (Table 1), the variable is added or removed, and to be checked after each step if the significance of other variables has changed. Based on Tian *et al*. (2005), it is assumed that three variables (x_1, x_5, x_6) have been added to the regression equation. Then the variables (x_4, x_3) are added and variable x_5 is removed, meanwhile, evaluation indexes in the regression diagnostics have been recorded in Table 2. The data presented in Table 2 indicates that R^2, Adj–R^2 increase and RMSE decrease, while variables x_4, x_3 have been added to the regression equation. While variable x_5 has been removed, R^2, Adj–R^2 and RMSE have not been changed remarkably but F statistics increase obviously. Therefore, four variables x_1, x_3, x_4, and x_6 are retained for the final solar radiation modeling.

Table 2 Summary statistics updated with each step for the solar radiation model based on MSRA

	x_1, x_5, x_6	x_1, x_4, x_5, x_6	x_1, x_3, x_4, x_5, x_6	x_1, x_3, x_4, x_6
RMSE	274.001	264.591	257.307	257.770
R^2	0.910927	0.917682	0.922853	0.921871
Adj–R^2	0.908541	0.914715	0.919346	0.919056
F	381.799	309.356	263.169	327.432

3.2.2 Coefficient of solar radiation modeling

After four predictor variables selected, the parameters need to be determined. The estimation of parameters for the solar radiation model is carried out here using mean square error (MSE), relative mean square error (R-MSE) and computational efficiency (CE) as evaluation indexes in the error analysis. If the sample sizes are large enough, its mean square error is the same as its standard deviation, and its relative mean square error is the proportion of the absolute value of mean square error to the corresponding observed value.

$$Q_{\mathrm{Ref}} = 170292 + 20.73189x_1 - 0.19171x_1 x_6 + 0.07212x_5 x_6 \quad (2)$$

$$Q_{\mathrm{Line}} = 299608.16779+18.20611x_1+1976.62392x_3-609.31017x_4-7310.72926x_6 \quad (3)$$

$$Q_{\mathrm{Int}} = 355453.51165+17.67786\, x_1-2057.18878\, x_3-355.29889\times x_4+767.34475\, x_6+0.11365\, x_1 x_3-0.01519x_1 x_4-0.235735\, x_1 x_6 +8.54718\, x_3 x_4-10.11656\, x_3 x_6-19.90513\, x_4 x_6 \quad (4)$$

Here we use variables x_1, x_3, x_4, x_6 and Q to build multiple linear regression Eq. (3) and binomial regression Eq. (4); and quote Eq. (2) from the reference (Tian *et al*., 2005) to compare with them. The scatter plots of observed and simulated annual global solar radiations is shown in Fig.2, in which (a) comes from the Eq. (2) model; (b) comes from the Eq. (3) model; (c) comes from the Eq. (4) model. Comparison of errors and efficiency among different models of solar radiation are shown in Table 3.

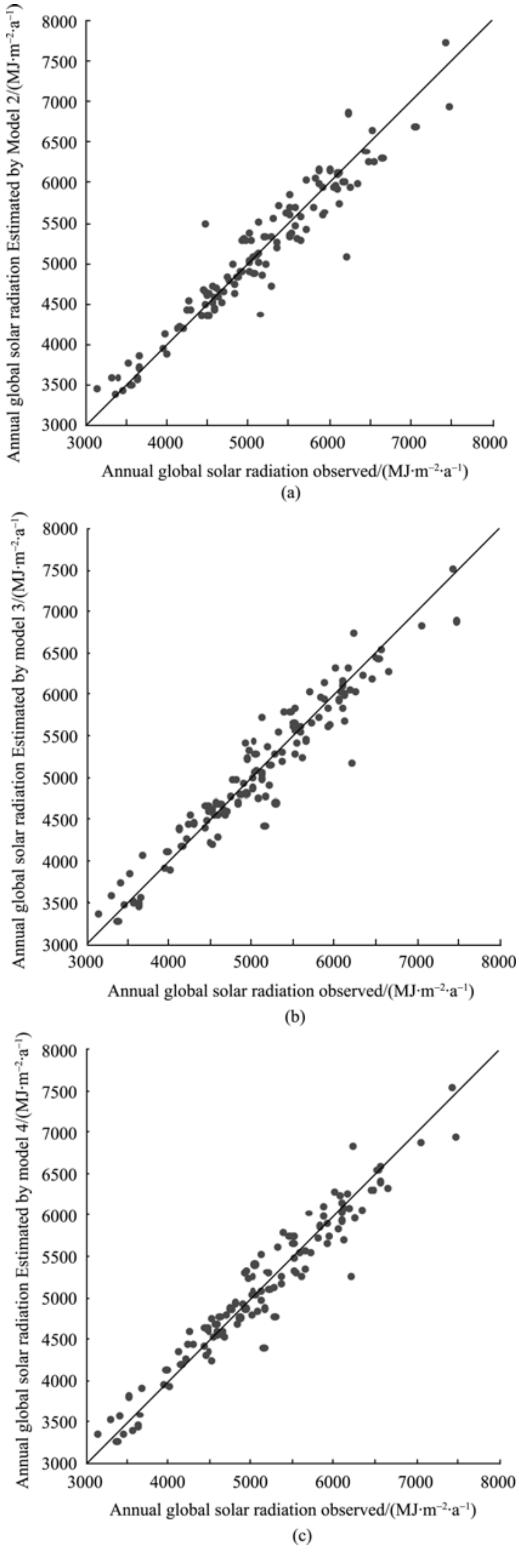

Fig. 2 Scatter plots of observed and simulated annual global solar radiations

(a) Estimated by Eq.(2); (b) Estimated by Eq.(3); (c) Estimated by Eq.(4)

Table 3 Comparison of errors and efficiency among different models of solar radiation*

	model 2	model 3	model 4
MSE	297.68	252.15	234.54
R-MSE	0.0580	0.0491	0.0457
CE	0.800403	1	0.294891

*MSE, mean square error ($MJ{\cdot}m^{-2}{\cdot}a^{-1}$); R-MSE, relative mean square error and CE is computational efficiency, mean value of the 18 samples with different sizes.

According to Fig.2, it shows that Fig.2 (c) from model 4 having lowest value of mean square error (234.54) and indicating it to be the most precise model Fig.2 (b) is better than Fig.2 (a). The relative mean square errors (R-MSE) presented in Table 3 shows that all the three models are reasonable, and that R-MSE of model 3 and model 4 decrease by 15.3% and 21.2% comparing with model 2. At the same time, the computational efficiency (CE) presented in Table 3 shows that CE of model 2 and model 4 decrease by 20% and 70.5% comparing with model 3. Based on the aforementioned analysis, it indicates model 3 to be the best model among the three models discussed above. So model 3 is finalling used to estimate the annul global solar radiation in China.

