Spatial scaling of net primary productivity model based on remote sensing

WANG Liwen[1], WEI Yaxing[1], NIU Zheng[2]

1. *College of Urban and Environment Science*, *Liaoning Normal University*, *Liaoning Dalian* 116029, *China*;
2. *The State Key Laboratory of Remote Sensing Science*, *Institute of Remote Sensing Applications*, *Chinese Academy of Sciences*, *Beijing* 100101, *China*

Abstract: Spatial scaling for net primary productivity (NPP) refers to the transferring process of establishing quantitative correlation between simulated NPP derived from data at different spatial resolutions. How to transfer NPP at one scale by the algorithm with smaller error to at another is the urgent problem. Nonlinearity and effects from land cover type are two main problems in NPP scaling. In this paper, the contextural approach based on mixed pixels and support vector machine (SVM) algorithm are used to make the scaling model from the fine resolution (TM) to the coarse resolution (MODIS). Spatial scaling from NPP retrieved from fine resolution data to NPP derived from coarse resolution images is performed, and the correction of scale effect to NPP retrieved from coarse resolution data of MODIS is accomplished. The result shows that the correlation between $R_{j\_corrected}$ of the correction factor for scale effect and 1-$F_{\text{middle density grassland}}$ estimated by SVM regression model is higher (R^2=0.81). Before the correction for scale effect, the correlation between NPP_{MODIS} and NPP_{TM} is lower (R^2=0.69; RMSE=3.47), while the correlation between NPP_{TM} and corrected $NPP_{MODIS\_corrected}$ is higher (R^2=0.84; RMSE=1.87). Therefore, NPP corrected for scale effect has been greatly improved in both correlation and error.

Key words: net primary productivity, light use efficiency model, remote sensing, scaling, support vector machine

CLC number: TP761 **Document code:** A

Citation format: Wang L W, Wei Y X and Niu Z. 2010. Spatial scaling of net primary productivity model based on remote sensing. *Journal of Remote Sensing*. **14**(6): 1074—1089

1 INTRODUCTION

The advantage of remote sensing is to provide land surface information frequently and longer. An important reason is that problems in remote sensing are complex and the inconsistent scale data are derived from the measured data, data from different methods of remote sensing and data for applications. How to scale between different scale data is the core of scale problem (Li *et al*., 2001; Turner *et al*., 1996). Spatial heterogeneity of the earth system limits the transferring from one scale to another. That is the major cause that data at different scale retrieved from satellite sensors need to be unified (Su *et al*., 2001; Pierce *et al*., 1995).

Spatial scaling for net primary productivity (NPP) refers to the transferring process of establishing quantitative correlation between simulated NPP derived from data at different spatial resolutions. In the process of scaling, simulated NPP at one resolution can be deduced from simulated NPP at another resolution (Ehleringer & Field 1993). Errors in land surface parameters at the coarse resolution are derived from the synthesis of various kinds of spectral signals from land surface objects received by sensors. In addition, the quantitative correlation between mixed spectral signals and land surface parameters is difficult to be precisely defined. Therefore, the urgent problem is how to transfer NPP at one scale by the algorithm with smaller error to at another. In scaling, nonlinearity and land cover types that have effects on scaling are two main problems (Chen *et al*., 1999; Nemani *et al*., 2003).

In this paper, a contextural approach based on mixed pixels was used to make the scaling model from the fine resolution (TM) to the coarse resolution (MODIS). The algorithm of support vector machine (SVM) was adopted to establish the regression model between correction factor (R_j) of scale effect for a pixel labeled as cover type j and F_{ij} of the nondominant cover type i in this pixel. That is, F_{ij} was the fraction of a nondominant cover type i in a pixel labeled as cover type j in fine resolution data of TM. Afterwards Spatial scaling from NPP retrieved at the fine resolution to NPP derived from the coarse resolution was performed, and the correction of scale effect to NPP retrieved at the coarse resolution was accomplished.

Received: 2009-12-14; **Accepted**: 2010-06-07

Foundation: Chinese Liaoning Province Education Bureau General Science Research Project (No. L2010226); Chinese Education Ministry Humanities and Social Sciences Key Research Base Project (No.08JJD790142); Chinese Liaoning Province Education Bureau Innovation Team Project (No. 2007T095); Chinese Special Funds for Major State Basic Research Project (No. 2007CB714406).

First author biography: WANG Liwen (1971—), female, lecturer. She graduated from Institute of Remote Sensing Applications, Chinese Academy of Sciences in 2009, majors in Cartography and Geographic Information System. She is focusing on the research of remote sensing applications. She has published more than 10 science papers. E-mail: wlw9585@163.com

2 STUDY AREA

This study is conducted in Dari County, located in the southeastern of Qinghai Province in West China (Fig.1). Dari County is geographically situated between 98°15′29″ and 100°32′41″ East longitude and between 32°36′42″ and 34°15′20″ North latitude. The length of Dari County from east to west is about 162 km, and the width from north to south is almost 126 km. The total area of Dari County is 14600 km^2, and the average altitude is 4426 m. The climate of Dari County is semi-humid plateau type without obvious season variation in addition to warm and cold seasons. The average temperature of Dari County ranges from −0.1℃ to −3.5℃. The annual precipitation is 569 mm usually taking place from June to September. The annual evaporation is 1119.07 mm, and the monthly highest is 150.1 mm occurring in May. The main vegetation in Dari County is alpine meadow dominated by Kobresia Pygmaea, K. Humilis and K. Tibetica, which are widely distributed on sunny sides of hills and wide valley. Alpine shrub is mainly dominated by Dasiphora Bsuticosa, Salix Cupularis, which are often located on shade sides of hills.

3 DATA PREPARATION

Satellite data used in this paper include MODIS data products and TM data. In the process of making the NPP model, MODIS data products were used to calculation and validation including MOD09A1 (surface reflectance 8-day L3 global 500 m), MOD11A2 (land surface temperature/emissivity 8-day L3 global 1 km), MOD12Q1 (land cover type yearly L3 global 1 km), MOD13A2 (vegetation indices 16-day L3 global 1 km), MOD15A2 (leaf area index/FPAR 8-day L4 global 1 km), MOD17A2 (gross primary productivity 8-day L4 global 1 km) and MOD43B3 (albedo 16-day L3 global 1 km), which were supplied by LPDAAC (Land Process Distributed Active Archive Center, U.S.A.). Geometric rectification, cloud elimination processsing and modifying abnormal pixel were made to MODIS data in advance. The study used 5 scenes of TM images derived from Landsat-5, of which path/row numbers were 132/37, 133/36, 133/37, 134/36 and 134/37. Image extent covered the whole Dari County. Geometric rectification and atmospheric correction was performed to TM data. The received date of MODIS and TM was in July 2006, which represented the period that vegetation grew better in a year.

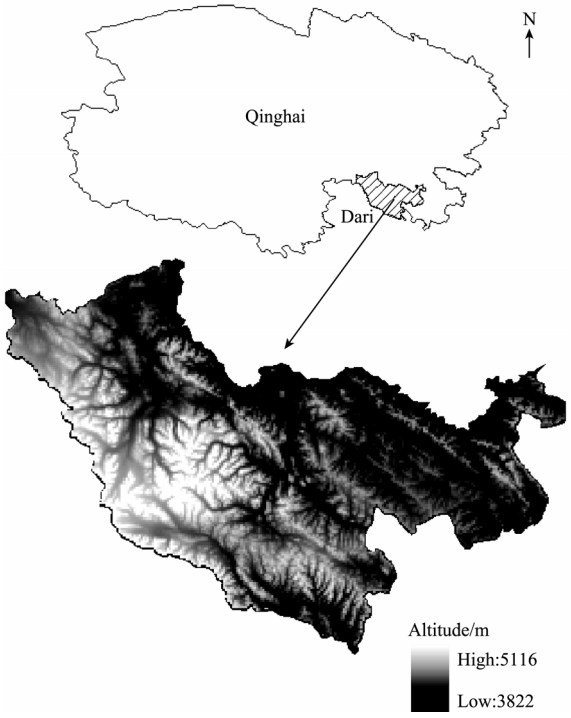

Fig. 1 Location map of the study area

The meteorology data including average temperature, precipitation, average atmospheric relative humidity, sunshine duration, air pressure and mean wind speed in the paper was derived from Chinese national meteorological center. The method of vegetation cover classification in Qinghai Province (Wang *et al*., 2008) was used in the research. DEM data were downloaded from Centro International de Agriculture Tropical (CIAT, http://srtm.csi.cgiar.org/).

Measured data in the paper was from experiment in Dari County in July 2006. Experimental instrument included LI-6400 portable photosynthesis system, LAI-2000 vegetation canopy analysis system, SUNSCAN canopy analysis system and so on.

4 METHODS

4.1 NPP simulation model

Remote sensing data can be directly used in the NPP simulation model based on light use efficiency (LUE). Thus spatio-temporal resolution of the model can be greatly improved. It makes it possible to simulate and dynamic monitor NPP for large scale. The NPP simulation model in the paper is based on LUE and learns from advantages of MODIS-PSN, CASA, GLO-PEM, VPM models. At the same time, the NPP model fully considers characteristics of vegetation LUE and environment in the study area. The NPP simulation model based on remote sensing can be expressed as follows.

$$\mathrm{NPP} = \mathrm{PAR} \times \mathrm{FPAR} \times \min\left(T_\mathrm{S},\ W_\mathrm{S}\right) \times \varepsilon_{\max} \qquad (1)$$

where PAR is photosynthetic active radiation (MJ·m^{-2}), FPAR is fractional photosynthetic active radiation (unitless), T_S is temperature stress factor (unitless), W_S is water stress factor (unitless), and $\varepsilon_{\max}$ is maximum LUE (gC·MJ^{-1}).

4.1.1 Maximum LUE ($\varepsilon_{\max}$)

Values of maximum LUE for various plants are different. They have great effect on final NPP simulation results. Optimal values of maximum LUE for different vegetation are simulated by maximum likelihood estimate based on literature and field experimental data in the study area (Wang *et al*., 2006; Zhu *et*

al., 2006; Li *et al*., 2004; Yang *et al*., 1987; Pu *et al*., 2005).

$$\hat{\varepsilon}_{\max}=\frac{\sum_{i=1}^{n}x_i y_i-n\overline{xy}}{\sum_{i=1}^{n}x_i^{\,2}-n\overline{x}^{2}} \tag{2}$$

where $\hat{\varepsilon}_{\max}$ is the simulated value of maximum LUE, x_i is the calculated value of $\mathrm{PAR}\times\mathrm{FPAR}\times\min(T_s, W_s)$ for the plant, y_i is the NPP value for the plant, i is the sample number for the plant, n is maximum sample number for the plant, $\overline{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$, and $\overline{y}=\frac{1}{n}\sum_{i=1}^{n}y_i$.

After Eq.(2) is used to simulate maximum LUE value of main type of grassland or shrub in the study area, Eq. (3) is adopted to calculate maximum LUE value of every kind of grassland or shrub remote sensing-based classification in order to improve simulated precision of maximum LUE.

$$\varepsilon_{\max}=\sum_{i=1}^{n}c_i(\varepsilon_{\max})_i \tag{3}$$

where c_i is area fraction of grassland or shrub labeled as cover type i, and $(\varepsilon_{\max})_i$ is maximum LUE of grassland or shrub labeled as cover type i.

