



HAL
open science

An Improved Computational Geometry Method for Obtaining Accurate Remotely Sensed Products via Convex Hulls With Dynamic Weights: A Case Study With Leaf Area Index

Hong Chen, Hua Wu, Zhao-Liang Li, Jienan Tu

► **To cite this version:**

Hong Chen, Hua Wu, Zhao-Liang Li, Jienan Tu. An Improved Computational Geometry Method for Obtaining Accurate Remotely Sensed Products via Convex Hulls With Dynamic Weights: A Case Study With Leaf Area Index. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2019, 12 (7), pp.2308-2319. 10.1109/JSTARS.2019.2906053 . hal-02377732

HAL Id: hal-02377732

<https://hal.science/hal-02377732v1>

Submitted on 25 Nov 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

An Improved Computational Geometry Method for Obtaining Accurate Remotely Sensed Products via Convex Hulls with Dynamic Weights: A Case Study with Leaf Area Index

Hong Chen, Hua Wu, Zhao-Liang Li and Jienan Tu

Abstract—Most retrieval functions used in remote sensing assume that the land surface is homogeneous. When those functions are used at coarse spatial resolution for heterogeneous surfaces, scale effects might appear. This paper tries to develop an improved computational geometry method (ICGM) upscaling model that takes into consideration the actual distribution of surface measurements by using dynamic weights for the upper and lower envelopes of a convex hull. By aggregating to a series of simulated data at coarse spatial resolution, the weight coefficients can be determined via a least square method. To evaluate the proposed upscaling model, the leaf area index (LAI) is used as an example. The results for three sites with different degrees of heterogeneity show that the ICGM upscaling model can effectively correct for the scale effects of the LAI and in most cases achieve an accuracy that is comparable to that of traditional upscaling models. The relative error of the estimated LAI for the selected sites decreases from 3.35%, 11.01%, and 19.62% to an average of 0.28%, 1.48%, and 5.16%, respectively, at kilometer scale. A determination of whether retrieval functions are continuous or derivable is no longer required. Furthermore, there is no need to rely upon synchronous high spatial resolution data. Because the weight coefficients vary little at different scales, those coefficients are thought to be insensitive to different scales and can be taken as constants for a given study site. This study indicates that the proposed method is promising and feasible even for a heterogeneous landscape.

Manuscript received September 15, 2018; this work was supported partially by the National Natural Science Foundation of China under Grant 41531174 and 41871267 and partially by the National Key R&D Program of China under Grant 2018YFB0504800 and 2016YFC0502506 (*Corresponding author: Hua Wu.*)

H. Chen is with China Aero Geophysical Survey & Remote Sensing Center for Nature Resources, Beijing, 100083, China (e-mail: chch1223@126.com).

H. Wu is with the State Key Laboratory of Resources and Environment Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China, with the University of Chinese Academy of Sciences, Beijing 100049, China, and also

Index Terms—Upscaling, leaf area index, Taylor Series Expansion, Computational Geometry

I. INTRODUCTION

Advances in remote sensing technology, which can quickly retrieve a wide range of real-time land surface spatial information, have provided a powerful way to generate regional and global products for resource surveys, environmental monitoring and disaster prediction. However, most of the retrieval functions in remote sensing assume that the land surface is homogeneous [1, 2]. If those functions are directly applied to generate products at coarse spatial resolution in areas where the land surface is heterogeneous, scale effects may be produced [1, 3]. Some researchers have indicated that using data and models at inappropriate scales without thorough consideration could lead to meaningless conclusions [4, 5]. The existence of scale effects has hampered improvements in the accuracy of retrieval and the development of remote-sensing applications [1, 6]. Therefore, the elimination of scale effects has become an important issue in quantitative remote sensing.

Because of the large discrepancies in the scales of different data sources, the effects of various land surface parameters detected by remote sensing are complex and diverse. Taking the leaf area index (LAI) product as an example, scale effects may be as high as 50% of the exact LAI value in extreme conditions [7]. Thus, LAI retrieval at coarse spatial resolution is unlikely to produce high-accuracy products unless those scale effects are

with the Jiangsu Center for Collaborative Innovation in Geographical Information Resource Development and Application, Nanjing 210023, China (e-mail: wuhua@igsnr.ac.cn)

Z.-L. Li is with the Key Laboratory of Agricultural Remote Sensing, Ministry of Agriculture/Institute of Agricultural Resources and Regional Planning, Chinese Academy of Agricultural Sciences, Beijing 100081, China, and also with ICube, UdS, CNRS, Bld Sebastien Brant, BP10413, 67412, Illkirch, France (e-mail: lizl@unistra.fr).