3.2.3 Simulation of solar radiation modeling

Two different methods are used to make the required estimates of global solar radiation in China. One is spatial interpolation based on 122 stations' data (global solar radiation observation record); the other is spatial interpolation based on 731 suppositional stations' data of global solar radiation, which are simulated by Eq. (3). Fig.3 (a) (1km×1km resolution, 4173×4847 grid) resulting from the first method shows that the spatial distribution of stations is uniform and the spatial distribution of global solar radiation is simulated generally and simply. However, it is difficult to reveal the spatial distribution characteristics of global solar radiation reasonablely, based on so few observation data in the whole country. Large errors appear in the result of many regions such as Qilian Mountains, Tarim Basin and southern Tibetan Plateau.

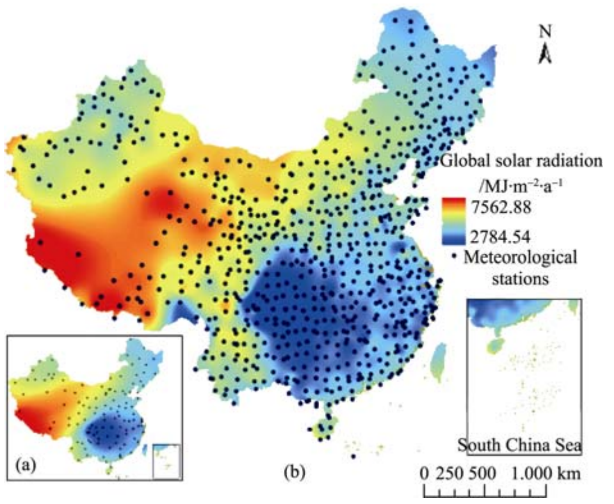

Fig. 3 Simulation results of annual global solar radiations in China

Compared with Fig.3 (a), Fig.3 (b) has a spatial resolution of 1km×1km, and a size of 4173×4847 grid. Its extreme values observed against predicted ones for the same locations are unanimous. Moreover, the values estimated in Fig.3 (b) are closer to the observed (true) solar radiation measurements in the whole study area. It implies abundant information and portrays more details in the spatial distribution of global solar radiation. Classification map of annual global solar radiations in China is shown as Fig. 4. In western China with sparse measured sites, however, these results still leave much to be desired.

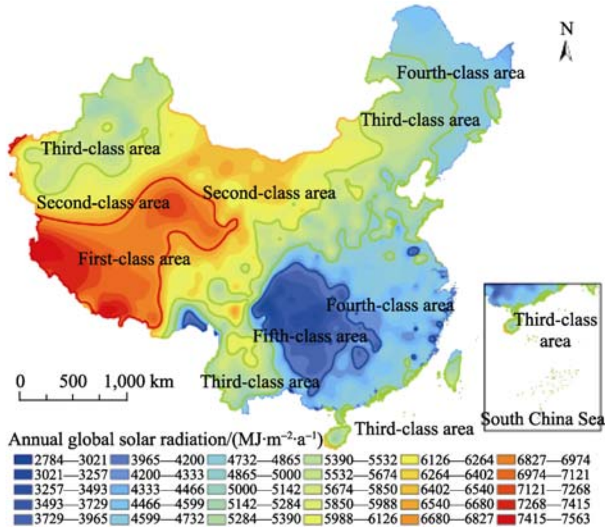

Fig. 4 Classification map of annual global solar radiations in China

4 CONCLUSIONS

(1) There are about 52.4×10^{18} kJ solar radiation energy (equivalent to 1790 billion tons of standard coal) to be obtained on the whole land surface of China. The spatial distribution of annual global solar radiation shows the regional characteristics obviously; as is abundant in western and northwestern China, and is scanty in the east and southwest China.

(2) The Tibet Autonomous Region, northeastern Qinghai and western borders of Gansu, whose areas accumulated is about 1.3 million km^2, are first-class areas. As shown in Fig. 4, annual global solar radiation energy in these areas reaches up to $6680MJ{\cdot}m^{-2}{\cdot}a^{-1}$. The lowest value area of solar radiation ($\leqslant4200MJ{\cdot}m^{-2}{\cdot}a^{-1}$) is in the Sichuan Basin and the gorge area of the Yarlung Zangbo Grand Canyon in southern Tibetan Plateau. They are fifth-class areas, whose total area is about 0.75 million km^2.

(3) The second-class areas of solar radiation (5850—$6680MJ{\cdot}m^{-2}{\cdot}a^{-1}$), whose total area is about 2.4 million km^2, are in the southern Inner Mongolia, northern Shanxi and Ningxia, the middle and northwest of Gansu, eastern Qinghai, the southeast of Tibet and southern Xinjiang. The third-class areas (5000—$5850MJ{\cdot}m^{-2}{\cdot}a^{-1}$), mainly distribute in the northeast of Inner

Mongolia, western Jilin and Liaoning, the southeast and north of Hebei, southeastern Shandong, southern Shanxi, northern Henan and Shaanxi, the southeast of Gansu, northern Xinjiang, Yunnan, southern Guangdong, southern Fujian and the southwest of Taiwan, whose area is about 2.8 million km^2. The fourth-class areas ($4200 - 5000 MJ\cdot m^{-2}\cdot a^{-1}$) are mainly in Heilongjiang, the northeast of Jilin and Liaoning, southern Shaanxi, Hunan, Hubei, Jiangsu, Anhui, Jiangxi, Zhejiang, Guangxi, northern Guangdong, northeastern Fujian and the north of Taiwan, whose total area is about 2.35 million km^2.

(4) A method put forward in this paper provides a new idea to integrate remote sensing information in agroclimatological resources, spatial interpolation or surface fitting. Based on the geographic regression method, using continuous relevant information implicated in satellite images, we are able to estimate the distribution of meteorological elements by the inferential measurement. So, it can improve the simulation results in data sparse regions in China.

In summary, based on stepwise regression analysis we put forward a new model to estimate solar radiation for data sparse regions in China. Modeling with common meteorological elements, considering the geographical factors influence, herein it has exerted a greater effect in simulating solar radiation continuously. It is not only able to reveal the distribution of the annual solar radiation, but also can provide a satisfactory numerical simulation. Thus, we suggest that the simulation results should be applied widely in agro-ecological regionalization, food provision, vegetation NPP, energy-saving and emission reduction and new energy development study areas and so on. Obviously, in addition to astronomical, geographical and atmospheric elements, solar radiation is also influenced by local topography. In the future, we will improve the effect of simulation taking into account the influence of terrain aspect and unobstructed factor, especially in regions with sparse measured sites.