Calculated results of maximum LUE for main vegetation in the study area are 0.908 gC·MJ^{-1} for broad-leaved forest, 0.725 gC·MJ^{-1} for needle-leaved and broad-leaved mixed forest, 0.645 gC·MJ^{-1} for needle-leaved forest, 0.312 gC·MJ^{-1} for dense grassland, 0.538 gC·MJ^{-1} for dense shrub, 0.234 gC·MJ^{-1} for middle density grassland, 0.208 gC·MJ^{-1} for grassland mixed with farmland, 0.175 gC·MJ^{-1} for sparse grassland, and 0.114 gC·MJ^{-1} for sparse shrub.

4.1.2 FPAR

The correlation equation between NDVI and FPAR is made based on data in the study area. FPAR in non-vegetation cover regions is assumed as a minimum value of 0, and FPAR in complete vegetation cover areas is assigned to a maximum value of 0.95 (Chen *et al*., 2007):

$$\mathrm{FPAR}=\min\left[\frac{\mathrm{SR}}{\mathrm{SR}_{\max}-\mathrm{SR}_{\min}}-\frac{\mathrm{SR}_{\min}}{\mathrm{SR}_{\max}-\mathrm{SR}_{\min}},0.95\right] \tag{4}$$

$$\mathrm{SR}=\frac{1+\mathrm{NDVI}}{1-\mathrm{NDVI}} \tag{5}$$

where $\mathrm{SR}_{\min}$ of all kinds of vegetation is assigned to a same value of 1.06, values of $\mathrm{SR}_{\max}$ are 6.14 for broad-leaved forest, 3.88 for needle-leaved and broad-leaved mixed forest, 3.44 for needle-leaved forest, 3 for dense grassland, 3.27 for dense shrub, 2.17 for middle density grassland, 1.94 for grassland mixed with farmland, 1.78 for sparse grassland, and 1.70 for sparse shrub.

4.1.3 Water stress factor (W_s)

In this paper, evaporative fraction (EF) which has been proven in many researches is used to calculate water stress factor (W_s) (Suleiman & Crago, 2004; Venturini *et al*., 2004; Wen *et al*., 2006; Zhang *et al*., 2004):

$$W_s=\mathrm{EF}=\frac{\mathrm{LE}}{\mathrm{LE}+H} \tag{6}$$

where LE is latent heat flux (W·m^{-2}), and H is sensible heat flux (W·m^{-2}).

Fig. 2 shows the flowchart of calculating EF.

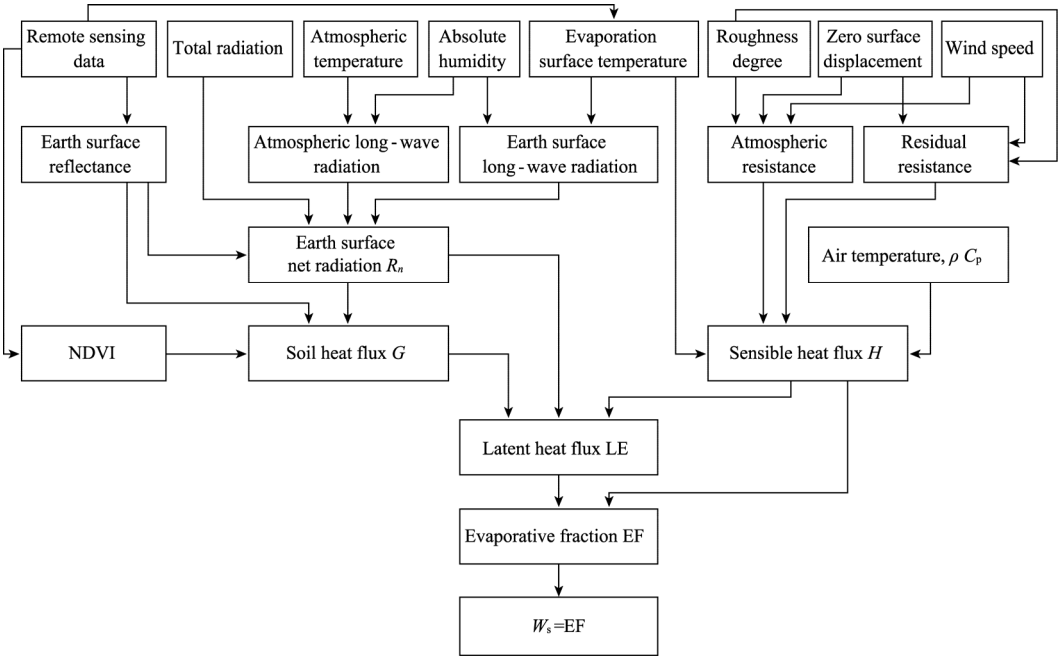

Fig. 2 Flowchart of EF

4.1.4 Temperature stress factor (T_s)

Temperature stress factor is estimated by the method from the Terrestrial Ecosystem Model (TEM) of Raich (1991).

$$T_s = \frac{(T - T_{min})(T - T_{max})}{(T - T_{min})(T - T_{max}) - (T - T_{opt})^2} \quad (7)$$

where T is air temperature, T_{min}, T_{max} and T_{opt} are minimum, maximum and optimum air temperature for photosynthetic activity, respectively. If air temperature is below T_{min} or beyond T_{max}, T_s is set to 0. In this paper, T_{min} and T_{max} are set to 0℃ and 36℃, respectively. Because vegetation have adapted well to the temperature of growth environment, T_{opt} may be set to the mean temperature of growing season.

4.2 Spatial scaling of NPP

4.2.1 Contextural approach based on mixed pixels

According to primary problems affecting scaling, contextural approach based on mixed pixels of Anita *et al.* (2004) is used to perform spatial scaling of NPP. The theory is as follows.

Mixed pixels are common phenomena in remote sensing images. The presence of mixed pixels is the main reason that precisions of land surface parameters can not meet requirements. In order to improve precisions of remote sensing applications, unmixing of pixels must be solved and remote sensing applications need to transfer from pixels level to sub-pixels level. A pixel in coarse resolution images can be deemed as a mixed pixel composed of many pixels at corresponding coordinates in fine resolution images, and these pixels in fine resolution images can be considered as sub-pixels of this pixel in coarse resolution images (Chen *et al*., 1999; Tong *et al*., 2006).

Spatial scaling based on contextural approach adopts the theory of analyzing area fractions for land cover types to compute land surface parameters at various resolutions. A pixel in remote sensing classifications is often defined as the pixel of dominant cover type whose area is the largest in the pixel, while other nondominant cover types in this pixel is ignored during calculating land surface parameters. This leads to every pixel only to be labeled as a kind of dominant cover type. Therefore, final calculation results contain much uncertainty (Chen *et al*., 1999). The ideal method is that land surface parameters in coarse resolution images should be calculated as mean values of parameters at corresponding coordinates in fine resolution images.

In this paper, NPP_{TM} retrieved from fine resolution data of TM is assumed to represent the real value of NPP for terrestrial vegetation (Fig.3(a)), and NPP_{MODIS} derived from coarse resolution data of MODIS is the approximate value which needs to be corrected for scale effect (Fig.3(b)). A pixel in MODIS (1×1 km^2) is supposed to contain 1089 pixels of TM (30×30 m^2). Due to the presence of mixed pixels, 1089 pixels in TM contain many kinds of land cover types. While a pixel at the corresponding coordinate in MODIS is labeled as a kind of dominant cover type which covers the largest area in this pixel, and other nondominant cover types in the same pixel are neglected. TM and MODIS data which all covers the study area are respectively input as parameters of the NPP model based on LUE, and NPP_{TM} and NPP_{MODIS} can be achieved respectively (Fig.3). It should be noticed that estimated NPP results for two kinds of scales need to be converted to the same resolution in order to contrastively analyze NPP results. That is, the resolution of $NPP_{TM'}$ retrieved from fine resolution data of TM needs to be switched to 1 km. when NPP_{TM} of 30 m resolution is converted to 1 km $NPP_{TM'}$, arithmetical mean value of NPP for many pixels at the corresponding coordinate in $NPP_{TM'}$ images is considered as estimated NPP for a pixel in NPP_{TM} images. Due to the resolution difference between TM and MODIS which are all input parameters of the model, much error appears between NPP_{TM} and NPP_{MODIS}. The reason that calculating results for two scales are different is that large amount of mixed pixels are in coarse resolution data of MODIS comparing with fine resolution data of TM and nondominant cover types in mixed pixels are ignored. Then NPP_{MODIS} is corrected for scale effect based on NPP_{TM} to retrieve $NPP_{MODIS\,j\_correcred}$ so that the precision of estimated NPP derived from coarse resolution data of remote sensing can be improved.

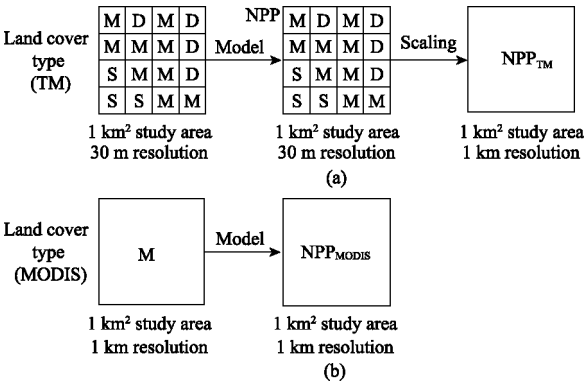

Fig. 3 Flowchart of calculating NPP by fine or coarse resolution data of remote sensing

(a) Flowchart of NPP scaling retrieved from fine resolution data of TM; (b) Flowchart of NPP retrieved from coarse resolution data of MODIS

NPP_{MODIS} retrieved from coarse resolution data of MODIS can be corrected as follows.

$$NPP_{TM\,j} = NPP_{MODIS\,j} \times R_j \quad (8)$$

where $NPP_{TM\,j}$ is estimated NPP result for a pixel labeled as cover type j derived from fine resolution data of TM, $NPP_{MODIS\,j}$ is estimated NPP result for a pixel labeled as cover type j retrieved from coarse resolution data of MODIS, and R_j is a correction factor for scale effect.

So R_j is given by:

$$R_j = \frac{NPP_{TM\,j}}{NPP_{MODIS\,j}} \quad (9)$$

The relationship is theoretically as follows.

$$R_j = 1 - \sum_{i=1}^{n} C_{ij} F_{ij} \tag{10}$$

where F_{ij} is the fraction of nondominant cover type i (30 m resolution TM) in a pixel labeled as cover type j (1 km resolution TM), C_{ij} is the regression coefficient between R_j for a pixel labeled as cover type j and F_{ij} of nondominant cover type i in this pixel, and n is the amount of nondominant cover types in a pixel labeled as cover type j.

Based on above formulas and area fractions of nondominant cover types in every pixel, $\mathrm{NPP}_{\mathrm{MODIS}\,j\_\mathrm{corrected}}$ which has been corrected for scale effect can be achieved as follows.

$$\mathrm{NPP}_{\mathrm{MODIS}\,j\_\mathrm{corrected}} = \mathrm{NPP}_{\mathrm{MODIS}\,j} \times \left(1 - \sum_{i=1}^{n} C_{ij} F_{ij}\right) \tag{11}$$

4.2.2 Support Vector Machine (SVM) algorithm

When NPP is performed scaling in this paper, regression coefficient C_{ij} between R_j for a pixel labeled as cover type j and F_{ij} of nondominant cover type i in this pixel needs to be calculated as in Eq. (10) and Eq. (11). The precision of estimated C_{ij} has direct effect on successful accomplishment of spatial scaling method for the NPP model based on remote sensing. Therefore, SVM algorithm is used to make the regression model between R_j and F_{ij} in the paper. Comparing with the traditional statistical studying theories, it is found that fewer adjustment parameters, faster studying speed, optimum process and better spread capability from technical performance make SVM predominant. Studying results of SVM are often better than that of other pattern identifications and regression prediction methods. SVM has unique advantages in solving small sample, nonlinearity and higher dimensional pattern identification (Li *et al*., 2006).