J. N. Tu is with China Aero Geophysical Survey & Remote Sensing Center for Nature Resources, Beijing, 100083, China (e-mail: tujianan@outlook.com).

successfully corrected. Therefore, the correction of scale effects should be carefully addressed [8, 9].

Many previous studies have successfully proven that scale effects are caused not only by surface spatial heterogeneity but also by the nonlinearity of the retrieval function [1, 2, 9, 10]. The nonuniformity and nonlinearity of land surface systems make it very difficult to solve the problem of scale. Many researchers have tried to use mathematical methods to quantitatively analyze scale effects and establish a semiempirical or physical upscaling method. A series of upscaling methods have been proposed, including the computational geometry method (CGM), the Taylor series expansion method (TSEM) and the contexture-based method (CM). Based on the theory of convex hulls in computational geometry, the CGM-based upscaling model assumes that retrieval values are evenly distributed between the upper and lower envelopes of the retrieval function to obtain accurate retrieval values at coarse spatial resolutions [2]. Regardless of whether retrieval functions are continuous or derivable, this upscaling model is widely applicable. However, the retrieval values are not always equal to the average values of the upper and lower envelopes. The CGM-based upscaling model, which does not consider the actual distribution of surface parameters and assigns a constant weight coefficient for upper and lower envelopes, may be inappropriate and lead to a low accuracy of correction [1]. Thus, determining reasonable weight coefficients for the upper and lower envelopes is the key to improving the accuracy of this upscaling model. Hu and Islam proposed a TSEM-based upscaling model that takes the impact of density changes into account and uses textural parameters (variance and covariance) as inputs to quantitatively characterize the scale effects [11]. By extending the TSEM-based upscaling model, Garrigues *et al.* effectively corrected for the scale effects of the multivariate retrieval function [12]. Consequently, the TSEM-based upscaling model has been applied in several studies to various surface parameters and has achieved satisfactory results [13-16]. However, it is rather difficult to obtain the textural parameters required by the TSEM-based upscaling model. Generally, synchronous high spatial resolution data are necessary to estimate textural parameters. In addition, the TSEM-based upscaling model can be used only under the condition that the retrieval function is continuous and derivable. When high spatial resolution data are not available or the retrieval function does not meet those requirements, the TSEM-based upscaling model is restricted in its use. Apart from that, the TSEM-based upscaling model does not consider the influence of pixels with different components on scale effects, which will reduce the accuracy of upscaling. To overcome this limitation of the TSEM-based upscaling model, Chen proposed a CM-based upscaling model in which the contextual parameters represented by the fractions of the components were adopted to quantify the scale effects [7]. Simic *et al.* then used such subpixel information to correct for the scale effects of the net primary productivity and obtained satisfactory results [17]. However, the physical mechanism of this method is not yet clear. Studies have tried to use both textural and contextual parameters to successfully remove the

scale effects of surfaces with contrasting cover types [6, 18]. However, these attempts still rely heavily on synchronous high spatial resolution data, which will restrict the corresponding applications. In addition to the upscaling models mentioned above, there are also a few studies on the upscaling of remote sensing products based on fractal theory [19-23]. However, the upscaling model based on fractal theory is not applicable when upscaling is applied beyond the domain of the similarity scale [9]. In recent years, some studies have focused on the correction of scale effects by applying mathematical tools, such as wavelet transforms; these approaches have weakened scale effects to a certain degree (For example, the relative error caused by scale effects can be reduced by half) but are still limited by the need for synchronous high spatial resolution data [24-26]. Although the accuracy of LAI products at coarse spatial resolution has been improved to a certain extent with these models, the upscaling models still need further improvements to expand their universality [27].

The objective of this paper is to improve the traditional CGM-based upscaling model and to use LAI upscaling as an example to show how retrieval accuracy at coarse spatial resolution can be improved by applying appropriate weight coefficients. The next sections will be organized as follows. In Section 2, the LAI upscaling problem and aggregation analysis using convexity will be described. A solution to the spatial upscaling problem of that LAI using convex hulls will be proposed. Section 3 describes the data used to evaluate the applicability of the improved CGM-based upscaling model. The results of the evaluation and the corresponding discussions will be provided in Section 4. The scale effects of three contrasting sites will be demonstrated, and the proposed method will be evaluated with nested multiscale aggregated data. The conclusions will be given in Section 5.

II. METHODOLOGY

A. LAI scale effects in retrieval

LAI products can be spatially estimated from remotely sensed data by using a retrieval function derived either from a radiative transfer model or from a calibrated semiempirical relationship [12]. The retrieval process of the LAI can be expressed abstractly as follows:

$$\text{LAI} = f(x) \quad (1)$$

where f is the retrieval function and x represents the surface reflectance or vegetation index.