REFERENCES

Ertekin C and Evrendilek F. 2007. Spatio-temporal modeling of global solar radiation dynamics as a function of sunshine duration for Turkey. *Agricultural and Forest Meteorology*, **145**(1/2): 36—47

Fu Y, Wang L J, Chai F M, Deng G, Chen Y and Chen J H. 2009. The application of multi-variate linear regression combined with stepwise regression analysis for Cu-Ni sulfide deposit of Baishiquan. *Earth Science Frontiers*, **16**(1): 373—380

He H L, Yu G R, Liu X A, Su W, Niu D and Yue Y Z. 2004. Study on spatialization technology of terrestrial eco-information in China (Ⅱ): Solar radiation. *Journal of Natural Resources*, **19**(5): 679—687

Ju X H, Tu Q P, and Li Q X. 2005. Discussion on the climatological calculation of solar radiation. *Journal of Nanjing Institute of Meteorology*, **28**(4): 516—521

Ng C M. 2003. Comparison of neural network, Bayesian and multiple stepwise regression-based limited sampling models to estimate area under the curve. *Pharmacotherapy*, **23**(8): 1044—1051

Petersen S L. and Stringham T K. 2008. Development of GIS-based models to predict plant community structure in relation to western juniper establishment. *Forest Ecology and Management*, **256**(5): 981—989

Robaa S M. 2009. Validation of the existing models for estimating global solar radiation over Egypt. *Energy Conversion and Management*, **50**(1): 184—193

Tian Y Z, Zhu L F, and Yue T X. 2005. Comparative study on spatial simulation of solar radiation on level surfaces. *Geo-information Science*, **7**(4): 95—100

Trnka Miroslav, Eitzinger Josef, Kapler Pavel, Dubrovský Martin, Semerádová Daniela, Žalud Zdeněk and Formayer Herbert. 2007. Effect of estimated daily global solar radiation data on the results of crop growth models. *Sensors*, **7**(10): 2330—2362

Wang Z Q, Chai S X, Zhong X M, Guo Y, Wang X Y and Wei L. 2007. Multivariate stepwise regression method use in correlation analysis of microstructure indices and strength of solidified soil. *Rock and Soil Mechanics*, **28**(8): 1650—1654

Wei H L, Xu Q S Zhang T S. 2003. Observation of solar irradiance at the surface from GMS-5. *Journal Of Remote Sensing*, **7**(6): 466—471

Xu S Y, Peddle D R, Coburn C A. and Kienzle S. 2008. Sensitivity of a carbon and productivity model to climatic, water, terrain, and biophysical parameters in a Rocky Mountain watershed. *Canadian Journal of Remote Sensing*, **34**(3): 245—258

Yue T X, Du Z P and Song D J. 2007. High accuracy surface modelling: HASM4. *Journal of Image and Graphics*, **12**(2): 343—348

Yue T X, Tian Y Z, Liu J Y, and Fan Z M. 2008. Surface modeling of human carrying capacity of terrestrial ecosystems in China. *Ecological Modelling*, **214**(2/4): 168—180

中国太阳总辐射的多元逐步回归模拟

卢毅敏[1,2]，岳天祥[1]，陈传法[1]，范泽孟[1]，王钦敏[2]

1. 中国科学院 地理科学与资源研究所，资源与环境信息系统国家重点实验室，北京 100101;
2. 空间数据挖掘与信息共享教育部重点实验室，福州大学省空间信息工程研究中心，福建 福州 350002

摘　要： 研究采用年日照时数、年均总云量、年均水汽压、维度等常规气象要素及地理要素为解析变量，建立起估算全国范围内太阳年辐射总量的多元逐步回归模拟模型；利用全国 730 多个地面站 1951—2002 年间的常规气象观测资料，通过所构建的模型对中国太阳年辐射总量进行模拟分析，模型模拟分辨率为 1km×1km，规模为 4173×4847 格网。模型数值计算结果显示：中国太阳辐射资源最丰富的地区为西藏自治区、青海东北部及甘肃西部边境等，高达 6700$MJ \cdot m^{-2} \cdot a^{-1}$ 以上(一类地区)，总面积约 130 万 km^2；最低的为四川盆地、青藏高原南部的雅鲁藏布峡谷地区，低于 4200$MJ \cdot m^{-2} \cdot a^{-1}$(五类地区)，总面积约 75 万 km^2；中国陆地表面每年接受的总太阳辐射能约为 52.4×10^{18} kJ；单位面积上太阳年辐射总量约介于 2780—7560$MJ \cdot m^{-2} \cdot a^{-1}$；分布具有明显的地域性，呈现西北、北部高，东南、东部及东北的部分地区低的特点，模拟值较好拟合实际的太阳年辐射总量。研究所构建模型实现了太阳年辐射总量空间分布的连续模拟，模型在保证足够的模拟精度条件下，提高了运算效率，简化了模拟过程，适于全国尺度下太阳年辐射总量的数值模拟计算。

关键词： 太阳总辐射，多元逐步回归，空间插值，中国

中图分类号： TP702　　**文献标志码：** A

引用格式： 卢毅敏，岳天祥，陈传法，范泽孟，王钦敏. 2010. 中国太阳总辐射的多元逐步回归模拟. 遥感学报, **14**(5): 852—864
Lu Y M, Yue T X, Chen C F, Fan Z M, Wang Q M. 2010. Solar radiation modeling based on stepwise regression analysis in China. *Journal of Remote Sensing*. **14**(5): 852—864

1　引　言

太阳辐射是地球表面的主要能量来源，一方面它是形成一定气候生态环境的重要因素，是生态系统碳循环模型、水文及气象模拟模型，清洁能源及节能减排等研究中的必不可少参数(Robaa, 2009; Xu 等, 2008)。另一方面太阳辐射与植物的生长有多方面的直接关系，植物体的干物质有 90%—95%系利用太阳能通过光合作用合成；太阳辐射是绿色植物通过光合作用制造有机养分的唯一能源 (Yue 等, 2008; Trnka 等, 2007)。因此模拟分析太阳辐射时空分布特征具有重要的科学价值和现实意义。然而，太阳辐射的观测由于成本费用高，与温度、降水等常规气象要素相比，只有极少数气象站进行太阳辐射观测，观测时期较短，分布集中在大中城市。利用已有的太阳辐射观测值，直接进行空间内插或外推技术不可能合理地揭示太阳辐射的空间分布特征。相对而言，中国气象部门的基准地面气象站较多，且对日照时数、总云量、水汽压等常规气象因素进行了多年观测，获取了多要素、长时间系列的数据。因此为了模拟广大无辐射观测地区的太阳年辐射总量，需要根据有辐射观测地区的资料建立太阳年辐射总量的多元回归模拟的最优模型。