In the paper, Libsvm is adopted as the SVM algorithm. Libsvm is a software package which has simple structure, easy usage and fast efficient support vector classification and regression. It can solve classification (e.g. C-SVM and nu-SVM), regression (e.g. e-SVM and nu-SVM), distributed estimation (one-class-SVM) and so on. Linearity, polynomial, RBF and S-shaped functions are four kinds of common kernel functions to be selected to solve multi-classification problems, choose parameters of cross validation, weight unbalance sample, estimate probability of multi-classification problems and so on (Lin, 2008). RBF kernel function is selected to run Libsvm program in this paper.

5 RESULTS

5.1 Land cover classification in the study area

Vegetation type can make great difference of estimating LUE and FPAR in the NPP model based on LUE, so it affects the precision of final NPP results. During NPP scaling, NPP at 30 m resolution is retrieved from 30 m resolution TM, and estimated NPP of a corresponding pixel at 1 km resolution can be acquired by calculating arithmetic mean value of NPP in many corresponding pixels at 30 m resolution. During the process, the principle of vegetation cover type with dominant area is used to define vegetation cover type of the pixel. In addition, various vegetation types have different contribution to NPP in scaling, and they have not linear or nonlinear relationship. So some errors are produced in scaling of vegetation classification data, and they are passed in NPP scaling. Through above analysis, estimated NPP results greatly rely on land cover type, classification precision needs to be enhanced to ensure successful accomplishment of NPP scaling. In this paper, based on literature of Dari County and field investigation, preprocessed MODIS at 1 km resolution and TM data at 30 m resolution of the study area in 2006 were respectively classified by high-precision synthetic classification method which can reflect the spatial distribution of vegetation. Fig.4 and Fig.5 show land cover classification maps retrieved from MODIS and TM data in the study area, respectively.

Land cover classification results in Dari County were statistically analyzed. The result shows that the area of middle density grassland is the largest (60.96% of the total area of the whole county). The second area is sparse grassland (16.81%). Sparse shrub covers 6.27% of the total area. Gobi and dense grassland cover approximate area (5.67% and 5.65%). Dense shrub only covers 3.25%. The area of forest in Dari County is sparse. Needle-leaved forest, needle-leaved and broad-leaved mixed forest and broad-leaved forest covers 0.46%, 0.40% and 0.53%, respectively.

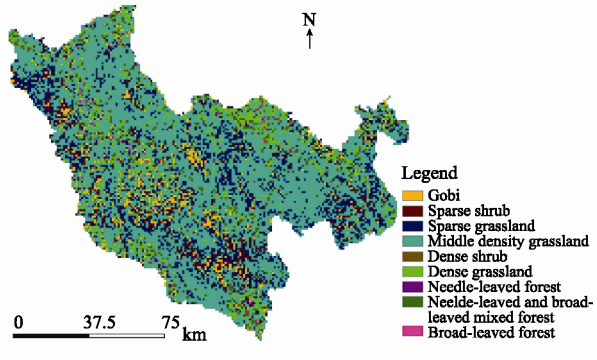

Fig. 4 Land cover classification map of the study area at 1 km resolution

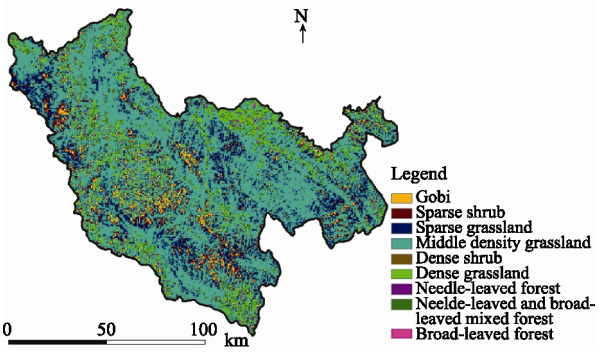

Fig. 5 Land cover classification map of the study area at 30 m resolution

5.2 Contrastive analysis of NPP_{TM} and NPP_{MODIS} map

NPP_{TM} (Fig.6) and NPP_{MODIS} (Fig.7) maps of the study area in July 2006 were respectively achieved based on the NPP model made in the paper. Fig.8 shows NPP_{TM} map at 1 km resolution which was converted from NPP_{TM} at 30 m resolution.

Fig.7 (NPP_{MODIS}) and Fig.8 (NPP_{TM}) at 1 km resolution were contrastively analyzed, and the difference between them was found. 30m resolution data and 1 km resolution data as input parameters of the NPP model are different, and during the course of the model calculating, the difference is enlarged due to nonlinearity of algorithm.

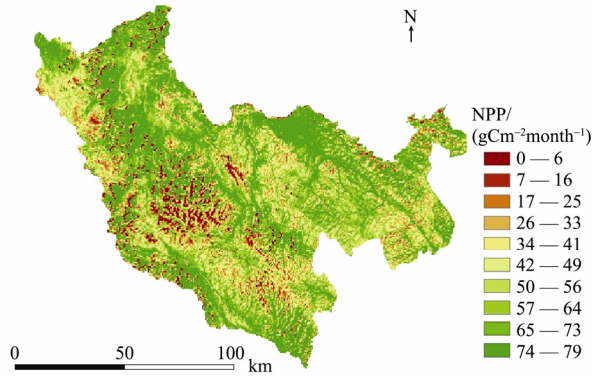

Fig. 6 NPP_{TM} map of the study area at 30 m resolution

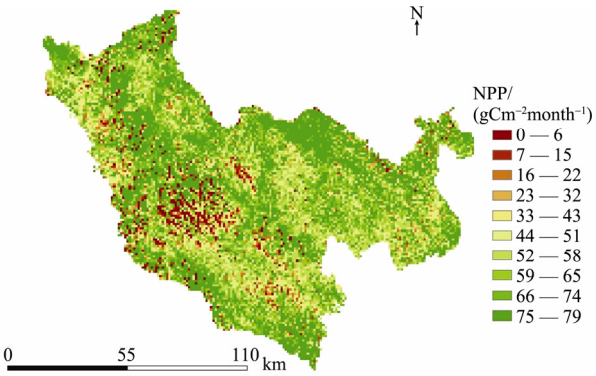

Fig. 7 NPP_{MODIS} map of the study area at 1 km resolution

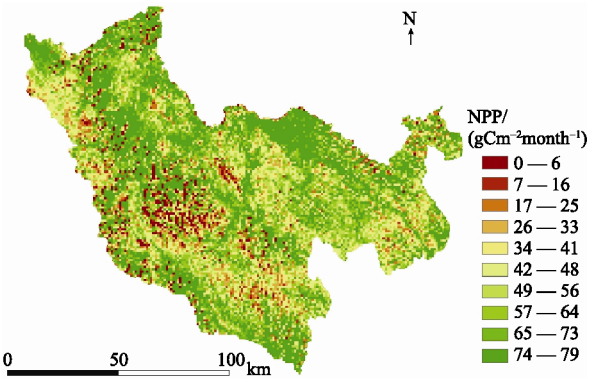

Fig. 8 NPP_{TM} map of the study area at 1 km resolution converted from NPP_{TM} by NPP scaling

According to statistics, the mean value of NPP_{TM} in the study area in July 2006 is 55 $gCm^{-2}month^{-1}$ (Fig.8). Mean values of NPP_{TM} for various vegetation types are 28 $gCm^{-2}month^{-1}$ for sparse shrub, 40 $gCm^{-2}month^{-1}$ for sparse grassland, 63 $gCm^{-2}month^{-1}$ for middle density grassland which is the main vegetation cover type in the study area, 67 $gCm^{-2}month^{-1}$ for dense shrub, 69 $gCm^{-2}month^{-1}$ for dense grassland, 72 $gCm^{-2}month^{-1}$ for needle-leaved forest, 75 $gCm^{-2}month^{-1}$ for needle-leaved and broad-leaved mixed forest, 78 $gCm^{-2}month^{-1}$ (the largest value) for broad-leaved forest, respectively.

Based on statistics analyzing to Fig.7, the average value of NPP_{MODIS} in the study area in July 2006 is 62 $gCm^{-2}month^{-1}$. Mean values of NPP_{MODIS} for various vegetation types are 34 $gCm^{-2}month^{-1}$ for sparse shrub, 45 $gCm^{-2}month^{-1}$ for sparse grassland, 65 $gCm^{-2}month^{-1}$ for middle density grassland which is the main vegetation cover type in the study area, 68 $gCm^{-2}month^{-1}$ for dense shrub, 71 $gCm^{-2}month^{-1}$ for dense grassland, 74 $gCm^{-2}month^{-1}$ for needle-leaved forest, 77 $gCm^{-2}month^{-1}$ for needle-leaved and broad-leaved mixed forest, 79 $gCm^{-2}month^{-1}$ (the largest value) for broad-leaved forest, respectively.

From the above calculating results, NPP_{MODIS} is bigger than NPP_{TM}, and variable range of the former is also bigger than that of the latter. Spatial heterogeneity and mixed pixels may be reasons. Nondominant cover types in the pixel are ignored in estimating NPP from MODIS data at the coarse resolution, while other cover types in many pixels are considered in estimating NPP from TM data at the fine resolution. The distribution characteristic of land cover type in the study area has a great effect on estimated NPP results, and it will lead that NPP_{MODIS} is bigger or smaller than NPP_{TM}. In this paper, the ratio of vegetation cover in Dari County is higher (94%). The primary vegetation cover is middle density grassland (60.96%), and the second is sparse grassland (16.81%). This distribution characteristic of vegetation leads that NPP_{MODIS} is a little bigger than NPP_{TM} in the whole study area.

5.3 Regression analysis between R_j and F_{ij}

During NPP scaling, the regression relationship between R_j for a pixel labeled as cover type j and F_{ij} of nondominant cover type i in this pixel needs to be quantified. In this paper, SVM algorithm is used to make the regression model between R_j and F_{ij}.

In the paper, RMSE is adopted to estimate the error between the calculated value and the real value. In the regression model of SVM, various model parameters determine different regression results, and model parameters include penalty coefficient C and parameter γ of RBF kernel function. Penalty coefficient C and parameter γ are defined by experiments. Selecting the value in the bigger extent is to acquire approximate optimum region, then the value is selected in the smaller extent. Considering synthetic performance and training time, values of parameter C

and γ are selected by the experiment while the error is less.

Fig.9 shows the scatter plot of $R_{j\_corrected}$ and 1-$F_{middle\ density\ grassland}$ estimated by the regression model based on SVM (for pixels labeled as middle density grassland). From Fig.9, the correlation between $R_{j\_corrected}$ and 1-$F_{middle\ density\ grassland}$ estimated by SVM regression model is higher (R^2=0.81).