Here, the cause of the scale effects is illustrated in an example of a semiempirical retrieval function between the normalized difference vegetation index (NDVI) and the LAI. The NDVI is derived from the reflectances of red and near infrared bands as [28]:

$$\text{NDVI} = \frac{\rho_{nir} - \rho_{red}}{\rho_{nir} + \rho_{red}} \quad (2)$$

where ρ_{nir} and ρ_{red} represent the near infrared and red reflectances, respectively. The nonlinear relationship between the NDVI and the reflectances of the red and near-infrared bands also contributes to the scale effects of LAI products [21].

However, for simplicity, only the scale effects associated with the nonlinear relationship between the NDVI and the LAI are quantified in this paper. The scale effects associated with the relationship between the LAI and the reflectances of the red and near-infrared bands can be quantified in a similar way.

As shown in Fig. 1, there are two distinct upscaling approaches that are used to obtain the LAI products at coarse spatial resolution. One is the spatial aggregation of LAI products at high spatial resolution, and the other is direct retrieval at coarse spatial resolution. In the first process, the exact value of the LAI (LAI_{EXA}) is obtained by first calculating the NDVI value at high spatial resolution and then by calculating the average (path I in Fig. 1):

$$LAI_{EXA} = \frac{1}{N} \sum_{i=1}^N LAI_i = \frac{1}{N} \sum_{i=1}^N f(NDVI_i) \quad (3)$$

where N is the number of pixels at high spatial resolution within the corresponding coarse spatial resolution pixel; LAI_i and $NDVI_i$ are the corresponding values for the i th high spatial resolution pixel; and f represents the LAI retrieval function at high spatial resolution. The pixel at high spatial resolution is considered homogenous without scale effects; therefore, this LAI value is regarded as the actual LAI value for high spatial resolution. Correspondingly, the aggregated LAI value is considered to be the exact value of the LAI at coarse spatial resolution.

In the second process with the coarse spatial resolution measurement, the approximated value of the LAI (LAI_{APP}) at coarse spatial resolution can be estimated as:

$$LAI_{APP} = f(NDVI_m) \quad (4)$$

where $NDVI_m$ is the NDVI value of the pixel at coarse spatial resolution. If the scale effect of the NDVI is ignored, the $NDVI_m$ of the coarse spatial resolution could be regarded as the average value of NDVI at high spatial resolution.

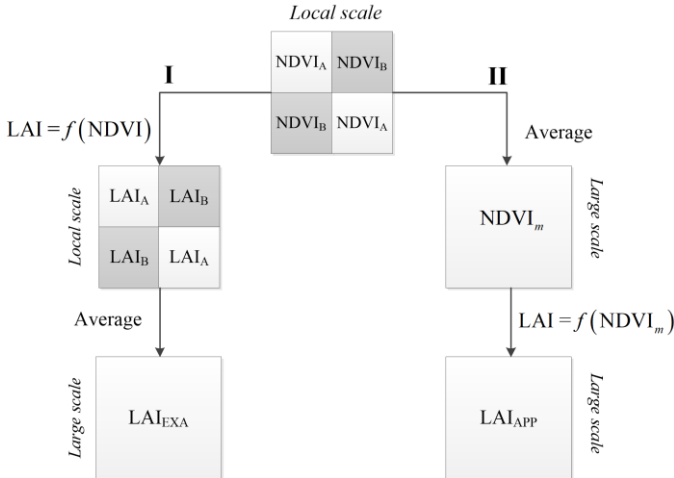


Fig. 1 Sketch map of the effect of spatial resolution on the nonlinear estimation of the LAI. Paths I and II represent two different upscaling processes. Path I (left) represents the application of the retrieval function f at high spatial resolution, in which the results are aggregated to obtain LAI_{EXA} ; Path II (right) entails aggregating the value of the NDVI before it is applied in the retrieval model f to obtain LAI_{APP} .

Since the retrieval function f is established for homogeneous pixels, it is suitable for high spatial resolution. When the

retrieval function f is used at coarse spatial resolution, scale effects will occur. The difference (bias) between the exact value of LAI_{EXA} and the approximated value of LAI_{APP} at coarse spatial resolution is regarded as the so-called scale effects and can be expressed as follows:

$$bias = LAI_{EXA} - LAI_{APP} \quad (5)$$

To show the scale effects of the LAI clearly, an example of a mixture scenario is demonstrated in Fig. 2. In this scenario, two components (A and B) with the same cover type but different densities are included. Here, the nonlinear relationship between the NDVI and the LAI is represented by the retrieval function f , which is generated at high spatial resolution and in which each component is assumed to be homogeneous.