本文采用年日照时数、年均总云量、年均水汽压、维度等常规气象及地理要素为解析变量，建立起太阳年辐射总量多元逐步回归模拟模型。利用全国 730 多个地面站的 1951—2002 年间的常规观测资料，通过所构建的模型对中国太阳年辐射总量进行模拟分析。在模型的实现中，利用 C#与 Matlab 混合编程实现空间插值模型(岳天祥等, 2007)、多元逐步回归模型等；并基于 ESRI 的 AO 组件，实现模型模拟结果的可视化。最终在.NET 编程环境中，对模型

收稿日期：2009-08-25；修订日期：2009-12-24

基金项目：国家杰出青年科学基金(编号：40825003)；国家科技支撑计划课题(编号：2006BAC08B04，2007BAH16B01)；国家高技术研究发展计划(编号：2006AA12Z219)。

第一作者简介：卢毅敏(1973—)，男，福建省，博士研究生，研究方向为资源环境数学模型与系统模拟。E-mail: luym@lreis.ac.cn。

进行构件化封装, 完成整个模型系统脱离 Matlab 及 ArcGIS 环境的开发集成。

2 数据与方法

2.1 数据来源及处理

本研究所用数据主要涉及观测站位置及高程等基础地理资料、太阳辐射观测资料、气温、云量等常规气象观测资料等。

2.1.1 数据来源

全国 122 台站(各省、自治区及直辖市), 1957—2001 年间太阳辐射月观测资料, 全国 735 个地面站的 1951—2002 年间的常规气象要素月观测资料, 气象观测站位置的经纬度数据等来源于国家气象信息中心气象资料室。日照百分率数据来源于中国科学院地理科学与资源研究所“中国自然资源数据库”中 603 个站, 累年各月日照百分率(分台站)资料。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。

2.1.2 数据分析处理

太阳辐射月观测数据有缺测情况, 观测记录出现间断点的年份, 122 台站记录年限从 10—44a 不等。常规气象要素逐月观测资料, 全国 735 地面台站记录年限从 10—50a 不等, 个别新建台站仅有 1—2a 记录。经过数据分析, 剔除了极少数观测有疑问的样本数据, 经整理后保留了 122 测站的太阳总辐射月值数据, 733 测站总云量的月值数据, 735 测站的水汽压、日照时数月值数据。

本研究的目标是建立起估算全国范围的太阳年辐射总量及其空间分布特征的回归模型, 进行模拟分析。因此如何利用长时间系列的日照时数、总云量、水汽压、太阳总辐射量等月观测数据, 获取合理、有代表性的要素相应年数据资料是构建理想模型, 并进行精确模拟的关键。为此对测站的原始观测数据进行分析处理: 一是对测站的太阳总辐射量、日照时数分别计算其历年观测记录的逐月平均, 12 个月等权累加, 获取相应测站的年太阳年辐射总量及年日照时数; 二是对测站的云量、水汽压分别计算其历年观测记录的逐月平均, 12 个月等权平均, 获取相应测站的年均总云量及年均水汽压。

2.2 研究方法

对全国范围上的太阳总辐射进行了研究, 主要有 3 种方式: 太阳辐射观测值的直接空间插值方法、卫星遥感方法(魏合理等, 2003)和气候学计算方法。直接的空间插值方法, 是一种最简单的方式, 但是, 目前国内与气象观测台站相比, 辐射观测站点有限, 精度难以满足实际应用。卫星遥感方法, 由于遥感反演技术本身的复杂性, 存在较大的不确定性(何洪林等, 2004), 这方面的研究相对较少。气候学计算方法, 虽然已经比较成熟, 但太阳辐射基数值的确定是相当不易, 而且模型的 a、b 经验系数不仅具有空间变化, 在许多地区具有明显的季节变化(鞠晓慧等, 2005), 这些决定了其在全国尺度上计算年辐射总量相当复杂。本文将通过逐步回归分析法, 建立太阳辐射观测值与其影响要素之间的“最优”回归模型, 确保模拟精度的条件下, 降低模型复杂性、提高运算效率。

多元逐步回归方法: 太阳年辐射总量多元回归模拟研究分析, 影响太阳年辐射总量 Q 的因子很多, 一般而言, 影响因子(自变量)越多, 回归平方和越大, 残差平方和越小。然而, 这些因子之间可能存在多重共线性, 这就给系数的估计带来不合理的解析。采用较多的变量来拟合回归方程, 得到的方程的稳定性差, 每个自变量的区间误差累积会影响总体误差, 用这样建立起来的回归方程作为预测可靠性下降、精度降低。另一方面, 如果采用了对 Q 影响甚小的变量而遗漏了重要变量, 可导致估计量产生偏离和不一致。为了得到一个可靠的回归模型, 需要一种方法能有效地从众多影响 Q 的因子中挑选出对 Q 贡献大的变量, 在它们和 Q 的观测数据基础上建立“最优”的回归方程。

目前的研究, 主要通过 3 类方法克服模型变量的多重共线性问题: 差分法、减小参数估计量的方差和排除引起共线性的变量。前两种方法只能减轻多重共线性对模型的影响, 而第 3 类方法, 从根本上寻找引起多重共线性的解释变量, 将其排除出原模型, 因而更为有效。通过排除引起共线性的变量, 建立“最优”的回归方程有以下几种方法: (1)根据一定准则(比如 Cp 准则、AIC 准则), 从所有可能的因子(变量)组合的回归方程中选择‘最优’回归模型。(2)向后剔除法, 从包含全部变量的回归方程中逐次剔除不显著因子; 其局限性是, 自变量高度相关时, 可能得不出正确的结果。(3)向前引入法: 自变量从无到有、由少到多一个一个引入回归方程; 其局限性, 即后续变量的引入可能会使先进入方程的自变量, 变得不重要。(4) 逐步回归分析法: 先规定两个阈值 $F_{引入}$和 $F_{剔除}$, 当候选变量中最大 $F \geqslant F_{引入}$时, 引入相应变量; 已进入方程的变量最小 $F \leqslant F_{剔除}$时, 剔除相应变量。如此交替进行直到无变量引入和无剔