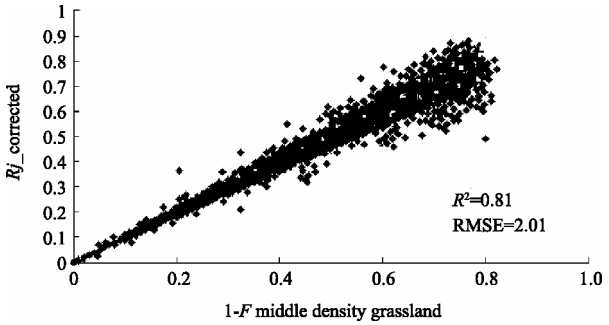

Fig. 9 Correlation analysis between $R_{j\_corrected}$ and 1-$F_{middle\ density\ grassland}$ for pixels labeled as middle density grassland

5.4 Correlation comparison between NPP_{TM} and NPP which has been uncorrected or corrected for scale effect

In NPP scaling, to multiply NPP_{MODIS} by $R_{j\_corrected}$ estimated from SVM regression model is used to calculate $NPP_{MODIS\_corrected}$ which has been corrected for scale effect. Fig.10 and Fig.11 are analysis maps of correlation comparison between NPP_{TM} and NPP which has been uncorrected or corrected for scale effect (samples used in the figures is different from samples adopted to make the regression model between R_j and F_{ij}). From the figures, the correlation between NPP_{TM} and uncorrected NPP_{MODIS} is lower (R^2=0.69; RMSE=3.47), while the correlation between NPP_{TM} and corrected $NPP_{MODIS\_corrected}$ is higher (R^2=0.84; RMSE=1.87).Therefore, NPP which has been corrected for scale effect has been greatly improved in both correlation and error.

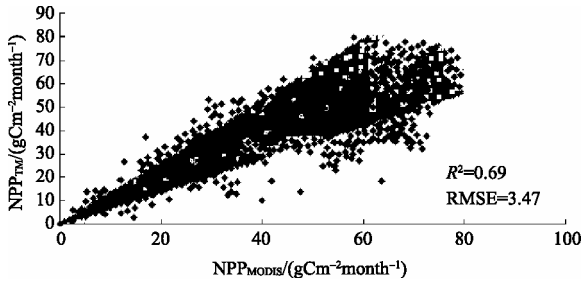

Fig. 10 Comparison between uncorrected NPP (NPP_{MODIS}) and NPP_{TM}

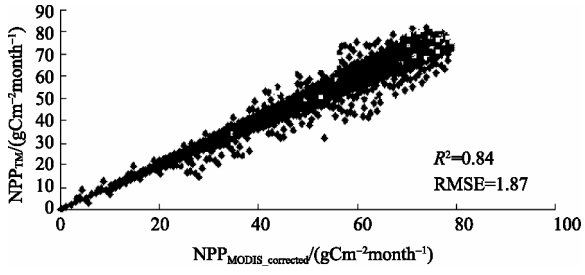

Fig. 11 Comparison between corrected NPP ($NPP_{MODIS\_corrected}$) and NPP_{TM}

6 CONCLUSION AND DISCUSSION

Spatial scaling of NPP model based on remote sensing is the main research content in this paper. The primary conclusions are drawn as follows.

(1) Contextural approach based on mixed pixels is ideal in conducting spatial scaling of NPP. The approach considers the effect of nondominant cover types in mixed pixels, and it used the theory of area fractions covered by various cover types to calculate land surface parameters at various resolutions. It realizes spatial scaling from NPP at the fine resolution to NPP at the coarse resolution based on remote sensing data. The study shows Contextural approach is a better method of scaling.

(2) The regression model between correction factor (R_j) of scale effect for a pixel labeled as cover type j and F_{ij} of nondominant cover type i in this pixel is made by support vector machine (SVM). Results show that SVM can efficiently convert realistic problems to high dimension characteristic space by nonlinear transform in order to construct the linear function to acquire optimum process result for limited samples. The correlation between $R_{j\_corrected}$ and 1-$F_{middle\ density\ grassland}$ estimated by SVM regression model is higher (R^2=0.81).

(3) The scaling model from fine resolution data of TM to coarse resolution data of MODIS is made, and NPP results retrieved from coarse resolution data of MODIS is corrected for scale effect. The correlation between NPP_{TM} and uncorrected NPP_{MODIS} is lower (R^2=0.69; RMSE=3.47), while the correlation between NPP_{TM} and corrected $NPP_{MODIS\_corrected}$ is higher (R^2=0.84; RMSE=1.87).Therefore, NPP which has been corrected for scale effect has been greatly improved in both correlation and error.

With the development of remote sensing technology, various kinds of spatial and temporal resolution sensors appear. This makes space scaling more important in quantified analysis of remote sensing. Land cover types are so stable for a period that they suit to be acquired by high spatial resolution images of remote sensing (e.g. Landsat). However, land surface variation by natural disaster (e.g. forest fire) needs to be found by high temporal resolution images of which occurring frequency is the same with that of natural disaster. Therefore, spatial scaling among remote sensing data at various resolutions in application needs to be solved. Due to the influence of sub-pixels heterogeneity, NPP retrieved from remote sensing data at coarse resolutions have considerable errors. In the study, the relative error of NPP derived from MODIS is about 13%. At present, this kind of error is common in regional and global NPP maps, but spatial scaling based on mixed pixels can efficiently reduce it. Contextural approach based on mixed pixels adopted in this

paper can also be used in spatial scaling among NPP at other resolutions.

REFERENCES

Anita S, Chen J M and Liu J. 2004. Spatial scaling of net primary productivity using subpixel information. *Remote Sensing of Environment*, **93**(1—2): 246—258

Chen J M, Liu J and Cihlar J. 1999. Daily canopy photosynthesis model through temporal and spatial scaling for remote sensing applications. *Ecological Modelling*, **124**(2—3): 99—119

Chen L F, Gao Y H, Li L, Liu Q H and Gu X F. 2007. Estimation of daily net primary productivity for forest based on white sky data of MODIS. *Science in China Ser. D*, **37**(11): 1515—1521

Ehleringer J R and Field C B. 1993. Scaling Physiological Processes: Leaf to Globe. Boston: Academic

Li G Z, Wang M and Zeng H J. 2006. An Introduction to Support Vector Machines. Beijing: Publishing House of Electronics Industry

Li X W, Wang J F, Wang J D and Liu Q H. 2001. Multi-Angle and Thermal Inferred Remote Sensing. Beijing: Science Press

Li Y N, Zhao X Q and Cao G M. 2004. Analyses on climates and vegetation productivity background at Haibei Alpine Meadow ecosystem research station. *Plateau Meteorology*, **23**(4): 558—567

Lin H D. 2008. A simple introduction to LIBSVM. Internet: www.csie.ntu.edu.tw/~cjlin/libsvm

Pierce L L and Running S W. 1995. The effects of aggregating sub-grid land surface variation on large-scale estimates of net primary production. *Landscape Ecology*, **10**(4): 239—253

Nemani R R, Keeling C D and Hashimoto H. 2003. Climate-driven increases in global terrestrial net primary production from 1982 to 1999. *Science*, **300**(5625): 1560—1563

Pu J Y, Li Y N, Zhao L and Yang S H. 2005. Seasonal changes of Kobresia Humilis Meadow biomass with climate factor. *Acta Agrestia Sinica*, **13**(3): 238—241

Raich J W, Rastetter E B and Melillo J M. 1991. Potential net primary productivity in South America: application of a global model. *Ecological Applications*, **1**(4): 399—429

Su L H, Li X W and Huang Y X. 2001. An review on scale in remote sensing. *Advance in Earth Sciences*, **16**(4): 544—548

Suleiman A and Crago R. 2004. Hourly and daytime evapotranspiration from grassland using radiometric surface temperatures. *Agronomy Journal*, **96**: 384—390

Tong Q X, Zhang B and Zheng L F. 2006. Hyperspectral Remote Sensing: Theory, Technology and Application. Beijing: Higher Education Press

Turner D P, Dodson R D and Marks D. 1996. Comparison of alternative spatial resolutions in the application of a spatially distributed biogeochemical model over complex terrain. *Ecological Modeling*, **90**(1): 53—67

Venturini V, Bisht G and Islam S. 2004. Comparison of evaporative fractions estimated from AVHRR and MODIS sensors over South Florida. *Remote Sensing of Environment*, **93**(1-2): 77—86

Wang L W, Wei Y X and Niu Z. 2008. Spatial and temporal variations of vegetation in Qinghai Province based on satellite data. *Journal of Geographical Sciences*, **18**(1): 73—84

Wang X Z, Tao B Z and Qiu W N. 2006. Advanced Surveying Adjustment. Beijing: Surveying and Mapping Press

Wen P Y, Shu G L and Guang S Z. 2007. Deriving a light use efficiency model from eddy covariance flux data for predicting daily gross primary production across biomes. *Agricultural and Forest Meteorology*, **143**(3—4): 189—207

Yang F T, Wang Q J and Shi S H. 1987. The allocation of the biomass and energy in Kobresia Humilis Meadow, Haibei district, Qinghai province. *Acta Phytoecologica Et Geobotanica Sinica*, **11**(2): 106—112

Zhang Y Q, Liu C M and Yu Q. 2004. Energy fluxes and the Priestley-Taylor parameter over winter wheat and maize in the North China Plain. *Hydrol. Process*, **18**: 2235—2246

Zhu W Q, Pan Y Z and He H. 2006. Simulation of maximum light use efficiency for main vegetation in China. *Chinese Science Bulletin*, **51**(6): 700—706

净初级生产力遥感估算模型空间尺度转换

王莉雯[1], 卫亚星[1], 牛 铮[2]

1. 辽宁师范大学 城市与环境学院, 辽宁 大连 116029;
2. 遥感科学国家重点实验室 中国科学院遥感应用研究所, 北京 100101

摘 要: 采用基于混合像元的结构分析方法和支持向量机(SVM)算法, 建立了高分辨率遥感数据(TM)向低分辨率遥感数据(MODIS)的尺度转换模型, 实现了由高分辨率遥感数据获得的 NPP 向低分辨率遥感数据获得的 NPP 的空间尺度转换。对低分辨率遥感数据(MODIS)估算的 NPP 结果进行了尺度效应校正。结果表明: SVM 回归模型模拟出的尺度效应校正因子 $R_{j\_corrected}$ 与 $1-F_{\text{中覆盖度草地}}$之间的相关性较高, R^2 达到 0.81。尺度效应校正前的 NPP_{MODIS} 与 NPP_{TM} 的相关性较低, R^2 仅为 0.69, RMSE 为 3.47; 尺度效应校正后的 $NPP_{MODIS\_corrected}$ 与 NPP_{TM} 的相关性较高, R^2 达到 0.84, RMSE 为 1.87。因此, 经过尺度效应校正后的 NPP 无论是在相关性还是在误差方面有了很大程度的提高。

关键词: 净初级生产力, 光能利用率模型, 遥感, 尺度转换, SVM

中图分类号: TP761 文献标志码: A

引用格式: 王莉雯, 卫亚星, 牛 铮. 2010. 净初级生产力遥感估算模型空间尺度转换. 遥感学报, **14**(6): 1074—1089
Wang L W, Wei Y X and Niu Z. 2010. Spatial scaling of net primary productivity model based on remote sensing. *Journal of Remote Sensing*. **14**(6): 1074—1089