The LAI_{EXA} and LAI_{APP} are the values of the LAI at coarse spatial resolution; these values are obtained by two different upscaling processes that use Eqs. (3) and (4), respectively.

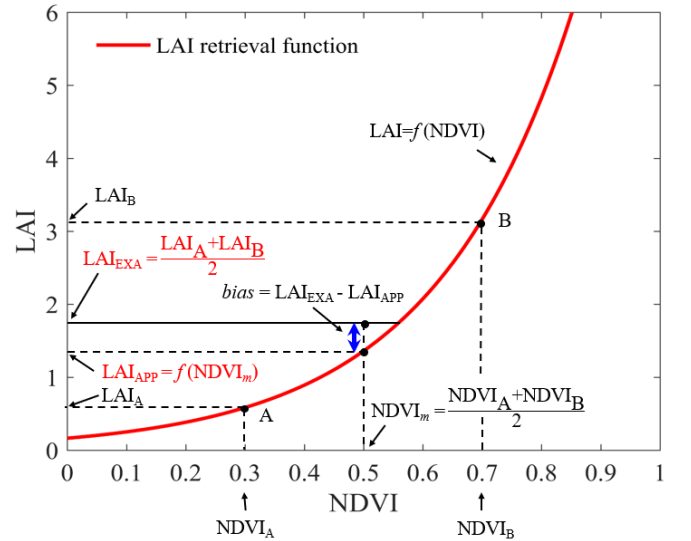


Fig. 2. Illustration of the scale effects of the LAI caused by function nonlinearity and spatial heterogeneity at coarse spatial resolution. f is a semiempirical retrieval function at high spatial resolution that relates the LAI and the NDVI. $NDVI_A$ and $NDVI_B$ are the NDVI values of components A and B, while LAI_A and LAI_B are their corresponding LAI values, respectively. $NDVI_m$ is the corresponding NDVI value of the mixed pixel at coarse spatial resolution and is defined as the average of $NDVI_A$ and $NDVI_B$. The LAI_{EXA} is the exact value of the LAI at coarse spatial resolution and is obtained by first applying f at high spatial resolution and then averaging the LAI_A and LAI_B at coarse spatial resolution. The LAI_{APP} is the approximated value of the LAI for the pixel at coarse spatial resolution and is derived from the retrieval result that directly applies the retrieval function f to $NDVI_m$. The difference between LAI_{EXA} and LAI_{APP} is regarded as the scale effect at coarse spatial resolution.

As shown in Fig. 2, the direct application of the retrieval function f at coarse spatial resolution will underestimate the exact LAI value because of the concave characteristics of the retrieval function [2]. The difference between the LAI_{EXA} and LAI_{APP} shown in Fig. 2 is considered to be the scale effect. It is crucial to minimize the influence of scale effects on the LAI product at coarse spatial resolution.

B. Aggregation analysis using convexity

According to the convex hull of computational geometry theory, the set of all possible values of LAI_{EXA} will fall into an interval when $NDVI_m \in D$ [2]:

$$LAI_{EXA} \in [f_{\vee}(NDVI_m), f^{\wedge}(NDVI_m)] \quad (6)$$

where $NDVI_m$ is the value of the NDVI at coarse spatial resolution, and D is the domain of the NDVI. The bounding functions f^{\wedge} and f_{\vee} are, respectively, the upper and lower envelopes of the retrieval function f . The determination of the upper and lower envelopes depends on the type of retrieval function and the scope of the space domain D .

To clearly demonstrate the difference between upper and lower envelopes in different conditions, three types of

landcover (bare soil, vegetation A and vegetation B) with different retrieval functions are taken as examples. As shown in Fig. 3 (a), the retrieval functions of bare soil, vegetation A and vegetation B are presented in green, blue and red, respectively. The NDVI values of the three different coverage types are as follows: bare soil (D_1 : $NDVI < 0.2$), vegetation A (D_2 : $0.2 < NDVI < 0.5$), and vegetation B (D_3 : $0.5 < NDVI < 0.8$). The corresponding upper and lower envelopes for the different conditions are shown in Fig. 3(b)-3(d). These graphs are only used to present the corresponding upper and lower envelopes of the function of the LAI and the NDVI under certain conditions. In fact, the retrieval functions for various land cover types may be more complex when the scope of D corresponding to each landcover type overlaps, as shown in Fig. 3.