除为止；其局限性，计算相对复杂。第1种方法自变量个数增加计算规模成几何级数增长。第2—4种的共同特点是每一步只引入或剔除一个自变量；以第4种方法，即逐步回归分析法在筛选变量方面较为理想。这种方法克服了变量多重共线性和解释的优良有效性，在地学(田永中等, 2005；付勇等, 2009)、生态(Petersen & Stringham, 2008)、气象(Ertekin & Evrendilek, 2007)、材料(王志强等, 2007)和医学(Ng, 2003)等领域得到广泛应用。

多元逐步回归的基本思想是从一个自变量开始，视自变量对因变量作用的显著程度，从大到小地依次逐个引入回归方程。在引入新变量后用偏回归平方和检验其显著性，如果显著才能将该变量加到方程中，方程中加入了新变量后，对原有的变量重新用偏回归平方和进行检验，若某个变量变得不显著时，将它从方程中删除。重复上述步骤，直至既无不显著的变量从回归方程中剔除，又无显著变量可引入回归方程时为止。逐步回归的每引入或剔除一个变量重复的步骤如图1，其中F值的计算公式如下：

$$F=\frac{U\ /k}{Q_{\mathrm{e}}/(n-k-1)}\sim F(k,n-k-1) \tag{1}$$

3 太阳总辐射的模拟及分析

3.1 影响因子的相关分析

相关分析的方法较多，常用的是绘制散点图，这种方法直观不很精确。通过计算相关系数进行相关分析能较为精确地描述变量之间的线性相关程度。

研究综合分析了太阳年辐射总量的影响因子资料的完整性和可获得性，从影响太阳年辐射总量($0.01\mathrm{MJ\cdot m^{-2}\cdot a^{-1}}$)的众多因子中，选取了6个影响因子：$x_1$年日照时数(0.1h)、$x_2$年均日照百分率(%)、$x_3$年均总云量(0.01)、$x_4$年均水汽压(hPa)、$x_5$海拔高度(0.1m)和$x_6$纬度(°)。以全国122个太阳辐射站点的数据为基础，选取观测时间长，且6个影响因子观测值齐全的共96个站点的观测数据，通过程序计算Pearson相关系数，衡量各个影响因子之间以及各因子与太阳年辐射总量Q之间的线性相关程度。程序运行结果如表1。x_1年日照时数和x_2年均日照百分率与其他各要素之间的相关系数非常接近，它们本身之间的相关系数高达0.9963，所以在选择因素时，两者可只取其一。在自由度为95，置信概率为0.99时，各因子与Q之间的相关性是显著的。因此，本

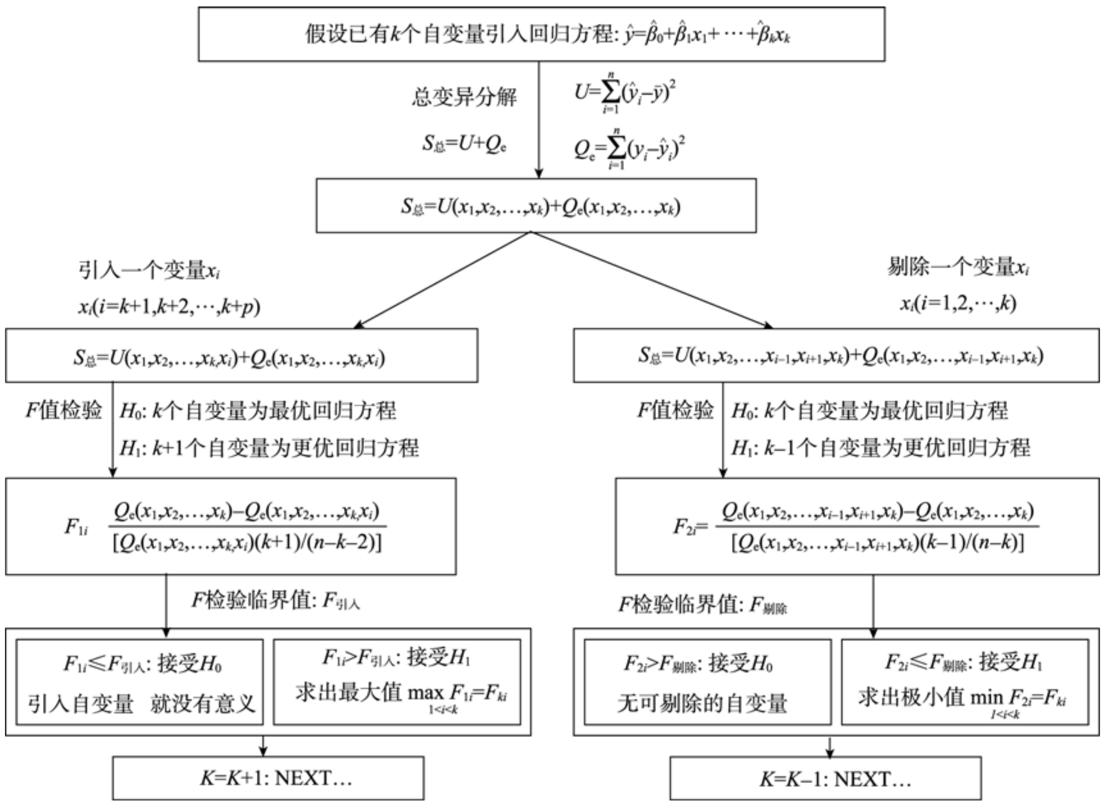

图1 多元逐步回归引入或剔除变量的程序流程图

表 1 太阳辐射影响因子之间的 Pearson 相关系数矩阵(N=95)

	x_1	x_2	x_3	x_4	x_5	x_6	Q
x_1	1.0000						
x_2	0.9963	1.0000					
x_3	−0.9158	−0.9189	1.0000				
x_4	−0.6677	−0.6525	0.7214	1.0000			
x_5	0.1824	0.1731	−0.0470	−0.4936	1.0000		
x_6	0.6306	0.6157	−0.7748	−0.7899	−0.0068	1.0000	
Q	0.8456	0.8435	−0.6446	−0.5119	0.4276	0.2347	1.0000

研究将排除 x_2 年均日照百分率因子，利用其他的 5 个因子进行多元逐步回归实验。

3.2 太阳总辐射模型的建立

太阳年辐射总量模型建立中，选择了 x_1 年日照时数、x_3 年均总云量、x_4 年均水汽压、x_5 海拔高度和 x_6 纬度 5 个因子的观测值保留完整的太阳辐射站点 116 个，依据表 1 逐步引入对太阳年辐射总量影响显著变量，剔除因引入其他因子而变不显著的变量。