1 引 言

遥感的优势在于频繁和持久地提供地表特征的面状信息。造成遥感问题复杂化的重要原因之一是地面测量数据、不同遥感测量方式获取的数据和应用需要的数据在尺度上的不一致。如何在不同尺度的数据间相互转换是尺度问题研究的核心之一(李小文等, 2001; Turner 等, 1996)。地球系统的空间异质性限制了从一个尺度到另一个尺度的信息转换, 这也是需要整合不同尺度上不同传感器遥感数据的重要原因(苏理宏等, 2001; Pierce 等, 1995)。

净初级生产力(NPP)空间尺度转换是指根据不同空间分辨率的数据所获得的 NPP 估算结果之间建立定量相关关系的分析转换过程。在这一空间转换过程中, 根据一定空间分辨率上估算的 NPP 预测另一空间分辨率上的 NPP 估值(Ehleringer & Field 1993)。根据低分辨率获得的地表参数中多存在误差, 这主要是因为传感器所获得的发射光谱信号是所对应的各种地物光谱信号的综合, 而且所获得的混合光谱信号与地表参数之间的定量相关关系难以准确确定。因此, 如何将某一尺度上获得的 NPP 通过一定的计算方法以较小的误差转换到另一空间尺度上, 这是急待解决的问题。尺度转换中的非线性问题和土地覆盖类型对尺度转换的影响是 NPP 空间尺度转换的两个主要问题(Chen 等, 1999; Nemani 等, 2003)。

采用基于混合像元的结构分析方法, 建立了高分辨率遥感数据(TM)向低分辨率遥感数据(MODIS)的尺度转换模型。采用支持向量机(SVM)算法, 建立了标注为 j 土地覆盖类型像元的尺度效应校正因子 R_j 和该像元中非主要土地覆盖类型 i 所对应的 F_{ij}(在高分辨率遥感数据 TM 中非主要土地覆盖类型 i 在标注为 j 土地覆盖类型的像元中所占的比例)之间的回归模型。实现了由高分辨率遥感数据获得的 NPP 向低分辨率遥感数据获得的 NPP 的空间尺度转换, 并对低分辨率遥感数据(MODIS)估算的 NPP 结果进行了尺度效应校正。

收稿日期: 2009-12-14; 修订日期: 2010-06-07

基金项目: 辽宁省教育厅科学研究一般项目(编号: L2010226); 教育部人文社会科学重点研究基地项目(编号: 08JJD790142); 辽宁省教育厅高等学校创新团队项目(编号: 2007T095)和国家重点基础研究发展规划(973)项目(编号: 2007CB714406)。

第一作者简介: 王莉雯(1971—), 女, 博士, 主要从事遥感和 GIS 的应用研究。发表了 10 多篇学术论文, 其中 4 篇被 SCI 收录。E-mail: wlw9585@163.com。

2 研究区域

研究区选定在青海省达日县。达日县位于青海省东南部(图 1), 东经 98°15′29″—100°32′41″, 北纬 32°36′42″—34°15′20″。全县东西长 162 km, 南北宽 126km, 总面积 14600 km^2。全县平均海拔 4426 m。县境属高寒半湿润性气候, 除冷暖两季外, 没有明显的四季之分。全县平均气温在–0.1℃至 –3.5℃之间。全县年降水量 569 mm, 多集中于 6 月—9 月。年蒸发量 1119.07 mm, 最高月份在 5 月, 达 150.1 mm。达日县植被以高山草甸为主, 其中以小嵩草、矮嵩草、藏嵩草等居多, 广泛分布于阳坡和宽谷。高山灌丛以金露梅、高山柳等较重要, 大部分分布在阴坡。

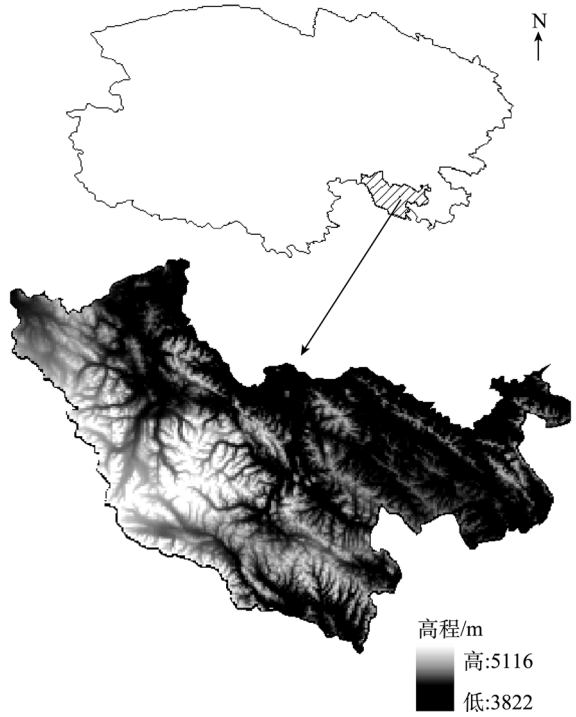

图 1 研究区位置示意图

3 数据准备

采用的遥感数据主要包括 MODIS 数据产品和 TM 数据。在 NPP 模型构建过程中用于计算和验证的 MODIS 数据产品主要包括 MOD09A1(全球每 8 天 500m 分辨率地表反射率 L3 级产品)、MOD11A2(全球每 8 天 1km 分辨率地表温度、发射率 L3 级产品)、MOD12Q1(全球每年 1km 分辨率土地覆盖类型 L3 级产品)、MOD13A2(全球每 16 天 1km 分辨率植被指数 L3 级产品)、MOD15A2(全球每 8 天 1km 分辨率 LAI、FPAR L4 级产品)、MOD17A2(全球每 8 天 1km 分辨率总初级生产力 L4 级产品)和 MOD43B3(全球每 16 天 1km 分辨率反照率 L3 级产品), 由美国 LP DAAC(Land Process Distributed Active Archive Center, U.S.A)提供。对上述的 MODIS 数据产品做了几何纠正、去云处理、修复异常值像元等预处理工作。研究中使用的 TM 数据来源于 Landsat-5。选择了 5 景 TM 影像, 轨道号分别为 132/37、133/36、133/37、134/36 和 134/37, 影像范围覆盖了整个青海省的达日县。对所选用的 TM 数据进行了几何校正和大气校正。所使用 MODIS 和 TM 影像的成像时间在 2006 年 7 月, 影像成像日期基本反映了一年中植被生长状况较好的时期, 具有较好的代表性。

研究所用气象数据由国家气象中心提供, 包括青海省内 41 个站点的平均气温、降水量、平均空气相对湿度、日照时数、气压和平均风速。研究采用了青海省植被覆盖分类方法(Wang 等, 2008)。所用的 DEM 数据是从国际热带农业中心网(Centro International de Agriculture Tropical, CIAT, http://srtm.csi. cgiar. org/)下载的。

地面实测数据来源于 2006 年 7 月青海省达日县的实验, 实验仪器为 LI-6400 便携式光合仪、LAI-2000 植物冠层分析仪和 SUNSCAN 冠层分析系统等。

4 方 法

4.1 NPP 估算模型构建

基于光能利用率的 NPP 估算模型的典型特点是直接利用遥感数据, 极大地提高模型的时空分辨率, 使大范围的 NPP 估算及其时空动态监测成为可能。构建的 NPP 遥感估算模型是基于光能利用率原理, 借鉴了 MODIS-PSN、CASA、GLO-PEM、VPM 等光能利用率 NPP 模型的优点, 同时考虑了特定区域其植物光能利用率和环境因素的典型特点, 建立了针对特定区域的 NPP 遥感估算模型。NPP 遥感估算模型的具体公式为:

$$\mathrm{NPP}=\mathrm{PAR}\times\mathrm{FPAR}\times\min\left(T_s,\ W_s\right)\times\varepsilon_{\max} \qquad (1)$$

式中, PAR 为光合有效辐射, 单位为 MJ·m^{-2}; FPAR 为植被层对入射光合有效辐射(PAR)的吸收分量, 无单位; T_s 为温度胁迫因子, 无单位; W_s 为水分胁迫因子, 无单位; $\varepsilon_{\max}$ 为最大光能利用率, 单位为 gC·MJ^{-1}。

4.1.1 最大光能利用率$\varepsilon_{\max}$估算

最大光能利用率$\varepsilon_{\max}$的取值因不同的植被类型

有所不同, 它的取值大小对最终的 NPP 估算结果影响很大。采用极大似然估计准则(王新洲等, 2006; 朱文泉等, 2006), 根据研究区相关文献资料和野外实验数据(李英年等, 2004; 杨福囤等, 1987; 蒲继延等, 2005), 采用式(2)模拟了各植被类型的最大光能利用率$\varepsilon_{\max}$的优化值。

$$\hat{\varepsilon}_{\max}=\frac{\sum_{i=1}^{n}x_i y_i-n\bar{x}\bar{y}}{\sum_{i=1}^{n}x_i{}^2-n\bar{x}^2} \tag{2}$$

式中, $\hat{\varepsilon}_{\max}$为最大光能利用率$\varepsilon_{\max}$的估值, x_i为某种植被类型 PAR×FPAR×min(T_s, W_s)的计算值; y_i为某种植被类型的 NPP 值; i为某种植被类型的样本数; n为某种植被类型的最大样本数; $\bar{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$; $\bar{y}=\frac{1}{n}\sum_{i=1}^{n}y_i$。

为了提高各种草地遥感分类类型最大光能利用率$\varepsilon_{\max}$的估算精度, 在采用式(2)模拟出研究区主要草地建群种类型和主要灌丛建群种类型的最大光能利用率的取值后, 再采用式(3)估算出各种草地遥感分类类型和各种灌丛遥感分类类型的最大光能利用率$\varepsilon_{\max}$。

$$\varepsilon_{\max}=\sum_{i=1}^{n}c_i\left(\varepsilon_{\max}\right)_i \tag{3}$$

式中, c_i为第 i 种草种类型(或灌丛类型)所占的面积比例, $(\varepsilon_{\max})_i$为第 i 种草种类型(或灌丛类型)的最大光能利用率。

经计算, 研究区主要植被类型最大光能利用率为: 阔叶林 0.908 gC·MJ^{-1}, 针阔混交林 0.725 gC·MJ^{-1}, 针叶林 0.645 gC·MJ^{-1}, 高覆盖度草地 0.312 gC·MJ^{-1}, 郁闭灌丛 0.538 gC·MJ^{-1}, 中覆盖度草地 0.234 gC·MJ^{-1}, 草地耕地混合体 0.208 gC·MJ^{-1}, 低覆盖度草地 0.175 gC·MJ^{-1}, 稀疏灌丛 0.114 gC·MJ^{-1}。

4.1.2 FPAR 估算

根据研究区数据构建了 NDVI 和 FPAR 之间的关系方程式, 设定无植被覆盖地区, FPAR 最小, 其值为 0; 植被完全覆盖地区, FPAR 达到最大值, 即 0.95(陈良富等, 2007):

$$\mathrm{FPAR}=\min\left[\frac{\mathrm{SR}}{\mathrm{SR}_{\max}-\mathrm{SR}_{\min}}-\frac{\mathrm{SR}_{\min}}{\mathrm{SR}_{\max}-\mathrm{SR}_{\min}},0.95\right] \tag{4}$$

$$\mathrm{SR}=\frac{1+\mathrm{NDVI}}{1-\mathrm{NDVI}} \tag{5}$$

式中, $\mathrm{SR}_{\min}$在所有植被类型中取同一个值 1.06, $\mathrm{SR}_{\max}$的取值与植被类型有关, 即: 阔叶林取 6.14; 针阔混交林取 3.88; 针叶林取 3.44; 高覆盖度草地取 3; 郁闭灌丛取 3.27; 中覆盖度草地取 2.17; 草地耕地混合体取 1.94; 低覆盖度草地取 1.78; 稀疏灌丛取 1.70。