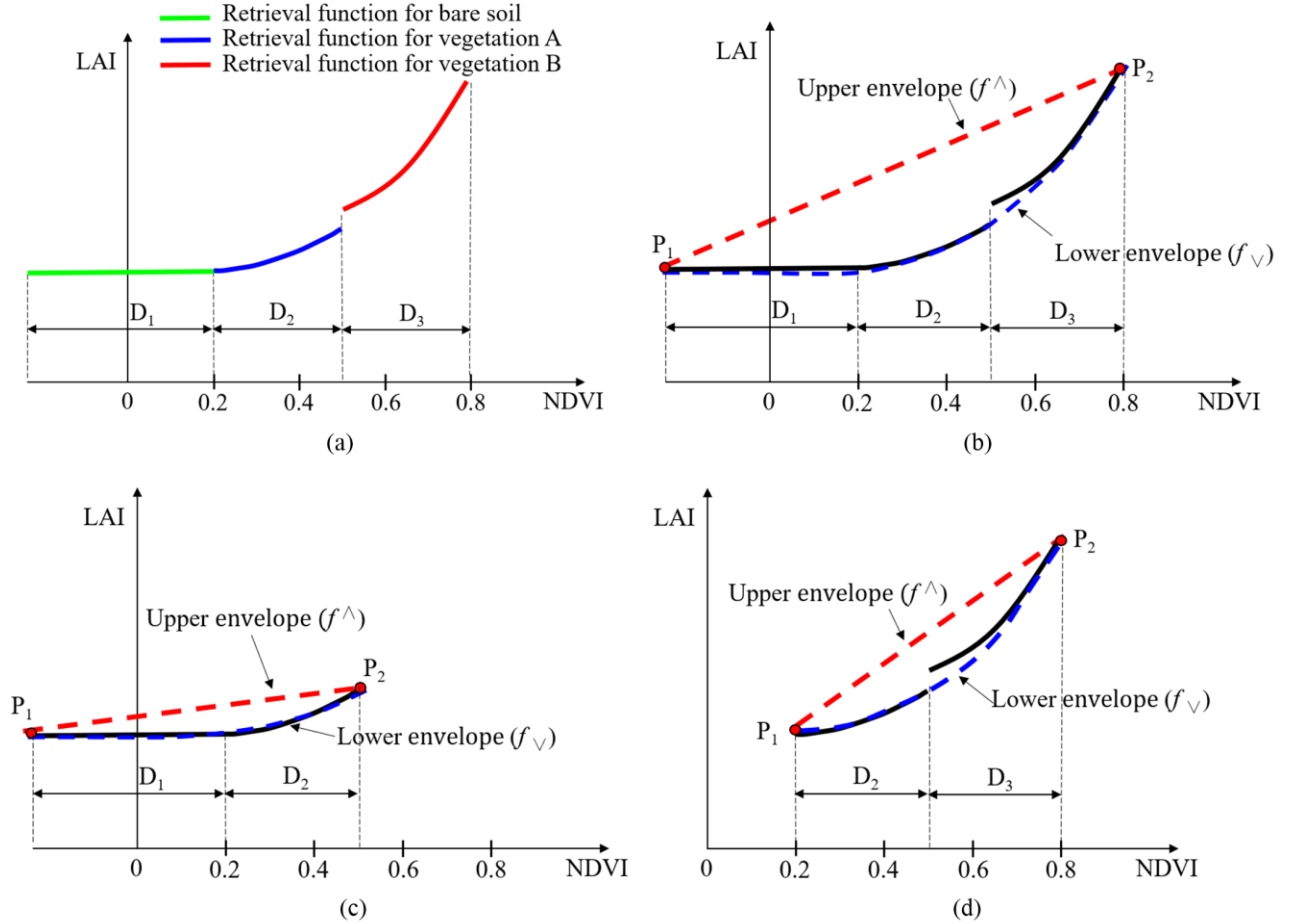


Fig. 3. The graphs of upper and lower envelopes for three different conditions. The bounding functions f^{\wedge} and f_{\vee} are the upper and lower envelopes of the retrieval function f , respectively. (a) Retrieval functions of three different types of land cover. The green, blue and red curves represent the retrieval functions for bare soil, vegetation A and vegetation B, respectively. D_1 , D_2 and D_3 are the respective NDVI domains of each landcover type; (b) the upper and lower envelopes when the three types of landcover are mixed; (c) the upper and lower envelopes when the land is covered by bare soil and vegetation A; (d) the upper and lower envelopes when the land is covered by vegetation A and vegetation B. P_1 and P_2 are the leftmost and rightmost endpoints of the retrieval curve, respectively.