3.2.1 模型变量确定

太阳年辐射总量变化受多个因素影响，在模型变量的引入与剔除中，为突出模型解释能力，保证逐步回归模型中因子的可靠性，显著性水平取 0.01；检验分析中选用了剩余标准差 RMSE，单位 $\mathrm{MJ \cdot m^{-2} \cdot a^{-1}}$、决定系数 R^2、校正决定系数 Adj-R^2 等评价指标。剩余标准差反映了应变量在扣除自变量的线性影响后的离散程度；决定系数表示回归平方和占总离差平方和的比例，反映回归直线的拟合程度；决定系数是复相关系数 R 的平方，而复相关系数 R 是随方程中的变量个数增加而增加的，即使因子对 Q 无显著性意义，为了克服这一缺点，对它进行校正，即使用校正决定系数 Adj-R^2，消除 R^2 中对模型没有解释力的新增变量。

研究中参考已有文献(田永中等, 2005)的研究成果，假设已有 3 个自变量 x_1,x_5,x_6 引入回归方程，利用逐步回归思想，依据自变量对因变量作用的显著程度，依次引入 x_4,x_3，剔除了 x_5 变量。在这些过程中，对每一个变量的引入与剔除，进行回归分析统计检验，对统计量及误差进行分析。评价指标及相应的统计量 F 值的比较如表 2。从表 2 中分析引入 x_4,x_3 的过程中剩余标准差减小，决定系数、校正决定系数增大。剔除 x_5 变量，虽然剩余标准差、相关系数及决定系数没有太大的变化，但是统计量 F 的值明显增大，因此对 Q 太阳年辐射总量影响显著变量确定为 x_1, x_3,x_4,x_6 等 4 个因子。

表 2 太阳年辐射总量多元逐步回归模型变量引入与剔除过程的统计检验分析

	x_1, x_5, x_6	x_1, x_4, x_5, x_6	x_1, x_3, x_4, x_5, x_6	x_1, x_3, x_4, x_6
RMSE	274.001	264.591	257.307	257.770
R^2	0.910927	0.917682	0.922853	0.921871
Adj-R^2	0.908541	0.914715	0.919346	0.919056
F	381.799	309.356	263.169	327.432

3.2.2 模型参数估计

建立回归模型时，既要尽可能地提高拟合的精度，又要尽可能地使模型简单。在确定模型变量后，进行模型构建及参数估计。模型的误差比较分析选用了中误差(mean square error, MSE)及相对中误差(relative mean square error, R-MSE)及模型解算效率(computational efficiency, CE)等评价指标。中误差即 n 次观测偶然误差平方和的均方根，当样本量 n 趋向于无穷大时，中误差即为标准差；相对中误差是中误差的绝对值与相应观测结果之比。

研究中对 Q 与变量 x_1, x_3,x_4,x_6 进行多元线性回归及多元二项式回归分析，建立起相应的模型并进行误差分析比较。太阳年辐射总量的多元线性回归及多元二项式回归式子如 (3)及(4)，(2)为已有文献(田永中等, 2005)的式子。图 2(a)、(b)、(c)分别采用式(2)、式(3) 和式(4)进行计算得的预测值与观测值对比散点图，对应建模方法的误差及解算效率比较分析如表 3。

表 3 太阳辐射不同模型的误差分析与效率比较

	模型 2	模型 3	模型 4
MSE	297.68	252.15	234.54
R-MSE	0.0580	0.0491	0.0457
CE	0.800403	1	0.294891

注: MSE($\mathrm{MJ \cdot m^{-2} \cdot a^{-1}}$)，中误差; R-MSE，相对中误差; CE，模型解算效率，取不同规模的 18 组数据解算效率的平均值。

从图 2 看出，图 2(c)结果最理想， 图 2(b)比 2(a)更趋于对角线。从表 3 分析，模型 3、模型 4 相对模型 2，相对中误差 R-MSE 降低 15.3%和 21.2%。而模型 2、模型 4 因为存在交叉项，解算效率 CE 较模型 3 降低约 20%和 70.5%。基于以上的分析，式(3)无论从运算效率上或运算结果都是理想的，下面采用模型 3 进行模拟分析。

$$Q_{\mathrm{Ref}}=170292 + 20.73189x_1 - 0.19171x_1 x_6+ 0.07212x_5 x_6 \tag{2}$$

$$Q_{\mathrm{Line}}=299608.16779+18.20611x_1+1976.62392x_3-609.31017x_4-7310.72926x_6 \tag{3}$$

$$Q_{\mathrm{Int}}=355453.51165+17.67786x_1-2057.18878x_3-355.29889 x_4+767.34475 x_6+0.11365 x_1 x_3-0.01519 x_1 x_4-0.235735 x_1 x_6+8.54718 x_3 x_4-10.11656 x_3 x_6-19.90513 x_4 x_6 \tag{4}$$