4.1.3 水分胁迫因子 W_s 的估算

采用一种经实践证明较为成熟的算法——蒸发比(EF)(Suleiman & Crago, 2004; Venturini 等, 2004; Wen 等, 2006; Zhang 等, 2004), 计算水分胁迫因子W_s, 计算公式如下:

$$W_s=\mathrm{EF}=\frac{\mathrm{LE}}{\mathrm{LE}+H} \tag{6}$$

式中, LE 为潜热通量单位为 W·m^{-2}; H为显热通量, 单位为 W·m^{-2}。

图 2 为蒸发比 EF 计算流程图。

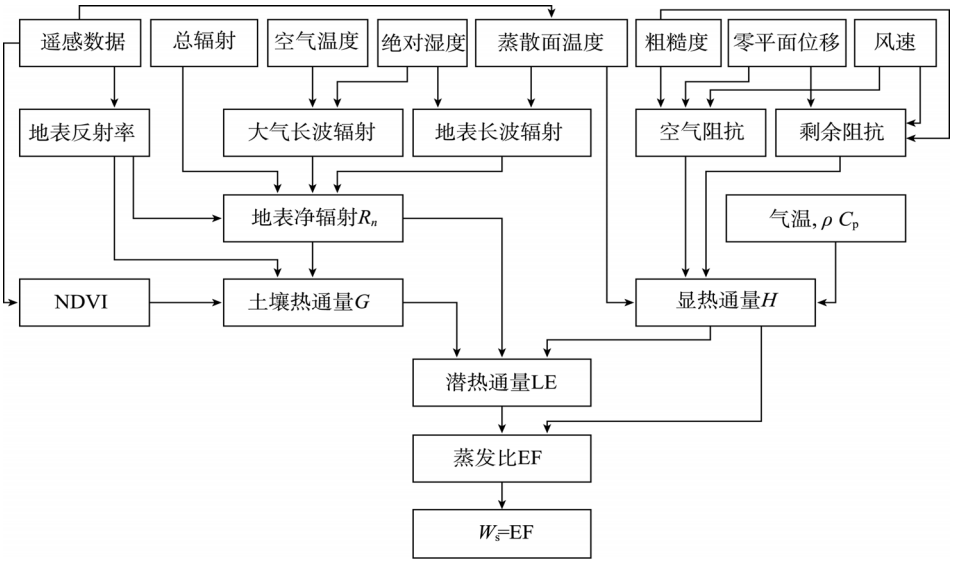

图 2 蒸发比 EF 计算流程图

4.1.4 温度胁迫因子 T_s 的估算

采用 Raich 等(1991)建立的陆地生态模型(the terrestrial ecosystem model, TEM)中的方法来计算 T_s:

$$T_s = \frac{(T-T_{min})(T-T_{max})}{(T-T_{min})(T-T_{max})-(T-T_{opt})^2} \quad (7)$$

式中, T 为大气温度, T_{min}、T_{opt} 和 T_{max} 分别为光合作用的最低、最适和最高温度。如果大气温度低于 T_{min} 或高于 T_{max}, 将 T_s 定为 0。在本研究中, T_{min} 和 T_{max} 分别被定为 0 和 36℃。同时, 假设植物已经适应了其所生长的环境的温度, T_{opt} 可定为生长季的长期平均温度。

4.2 NPP 空间尺度转换

4.2.1 基于混合像元的结构分析方法

根据影响尺度转换的主要问题, 采用 Anita 等人(2004)提出的基于混合像元的结构分析方法进行 NPP 的空间尺度转换, 其原理如下:

混合像元普遍存在于遥感图像中, 它的存在是各种地表参数的计算精度难以达到使用要求的主要原因。为了提高遥感应用的精度, 必须解决混合像元的分解问题, 使遥感应用由像元级达到亚像元级。因此, 低分辨率遥感图像中的一个像元可以认为是高分辨率遥感图像中对应坐标位置的多个像元的混合像元, 而该高分辨率遥感图像中这些像元可以认为是低分辨率遥感图像中该像元的亚像元(Chen 等, 1999; 童庆禧等, 2006)。

基于结构分析的空间尺度转换方法是采用了分析各土地覆盖类型所占面积比的原理计算不同分辨率下的地表参数。遥感分类一般是将一个像元定义为一种面积占优的主要土地覆盖类型, 在计算地表参数时忽略了该像元中其他非主要土地覆盖类型的影响。这样就导致在低分辨率影像中每个像元只被标注为一种主要土地覆盖类型, 从而使最终的计算结果存在较大的不确定性(Chen 等, 1999)。比较理想的做法是, 低分辨率影像的地表参数计算应该取相应高分辨率影像对应坐标位置地表参数的平均值。

在进行 NPP 空间尺度转换时, 假定由高分辨率 TM 估算出的 NPP_{TM} 代表了地表植被 NPP 的真实值(图 3(a)); 由低分辨率 MODIS 估算出的 NPP_{MODIS} 是需要经过尺度效应纠正的 NPP 近似值(图 3(b))。研究假设 1 个 MODIS 像元(1 km×1 km)包含了 1089 个 TM 像元(30 m×30 m)。由于存在混合像元, 1089 个 TM 像元包含了多种土地覆盖类型。在对应坐标位置的 1 个 MODIS 像元, 则被标注为一种面积占优的主要土地覆盖类型, 其他的非主要土地覆盖类型的信息被忽略了。将覆盖相同区域、不同分辨率的遥感数据 TM 和 MODIS 分别作为模型的输入参数, 利用本文所构建的基于光能利用率的 NPP 遥感估算模型模拟, 分别得到该区域的 NPP 模拟结果 $NPP_{TM'}$ 和 NPP_{MODIS}(图 3)。需要注意的是, 为了对比分析两个尺度下估算的 NPP 结果, 需要将两个尺度下估算的 NPP 结果转换为相同的空间分辨率, 即需要将由高分辨率 TM 估算出的 $NPP_{TM'}$ 的分辨率转换为 1km。在将分辨率为 30m 的 $NPP_{TM'}$ 转换为分辨率为 1km 的 NPP_{TM} 时, 采用取 $NPP_{TM'}$ 影像中相应坐标位置多个像元估算出的 NPP 的算术平均值, 作为 NPP_{TM} 影像中对应的一个像元的 NPP 估值。由于模型输入参数 TM 和 MODIS 空间分辨率的差异, 使得计算结果 NPP_{TM} 和 NPP_{MODIS} 之间存在较大的误差。分析两个尺度下计算结果产生差异的原因, 主要是由于相对于高分辨率 TM, 低分辨率 MODIS 中存在大量的混合像元, 混合像元中非主要土地覆盖类型的信息被忽略了。根据高分辨率 TM 的估算结果 NPP_{TM}, 对低分辨率 MODIS 计算的结果 NPP_{MODIS}, 进行尺度效应校正得到 $NPP_{MODIS\,j\_corrected}$, 提高了低分辨率遥感数据估算 NPP 的精度。

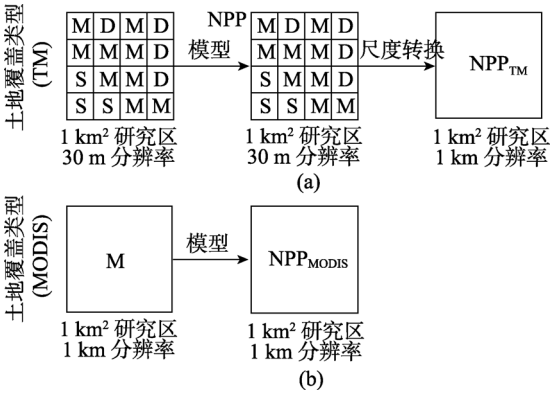

图 3 高分辨率和低分辨率遥感数据 NPP 计算流程图

(a) 高分辨率数据 TM 估算 NPP 尺度转换流程; (b) 低分辨率数据 MODIS 估算 NPP 流程

采用下述方法对由低分辨率 MODIS 估算的 NPP_{MODIS} 结果进行尺度效应校正:

$$NPP_{TM\,j} = NPP_{MODIS\,j} \times R_j \quad (8)$$

式中, $NPP_{TM\,j}$ 为土地覆盖类型为 j 的像元由高分辨率 TM 估算出的 NPP 结果, $NPP_{MODIS\,j}$ 为土地覆盖类型为 j 的像元由低分辨率 MODIS 估算出的 NPP 结果, R_j 为尺度效应校正因子。

因此, R_j 可由式(a)计算得到:

$$R_j = \frac{NPP_{TM\,j}}{NPP_{MODIS\,j}} \quad (9)$$

在理论上又存在如下关系式:

$$R_j = 1 - \sum_{i=1}^{n} C_{ij} F_{ij} \tag{10}$$

式中, F_{ij} 为在高分辨率 TM(分辨率为 30 m)中非主要土地覆盖类型 i 在标注为 j 土地覆盖类型的像元(将高分辨率 TM 的分辨率转换为 1km 时所对应的影像)中所占的比例, C_{ij} 为标注为 j 土地覆盖类型像元的尺度效应校正因子 R_j 与该像元中非主要土地覆盖类型 i 所对应的 F_{ij} 的回归系数, n 为标注为 j 土地覆盖类型像元内非主要土地覆盖类型的数目。

根据上述公式和已知的每个像元中非主要土地覆盖类型所占的面积比例, 用式(11)计算经尺度效应校正的 $\mathrm{NPP}_{\mathrm{MODIS}\,j\_\mathrm{corrected}}$:

$$\mathrm{NPP}_{\mathrm{MODIS}\,j\_\mathrm{corrected}} = \mathrm{NPP}_{\mathrm{MODIS}\,j} \times \left(1 - \sum_{i=1}^{n} C_{ij} F_{ij}\right) \tag{11}$$

4.2.2 支持向量机(SVM)算法

在进行 NPP 空间尺度转换时, 需要计算出标注为 j 土地覆盖类型像元的尺度效应校正因子 R_j 与该像元中非主要土地覆盖类型 i 所对应的 F_{ij} 的回归系数 C_{ij}, 如公式(10) (11)中的回归系数 C_{ij}。C_{ij} 参数估算的精度, 直接影响 NPP 遥感估算模型空间尺度转换方法的成功实施。在本文中采用支持向量机(SVM)算法建立了 R_j 和 F_{ij} 之间的回归模型。与传统的统计学习理论相比, 其性能的突出优越性在于: 调节参数较少, 学习速度快, 全局最优, 技术性能尤其使推广能力明显提高, 其学习结果经常明显好于其他它的模式识别和回归预测方法, 在解决小样本、非线性及高维模式识别问题中表现出许多特有的优势(李国正等, 2006)。

采用的 SVM 算法是 Libsvm。Libsvm 是一个结构较为简单、易于使用、快速有效的支持向量分类和回归的软件包, 可以解决分类问题(包括 C-SVM、nu-SVM)、回归问题(包括 e-SVM、nu-SVM)以及分布估计(one-class-SVM)等问题, 提供了线性、多项式、径向基和 S 形函数 4 种常用的核函数供选择, 有效地解决多类问题、交叉验证选择参数、对不平衡样本加权、多类问题的概率估计等(林弘德, 2008)。在运行 Libsvm 程序时选用了 RBF 核函数。