As shown in Fig. 3, the space domain $D \in [D_1 \cup D_2 \cup D_3]$. The upper and lower envelopes f^{\wedge} and f_{\vee} are usually determined from the corresponding retrieval function f with the space domain of the distribution of the NDVI. The corresponding upper and lower envelopes change with different conditions. When $NDVI \in D$, the land cover includes the three different types; where the retrieval functions are discontinuous,

the upper envelope can be treated as a straight line connecting the leftmost and rightmost endpoints, P_1 and P_2 , of the curve. The curves are wrapped by the lower envelope on the bottom. When $NDVI \in [D_1 \cup D_2]$, the upper envelope can be expressed as a linear function between P_1 and P_2 . For the curve to be continuous under this condition, the lower envelope must coincide with the retrieval function.

When $NDVI \in [D_2 \cup D_3]$, the upper envelope can also be expressed as a linear function, and the discontinuous curves are wrapped by the lower envelope from the bottom. As shown in Fig. 3(b)-(d), the corresponding upper and lower envelope functions can be established regardless of whether the retrieval function is continuous or discontinuous.

It is worth noting that the described upper and lower envelopes are usually determined by the spatial domain. When the surface is homogenous and where the NDVI values of all pixels are equal, the difference between LAI_{EXA} and LAI_{APP} is zero. In the case of a linear function, both the upper and lower envelopes are coincident with the retrieval function, $f \equiv f_v \equiv f^{\wedge}$, and the bias between LAI_{EXA} and LAI_{APP} will also be zero. Generally, there are no scale effects when the surface is homogenous or the retrieval function is linear. The difference between the envelopes f^{\wedge} and f_v is an estimate of the maximum bias between LAI_{EXA} and LAI_{APP} . Although the CGM-based upscaling model cannot quantitatively express scale effects as well as the TSEM-based upscaling model can, the CGM-based upscaling model can still reflect the combined scale effects of surface heterogeneity and the nonlinearity of the retrieval function.

C. Solution to the spatial upscaling problem of the LAI using convex hulls

Without obtaining the spatial distribution of the surface reflectance of the pixels, Raffy assumed that the LAI values at coarse spatial resolution are evenly distributed between the upper and lower envelopes [2]. By taking the average value of the upper and lower envelopes, the scale effects were reduced to a certain extent. The corresponding traditional CGM-based upscaling model can be expressed as follows:

$$LAI_{cor}^{CGM} = \frac{1}{2} (f_v(NDVI_m) + f^{\wedge}(NDVI_m)) \quad (7)$$

where LAI_{cor}^{CGM} is the scale-effect-corrected LAI value that is obtained using the traditional CGM-based upscaling model. However, the results show that the scale effects correction that is obtained with a constant coefficient of 1/2 is not satisfactory; this will be demonstrated in detail in the following sections. Under certain conditions, the accuracy of the retrieved LAI after correction with this upscaling model is not as good as it was before. This is likely caused by the assumption of uniform distribution, which is not consistent with the actual distribution of the NDVI [6]. Consequently, the traditional GCM-based upscaling model that uses a constant coefficient should be improved.

In this paper, we assume that the values of the weight coefficients of the upper and lower envelopes will change dynamically and might be related to the nonlinearity of the retrieval function and the heterogeneity of the land surface. An improved computational geometry method (ICGM)-based upscaling model with dynamic weights for the upper and lower envelopes of the convex hull is proposed as follows:

$$LAI_{cor}^{ICGM} = a \times f_v(NDVI_m) + b \times f^{\wedge}(NDVI_m) \quad (8)$$

where LAI_{cor}^{ICGM} is the scale-effect-corrected LAI value using the ICGM-based upscaling model; a and b are the weight coefficients of the upper and lower envelopes, respectively, $a, b \in [0, 1]$ and $a + b = 1$. Therefore, the key issue associated with the ICGM-based upscaling model is how to determine the weight coefficients. As described in the analysis in the following section, the weight coefficients are thought to be insensitive to changes in scale. Therefore, autocorrelation and self-similarity characteristics of the data can be used to determine the weight coefficients. These weight coefficients can be estimated by applying a least square method to a nested multiscale data set, which can be generated either from nonsynchronous historical high spatial resolution data or from coarse spatial resolution data by using a simple aggregation strategy.

III. STUDY AREAS AND DATA

To evaluate the performance of the ICGM-based upscaling model, three sites (Zhangbei, Haouz and Fundulea) with different degrees of surface heterogeneity were selected. The Zhangbei site, located in northern China (41°16'44"N, 114°41'16"E), is composed of plateaus (1400 m elevation) with pastures (for sheep and cows) and a few crops near villages. The Haouz site, located in the middle of Morocco (31°39'33"N, 7°36'01"W), is mainly covered by cropland such as wheat, fallow, barley and other crops. The Fundulea site is the most heterogeneous of the three sites and is located in the south of Romania (44°24'21"N, 26°35'06"E); it consists mainly of large crop fields and some forests.