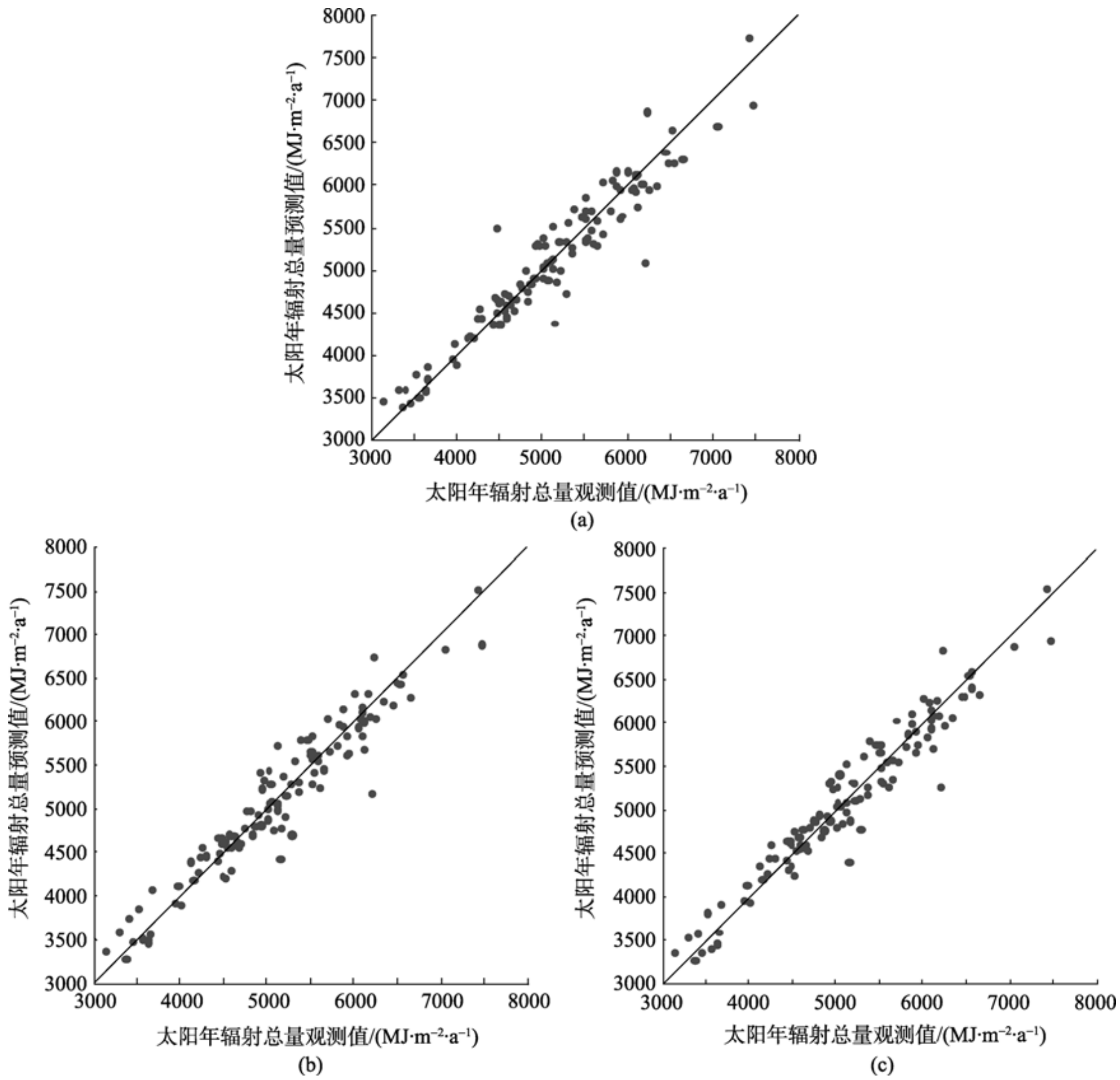

图 2 太阳年辐射总量的观测值与模拟值对比散点图

(a) 采用式(2); (b) 采用式(3); (c) 采用式(4)

3.2.3 模拟结果对比

研究中，为检验模型模拟的有效性，分别采用观测值直接空间插值与逐步回归模型后插值，对全国范围内的太阳年辐射总量进行模拟对比分析。模型模拟分辨率为 1km×1km，规模为 4173×4847 网格。

图 3(a)为 122 太阳辐射观测台站位置分布及其观测值直接插值获取的模拟结果，这些测站尽管其空间分布较为均匀，但在中国这样大的幅员范围内还是太少；模拟结果仅仅反映中国太阳年辐射总量的总体概貌，许多区域如祁连山脉、塔里木盆地、青藏高原南部等误差较大。相比图 3(b)，由 731 地面站的常规气象观测资料利用逐步回归模型 3 获取太阳年辐射总量估计值，再进行空间插值获取模拟结果，其模拟值能较好拟合实际的太阳年辐射总量。模拟结果展示许多细节，层次更加丰富，包含的信息量更大，突出了太阳辐射总量的空间变化及分布特征；模拟结果的分类专题图如图 4。在常规气象观测点相对稀疏的西部地区，模拟效果有所改善，仍然不够理想。

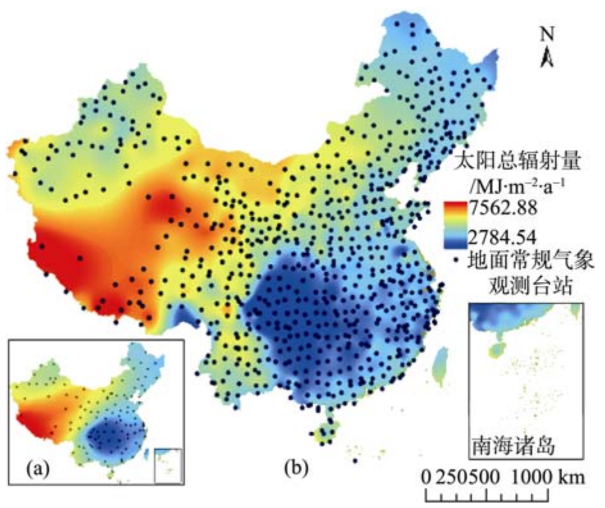

图 3 中国太阳年辐射总量的模拟结果

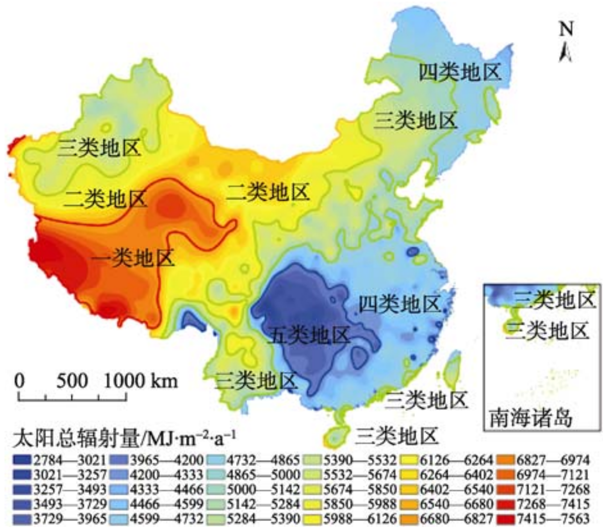

图 4 中国太阳年辐射总量分类图

4 结论与讨论

(1) 中国陆地表面每年接受的总太阳辐射能约为 52.4×10^{18} KJ(相当于 1790 亿 t 标准煤); 单位面积上太阳年辐射总量约介于 2780—7560MJ·m^{-2}·a^{-1}; 分布具有明显的区域性, 呈现西北、北部高, 东南、东部及东北的部分地区低的特点(图 4)。

(2) 太阳辐射资源最丰富的地区为西藏自治区、青海东北部及甘肃西部边境, 达 6680MJ·m^{-2}·a^{-1} 以上, 为一类地区, 面积约 130 万 km^2。太阳辐射年总量最低的为四川盆地、青藏高原南部的雅鲁藏布峡谷地区, 低于 4200MJ·m^{-2}·a^{-1}, 太阳能资源最贫乏地区, 即五类地区, 面积约 75 万 km^2。