5 结果分析

5.1 研究区土地覆盖分类

植被类型会使得基于光能利用率的 NPP 估算模型中主要参数因子光能利用率和 FPAR 的估算存在很大的差异, 因此植被类型会影响最终 NPP 估算结果的精度。在对 NPP 进行空间尺度转换时, 从 30 m 分辨率 TM 数据计算得到了 30 m 分辨率 NPP, 通过计算 30 m 分辨率相应位置多个像元 NPP 的算术平均, 得到 1 km 分辨率对应位置一个像元的 NPP 模拟值。在这一过程中, 需要采用植被覆盖类型面积占优原则, 确定这个像元的植被覆盖类型。另外, NPP 空间尺度转换过程中, 不同的植被类型对 NPP 贡献是不一样的, 他们之间不存在线性或者非线性的关系。植被分类数据在进行尺度转换时产生了一定程度的误差, 该误差在 NPP 尺度转换过程中进行传递。从上面的分析看出, NPP 的估算结果在很大程度上依赖土地覆盖类型, 需要提高遥感数据的分类精度, 才能保证 NPP 空间尺度转换方法的成功运用。本文采用能突出植被空间分布状况的高精度综合分类方法, 在达日县历史资料和野外样点调查的支持下, 对做过预处理的 2006 年覆盖研究区的空间分辨率为 1 km 的 MODIS 数据和空间分辨率为 30 m 的 TM 数据分别进行了分类。从 MODIS 和 TM 数据生成的研究区土地覆盖分类图分别如图 4 和图 5。

对研究区达日县土地覆盖分类结果进行统计分析。结果表明, 中覆盖度草地的面积最大, 约占全县

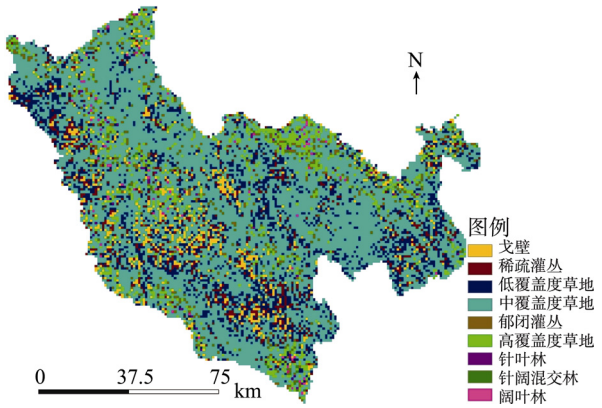

图 4 研究区 1km 分辨率土地覆盖分类图

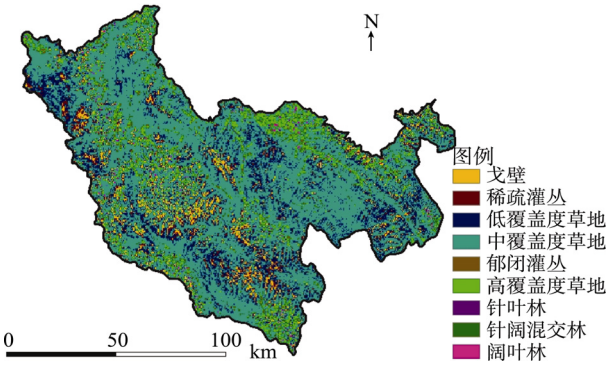

图 5 研究区 30m 分辨率土地覆盖分类图

总面积占 60.96%；其次为低覆盖度草地，占 16.81%；稀疏灌丛占 6.27%；戈壁和高覆盖度草地的面积近似，分别占 5.67%和 5.65%；郁闭灌丛仅占 3.25%；该县森林面积稀少，针叶林、针阔混交林和阔叶林分别占 0.46%、0.40%和 0.53%。

5.2 NPP_{TM} 和 NPP_{MODIS} 分布图的对比分析

首先，采用本文所构建的基于光能利用率的 NPP 遥感估算模型模拟，分别得到该区域 2006 年 7 月的 NPP 模拟结果 $NPP_{TM'}$(图 6)和 NPP_{MODIS}(图 7)。

再将分辨率为 30 m 的 $NPP_{TM'}$转换为分辨率为 1km 的 NPP_{TM}(图 8)。

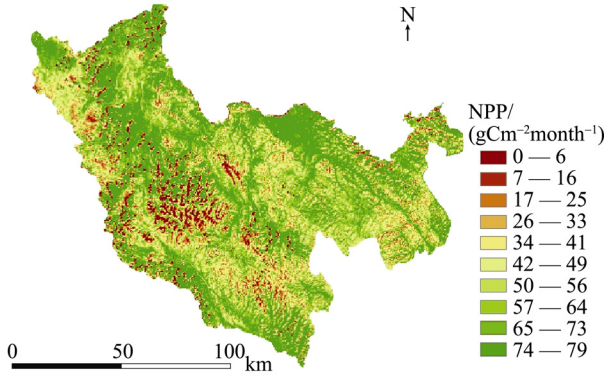

图 6 研究区 30m 分辨率 NPP 分布图($NPP_{TM'}$)

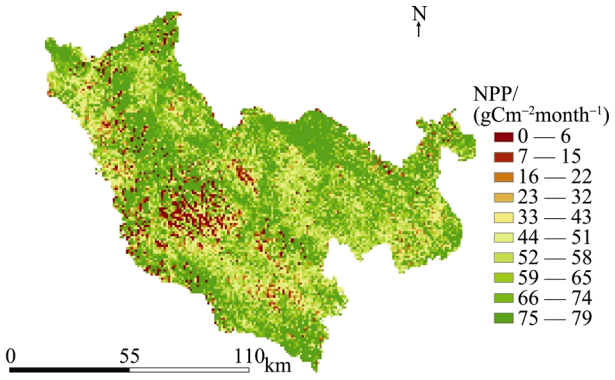

图 7 研究区 1km 分辨率 NPP 分布图(NPP_{MODIS})

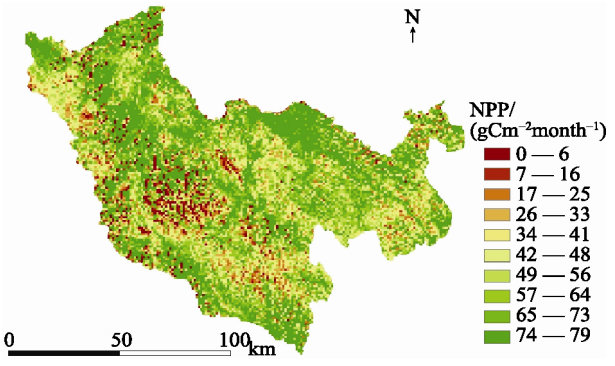

图 8 研究区由 $NPP_{TM'}$经 NPP 空间尺度转换算法得到的 1km 分辨率 NPP 分布图(NPP_{TM})

将分辨率 1km 的图 8(NPP_{TM})和图 7(NPP_{MODIS})对比分析，发现他们二者存在差异。这是由于作为 NPP 遥感估算模型输入参数的分辨率 30 m 和 1 km 本身存在差异，在模型的运算过程中，由于算法的非线性，加剧了这种差异。

经统计，整个研究区 2006 年 7 月 NPP_{TM} 的平均值为 55 $gCm^{-2}month^{-1}$(图 8)。其中，稀疏灌丛 NPP_{TM} 的平均值为 28 $gCm^{-2}month^{-1}$；低覆盖度草地的平均 NPP_{TM} 为 40 $gCm^{-2}month^{-1}$；试验区域的主要植被覆盖类型是中覆盖度草地，NPP_{TM} 的平均值为 63 $gCm^{-2}month^{-1}$；郁闭灌丛 NPP_{TM} 的平均值为 67 $gCm^{-2}month^{-1}$；高覆盖度草地的平均 NPP_{TM} 为 69 $gCm^{-2}month^{-1}$；阔叶林、针阔混交林和针叶林的平均 NPP_{TM} 值最高，分别为 78、75 和 72 $gCm^{-2}month^{-1}$。

同时，对图 7 进行统计分析，得出整个研究区 2006 年 7 月份 NPP_{MODIS} 的平均值为 62 gCm^{-2} $month^{-1}$。其中，稀疏灌丛 NPP_{MODIS} 的平均值为 34 $gCm^{-2}month^{-1}$；低覆盖度草地的平均 NPP_{MODIS} 为 45 $gCm^{-2}month^{-1}$；研究区的主要植被覆盖类型中覆盖度草地的 NPP_{MODIS} 平均值为 65 $gCm^{-2}month^{-1}$；郁闭灌丛 NPP_{MODIS} 的平均值为 68 $gCm^{-2}month^{-1}$；高覆盖度草地的平均 NPP_{MODIS} 为 71 $gCm^{-2}month^{-1}$；阔叶林、针阔混交林和针叶林的平均 NPP_{MODIS} 值最高，分别为 79、77 和 74 $gCm^{-2}month^{-1}$。

从上面的计算结果看出，NPP_{MODIS} 的值略高于 NPP_{TM}，而且发现前者的波动范围大于后者。这可能是因为，由于空间异质性和混合像元现象，低分辨率 MODIS 数据在估算 NPP 时忽略了像元中非主要土地覆盖类型的信息，而高分辨率 TM 数据在估算 NPP 时考虑了较多的像元中其他土地覆盖类型的信息。研究区土地覆盖类型的分布特点会在很大程度上影响 NPP 的估算结果，根据研究区各种植被类型分布特点的不同，会导致 NPP_{MODIS} 大于(或小于)NPP_{TM}。在本文中，研究区达日县的植被覆盖率比较高，达到 94%，而且主要的植被覆盖类型为中覆盖度草地(60.96%)，其次为低覆盖度草地(16.81%)，这种植被的分布特征，使得整个试验区域 NPP_{MODIS} 的值略大于 NPP_{TM}。

5.3 R_j 和 F_{ij} 的回归分析

在 NPP 空间尺度转换过程中，需要确定标注为 j 土地覆盖类型像元的尺度效应校正因子 R_j 与该像元中非主要土地覆盖类型 i 所对应的 F_{ij} 的回归关系，采用支持向量机(SVM)算法建立了 R_j 和 F_{ij} 之间的回归模型。

采用均方根误差(RMSE)来衡量估计值与真实值之间的误差。在 SVM 回归模型中，不同的模型参数决定了不同的回归效果，模型参数包括惩罚参数 C 和 RBF 核函数的参数 γ。通过实验来确定惩罚参数 C 和参数 γ 的值，先在较大范围取值，确定大致的最优区域后，再在较小范围内取值。从综合性能和训练时间两个方面考虑，通过实验选择误差较小时参数 C 和 γ 的取值。

图 9 是采用 SVM 算法所建回归模型模拟出的尺度效应校正因子 $R_{j\_corrected}$ 与 $1-F_{\text{中覆盖度草地}}$ 的散点分布图(以标注为中覆盖度草地的像元为例)。从图 9 看出，SVM 回归模型模拟出的尺度效应校正因子 $R_{j\_corrected}$ 与 $1-F_{\text{中覆盖度草地}}$ 之间的相关性较高，R^2 达到 0.81。