The ground sampling data obtained on 8-10 August 2002 for Zhangbei, 10-14 March 2003 for Haouz, and 9-10 May 2001 for Fundulea were all derived from the VALERI (VALidation of Land European Remote sensing Instruments) dataset. Accordingly, to establish a retrieval function and further evaluate the proposed model, three images of 2000 × 2000 pixels with high spatial resolution (30 m) obtained close to the sampling time (17 August 2002 for Zhangbei, 15 March 2003 for Haouz, 30 April 2001 for Fundulea) were collected from the Landsat 7 Enhanced Thematic Mapper Plus (ETM+) satellite. Because the image of Fundulea was obscured by clouds, this image was only used to estimate the model coefficients of the retrieval function. An image of Fundulea on 13 March 2001 was selected for further evaluation.

The acquisition images from the ETM+ were surface reflectance products directly downloaded from the website of the USGS (<http://earthexplorer.usgs.gov/>). The images had undergone level 2 processing (radiometric correction, geometric correction and atmospheric correction) with the Universal Transverse Mercator (UTM) projection. The distributions of the elementary sampling units (ESUs) around the three sites are shown in Fig. 4.

A regression analysis is used to determine the relationship between the NDVI and the LAI from the ground sampling data. At the same time, we only consider the influence of the spatial heterogeneity of the NDVI and the retrieval function nonlinearity on scale effects; we ignore the scale effects of the

NDVI itself. To avoid the impact of registration errors and ensure the accuracy of regression, the sampling points obscured by clouds are removed. Any points with an NDVI standard deviation of approximately 3×3 pixels greater than 0.05 are removed from the sampling data. As shown in Fig. 4, the removed sampling points are marked in black and are not considered in the regression.

After comparing the accuracy of multiple types of retrieval functions, an exponential empirical function representing the relationship between the NDVI and the LAI is chosen, mainly

because it has a higher fitting accuracy than other types of functions. The exponential LAI-NDVI function has been adopted in other studies as well [29, 30] and can be expressed as follows:

$$LAI = m \times e^{n \times NDVI} \quad (9)$$

where m and n are coefficients of the retrieval function and depend on the characteristics of the sites.