(3) 内蒙古南部、山西及宁夏北部、甘肃中部及西北部、青海东部、西藏东南部和新疆东南部等地, 为二类地区(5850—6680MJ·m^{-2}·a^{-1}), 面积约 240 万 km^2。三类地区(5000—5850MJ·m^{-2}·a^{-1}), 主要分布内蒙古东北部、吉林与辽宁西部、河北东南部及北部、山东东南部、山西南部、河南及陕西北部、甘肃东南部、新疆北部、云南、广东南部、福建南部和台湾西南部等地, 面积约 280 万 km^2。四类地区(4200—5000MJ·m^{-2}·a^{-1}), 主要包括黑龙江、吉林与辽宁东北部、陕西南部、湖南、湖北、江苏、安徽、江西、浙江、广西、广东北部、福建东北部、台湾北部等地, 面积约 235 万 km^2。

(4) 本研究方法, 为融合遥感信息进行农业气象要素空间插值拟合提供新思路。使用地理回归拟合方法, 通过引入卫星影像等具有延续分布的相关信息要素, 利用该要素的空间变异性来推测插值要素的空间变异性, 从而克服局部研究区因采样数据稀少影响模拟结果。

总之, 本研究建立起全国太阳年辐射总量多元逐步回归模型, 为模拟广大无辐射观测地区的太阳年辐射总量提供确实可行的方法。所构建的模型考虑了气象要素、地理因子等的影响, 实现太阳年辐射总量空间分布的连续模拟结果; 不仅适应于太阳辐射能量的区域分布研究, 在模拟数值上也能较好反映实际的太阳年辐射总量。因此, 可强化模拟结果在农业生态区划, 陆地生态系统中食物供给与承载功能变化监测, 植被净第一性生产力(NPP), 节能减排、新型能源开发等研究中的应用。显然, 太阳总辐射除了受到天文、地理因子和大气状况等的影响外, 还受到局部地形的影响; 今后, 在一些地形复杂(起伏大)特殊区域, 当常规气象数据的采样率不足情况下, 应当考虑地形复杂度等影响, 消除可能造成的较大误差, 使太阳总辐射值的模拟效果得到进一步完善。

REFERENCES

Ertekin C and Evrendilek F. 2007. Spatio-temporal modeling of global solar radiation dynamics as a function of sunshine duration for Turkey. *Agricultural and Forest Meteorology*, **145**(1/2): 36—47

Fu Y, Wang L J, Chai F M, Deng G, Chen Y and Chen J H. 2009. The application of multi-variate linear regression combined with stepwise regression analysis for Cu-Ni sulfide deposit of Baishiquan. *Earth Science Frontiers*, **16**(1): 373—380

He H L, Yu G R, Liu X A, Su W, Niu D and Yue Y Z. 2004. Study on spatialization technology of terrestrial eco-information in China(Ⅱ): Solar radiation. *Journal of Natural Resources*, **19**(5): 679—687

Ju X H, Tu Q P, and Li Q X. 2005. Discussion on the climatological calculation of solar radiation. *Journal of Nanjing Institute of Meteorology*, **28**(4): 516—521

Ng C M. 2003. Comparison of neural network, Bayesian and multiple stepwise regression-based limited sampling models to estimate area under the curve. *Pharmacotherapy*, **23**(8): 1044—1051

Petersen S L. and Stringham T K. 2008. Development of GIS-based models to predict plant community structure in relation to western juniper establishment. *Forest Ecology and Management*, **256**(5): 981—989

Robaa S M. 2009. Validation of the existing models for estimating global solar radiation over Egypt. *Energy Conversion and Management*, **50**(1): 184—193

Tian Y Z, Zhu L F, and Yue T X. 2005. Comparative study on spatial simulation of solar radiation on level surfaces. *Geo-information Science*, **7**(4): 95—100

Trnka Miroslav, Eitzinger Josef, Kapler Pavel, Dubrovský Martin, Semerádová Daniela, Žalud Zdeněk and Formayer Herbert. 2007. Effect of estimated daily global solar radiation data on the results of crop growth models. *Sensors*, **7**(10): 2330—2362

Wang Z Q, Chai S X, Zhong X M, Guo Y, Wang X Y and Wei L. 2007. Multivariate stepwise regression method use in correlation analysis of microstructure indices and strength of solidified soil. *Rock and Soil Mechanics*, **28**(8): 1650—1654

Wei H L, Xu Q S and Zhang T S. 2003. Observation of solar irradiance at the surface from GMS-5. *Journal of Remote Sensing*, **7**(6): 466—471

Xu S Y, Peddle D R, Coburn C A. and Kienzle S. 2008. Sensitivity of a carbon and productivity model to climatic, water, terrain, and biophysical parameters in a Rocky Mountain watershed. *Canadian Journal of Remote Sensing*, **34**(3): 245—258

Yue T X, Du Z P and Song D J. 2007. High accuracy surface modelling: HASM4. *Journal of Image and Graphics*, **12**(2): 343—348

Yue T X, Tian Y Z, Liu J Y and Fan Z M. 2008. Surface modeling of human carrying capacity of terrestrial ecosystems in China. *Ecological Modelling*, **214**(2/4): 168—180

附中文参考文献

付勇, 汪立今, 柴凤梅, 邓刚, 陈勇, 陈俊华. 2009. 多元线性回归和逐步回归分析在白石泉 Cu-Ni 硫化物矿床研究中的应用. 地学前缘, **16**(1): 373—380

何洪林, 于贵瑞, 刘新安, 苏文, 牛栋, 岳燕珍. 2004. 中国陆地生态信息空间化技术研究(Ⅱ)——太阳辐射要素. 自然资源学报, **19**(5): 679—687

鞠晓慧, 屠其璞, 李庆祥. 2005. 我国太阳总辐射气候学计算方法的再讨论. 南京气象学院学报, **28**(4):516—521

田永中, 朱莉芬, 岳天祥. 2005. 水平面上太阳辐射空间模拟对比研究. 地球信息科学, **7**(4): 95—100

王志强, 柴寿喜, 仲晓梅, 郭英, 王晓燕, 魏丽. 2007. 多元逐步回归分析应用于固化土强度与微结构参数相关性评价. 岩土力学, **28**(8): 1650—1654

魏合理, 徐青山, 张天舒. 2003. 用 GMS-5 气象卫星遥测地面太阳总辐射. 遥感学报, **7**(6): 466—471

岳天祥, 杜正平, 宋敦江. 2007. 高精度曲面建模: HASM4. 中国图象图形学报, **12**(2): 343—348