5.4 尺度效应校正前后的 NPP 与 NPP_{TM} 的相关对比

在 NPP 空间尺度转换过程中，采用 SVM 回归模型模拟出的尺度效应校正因子 $R_{j\_corrected}$ 与 NPP_{MODIS} 相乘，计算得到尺度效应校正后的 $NPP_{MODIS\_corrected}$。图 10 与图 11 分别是尺度效应校正前后的 NPP 与 NPP_{TM} 的相关性对比分析图(图 9 所用样本与建立 R_j 和 F_{ij} 之间回归模型时所用样本是不同的)。由图 9 可知，尺度效应校正前的 NPP_{MODIS} 与 NPP_{TM} 的相关性较低，R^2 仅为 0.69，RMSE 为 3.47；尺度效应校正后的 $NPP_{MODIS\_corrected}$ 与 NPP_{TM} 的相关性较高，R^2 达到 0.84，RMSE 为 1.87。因此，经过尺度效应校正后的 NPP 在相关性及误差方面有了很大程度的提高。

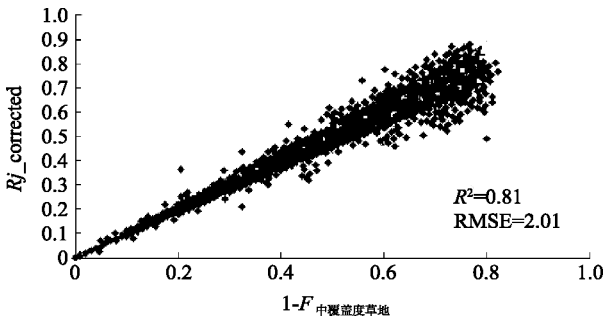

图 9 标注为中覆盖度草地的像元 $R_{j\_corrected}$ 与 $1-F_{\text{中覆盖度草地}}$ 的相关分析

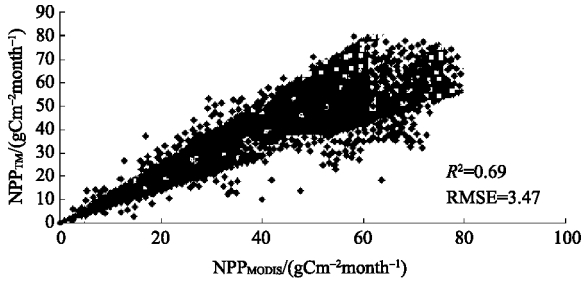

图 10 尺度效应校正前的 NPP(NPP_{MODIS})与 NPP_{TM} 的对比

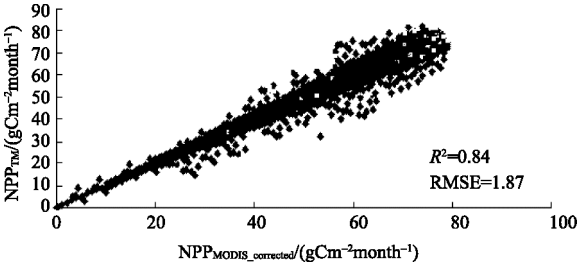

图 11 尺度效应校正后的 NPP($NPP_{MODIS\_corrected}$)与 NPP_{TM} 的对比

6 结论与讨论

NPP 遥感估算模型空间尺度转换方法是本文的主要研究内容，主要结论为：

(1) 基于混合像元的结构分析方法是进行 NPP 空间尺度转换研究的比较理想的方法。该方法考虑了混合像元中非主要土地覆盖类型的影响，采用分析各土地覆盖类型所占面积比的原理来计算不同分辨率下的地表参数，实现了由高分辨率遥感数据向低分辨率遥感数据获得的 NPP 之间的空间尺度转换。研究表明该方法是一种比较好的尺度转换方法。

(2) 采用支持向量机(SVM)算法建立了标记为 j 地表覆盖类型像元的尺度效应校正因子 R_j 和该像元中非主要地表覆盖类型 i 所对应的 F_{ij} 之间的回归模型。计算结果表明，SVM 算法能够有效地将实际问题通过非线性变换转换到高维的特征空间来构造线性函数，得到有限样本的全局最优解，SVM 回归模型模拟出的尺度效应校正因子 $R_{j\_corrected}$ 与 $1-F_{\text{中覆盖度草地}}$ 之间的相关性较高，R^2 达到 0.81。

(3) 建立了高分辨率遥感数据(TM)向低分辨率遥感数据(MODIS)的尺度转换模型，并对低分辨率遥感数据(MODIS)估算的 NPP 结果进行了尺度效应校正。尺度效应校正前的 NPP_{MODIS} 与 NPP_{TM} 的相关性较低，R^2 仅为 0.69，RMSE 为 3.47；尺度效应校正后的 $NPP_{MODIS\_corrected}$ 与 NPP_{TM} 的相关性较高，R^2 达到 0.84，RMSE 为 1.87。因此，经过尺度效应校正后的 NPP 无论是在相关性还是在误差方面都有了很大程度的提高。

随着遥感技术的发展出现了多种不同空间和时间分辨率的传感器，这使得空间尺度转换过程在遥感定量分析中突显其重要性。土地覆盖类型在一定时间段较为稳定，所以适合使用高空间分辨率遥感影像(如 Landsat)来获取。但是，自然灾害(如林火)

造成的地表变化需要与之发生频率相适合的高时间分辨率影像来获知。因此, 不同分辨率的遥感数据结合应用需要解决它们之间的空间尺度转换问题。由于亚像元空间异质性的影响, 由低分辨率遥感数据获得的 NPP 存在相当量的偏差。在本研究中, 由 MODIS 数据获得的 NPP 其相对偏差大约为 13%。这种偏差在目前区域和全球 NPP 分布图中应该较为普遍。而基于混合像元信息的空间尺度转换方法可以有效地降低该种偏差。本文采用的基于混合像元的结构分析方法也适用于其他分辨率获得的 NPP 之间的空间尺度转换。

REFERENCES

Anita S, Chen J M and Liu J. 2004. Spatial scaling of net primary productivity using subpixel information. *Remote Sensing of Environment*, **93**(1—2): 246—258

Chen J M, Liu J and Cihlar J. 1999. Daily canopy photosynthesis model through temporal and spatial scaling for remote sensing applications. *Ecological Modelling*, **124**(2—3): 99—119

Chen L F, Gao Y H, Li L, Liu Q H and Gu X F. 2007. Estimation of daily net primary productivity for forest based on white sky data of MODIS. *Science in China Ser. D*, **37**(11): 1515—1521

Ehleringer J R and Field C B. 1993. Scaling Physiological Processes: Leaf to Globe. Boston: Academic

Li G Z, Wang M and Zeng H J. 2006. An Introduction to Support Vector Machines. Beijing: Publishing House of Electronics Industry

Li X W, Wang J F, Wang J D and Liu Q H. 2001. Multi-Angle and Thermal Inferred Remote Sensing. Beijing: Science Press

Li Y N, Zhao X Q and Cao G M. 2004. Analyses on climates and vegetation productivity background at Haibei Alpine Meadow ecosystem research station. *Plateau Meteorology*, **23**(4): 558—567

Lin H D. 2008. A simple Introduction to LIBSVM. Internet: www.csie.ntu.edu.tw/~cjlin/libsvm

Pierce L L and Running S W. 1995. The effects of aggregating sub-grid land surface variation on large-scale estimates of net primary production. *Landscape Ecology*, **10**(4): 239—253

Nemani R R, Keeling C D and Hashimoto H. 2003. Climate-driven increases in global terrestrial net primary production from 1982 to 1999. *Science*, **300**(5625): 1560—1563

Pu J Y, Li Y N, Zhao L and Yang S H. 2005. Seasonal changes of Kobresia Humilis Meadow biomass with climate factor. *Acta Agrestia Sinica*, **13**(3): 238—241

Raich J W, Rastetter E B and Melillo J M. 1991. Potential net primary productivity in South America: application of a global model. *Ecological Applications*, **1**(4): 399—429

Su L H, Li X W and Huang Y X. 2001. An review on scale in remote sensing. *Advance in Earth Sciences*, **16**(4): 544—548

Suleiman A and Crago R. 2004. Hourly and daytime evapotranspiration from grassland using radiometric surface temperatures. *Agronomy Journal*, **96**: 384—390

Tong Q X, Zhang B and Zheng L F. 2006. Hyperspectral Remote Sensing: Theory, Technology and Application. Beijing: Higher Education Press

Turner D P, Dodson R D and Marks D. 1996. Comparison of alternative spatial resolutions in the application of a spatially distributed biogeochemical model over complex terrain. *Ecological Modeling*, **90**(1): 53—67

Venturini V, Bisht G and Islam S. 2004. Comparison of evaporative fractions estimated from AVHRR and MODIS sensors over South Florida. *Remote Sensing of Environment*, **93**(1—2): 77—86

Wang L W, Wei Y X and Niu Z. 2008. Spatial and temporal variations of vegetation in Qinghai Province based on satellite data. *Journal of Geographical Sciences*, **18**(1): 73—84

Wang X Z, Tao B Z and Qiu W N. 2006. Advanced Surveying Adjustment. Beijing: Surveying and Mapping Press

Wen P Y, Shu G L and Guang S Z. 2007. Deriving a light use efficiency model from eddy covariance flux data for predicting daily gross primary production across biomes. *Agricultural and Forest Meteorology*, **143**(3—4): 189—207

Yang F T, Wang Q J and Shi S H. 1987. The allocation of the biomass and energy in Kobresia Humilis Meadow, Haibei district, Qinghai province. *Acta Phytoecologica Et Geobotanica Sinica*, **11**(2): 106—112

Zhang Y Q, Liu C M and Yu Q. 2004. Energy fluxes and the Priestley-Taylor parameter over winter wheat and maize in the North China Plain. *Hydrol. Process*, **18**: 2235—2246

Zhu W Q, Pan Y Z and He H. 2006. Simulation of maximum light use efficiency for main vegetation in China. *Chinese Science Bulletin*, **51**(6): 700—706

附中文参考文献

陈良富, 高彦华, 李丽, 柳钦火, 顾行发. 2007. 基于 MODIS 晴空数据的森林日净第一性生产力估算. 中国科学 D 辑: 地球科学, **37**(11): 1515—1521

李国正, 王猛, 曾华军. 2006. 支持向量机导论. 北京: 电子工业出版社

李小文, 汪骏发, 王锦地, 柳钦火. 2001. 多角度与热红外对地遥感. 北京: 科学出版社

李英年, 赵新全, 曹广民. 2004. 海北高寒草甸生态系统定位站气候、植被生产力背景的分析. 高原气象, **23**(4): 558—567

林弘德. 2008. LIBSVM 的简易入门. www.csie.ntu.edu.tw/～cjlin/libsvm

蒲继延, 李英年, 赵亮, 杨时海. 2005. 矮嵩草草甸生物量季节动态及其与气候因子的关系. 草地学报, **13**(3): 238—241

苏理宏, 李小文, 黄裕霞. 2001. 遥感尺度问题研究进展. 地球科学进展, **16**(4): 544—548

童庆禧, 张兵, 郑兰芬. 2006. 高光谱遥感——原理、技术与应用. 北京: 高等教育出版社

王新洲, 陶本藻, 邱卫宁. 2006. 高等测量平差. 北京: 测绘出版社

杨福囤, 王启基, 史顺海. 1987. 青海海北地区矮嵩草草甸生物量和能量的分配. 植物生态学与地植物学学报, **11**(2): 106—112

朱文泉, 潘耀忠, 何浩. 2006. 中国典型植被最大光能利用率模拟. 科学通报, **51**(6): 700—706