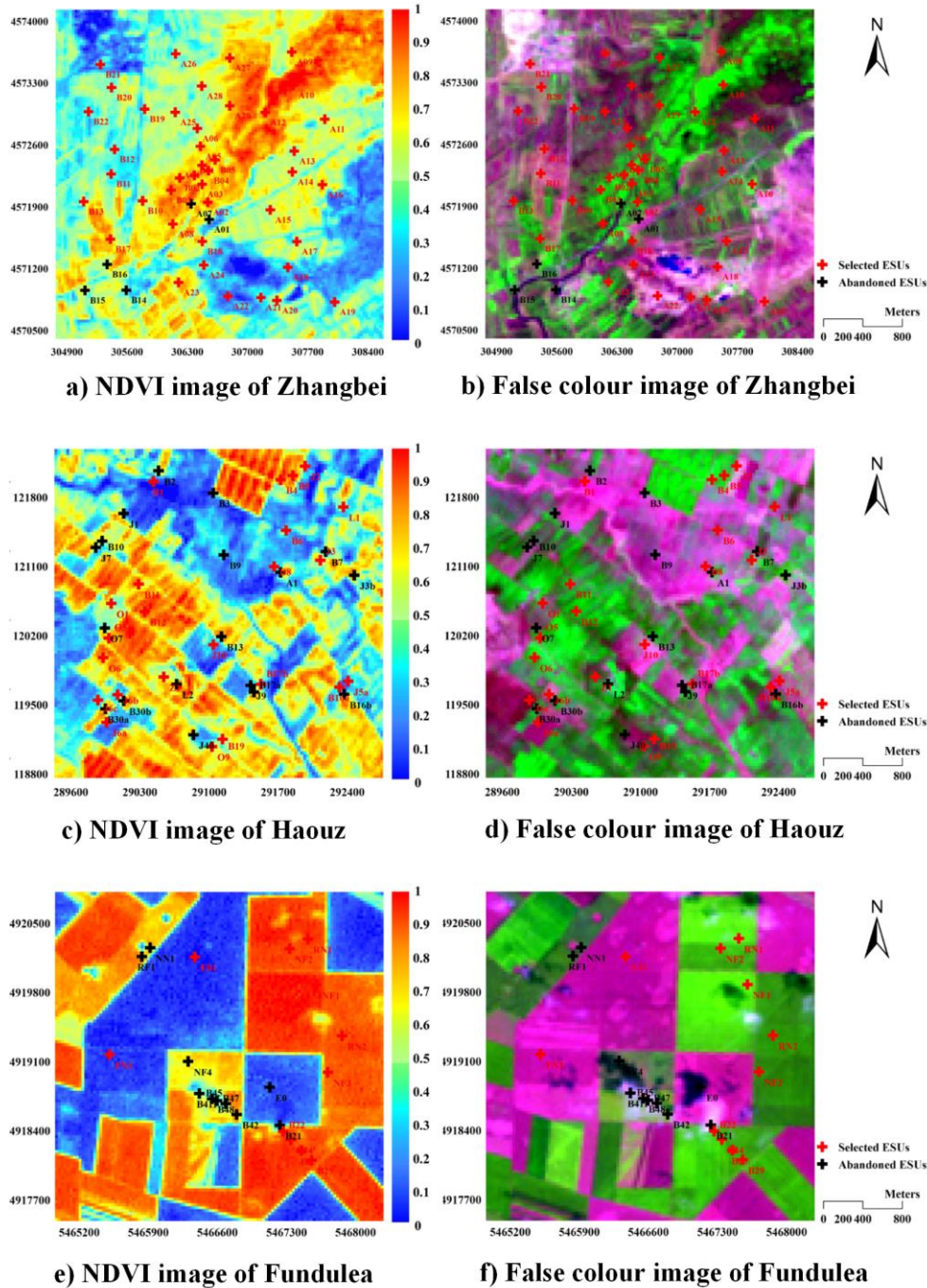


Fig. 4. The distribution of ESUs in three study sites. (a), (c) and (e) are NDVI images corresponding to Zhangbei, Haouz and Fundulea, respectively. (b), (d) and (f) are false color images composed of bands 5, 4 and 3 and correspond to Zhangbei, Haouz and Fundulea, respectively. The red and black points represent the selected and abandoned sampling points in the regression analysis, respectively.

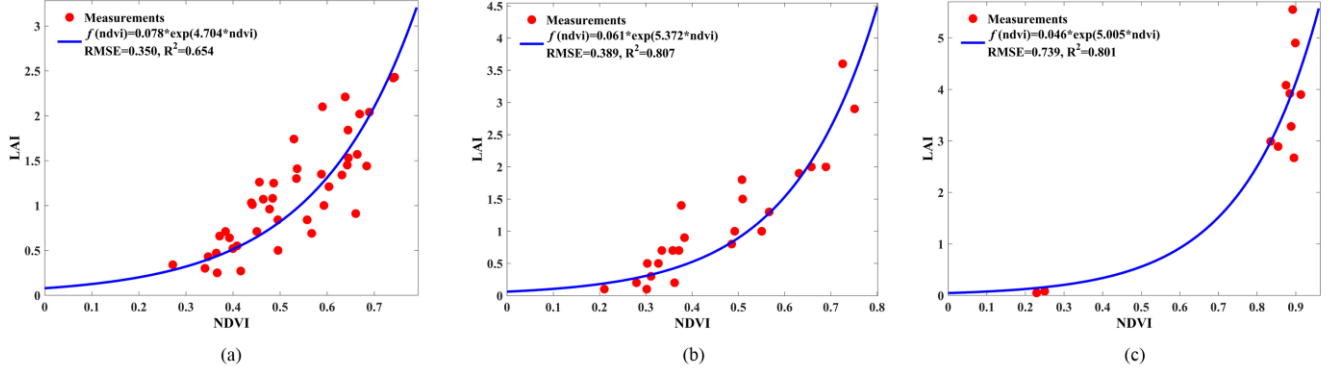


Fig. 5. The retrieval functions of the three selected study sites. (a), (b) and (c) correspond to Zhangbei, Haouz and Fundulea, respectively. The red points represent the ground sampling data, and the blue lines represent the curves of the retrieval functions.

The accuracies of the retrieval functions will not be considered in this paper because the function is just taken as a case study by which to analyze the scale effects of the LAI caused by the nonlinearity of the retrieval function and by spatial heterogeneity. The NDVI was calculated based on the selected ETM+ image. The retrieval functions are generated with the ground-measured LAI data and corresponding NDVI values for each site. As presented in Fig. 5, the root mean square errors (RMSEs) are 0.350, 0.389 and 0.739, corresponding to Zhangbei, Haouz and Fundulea, respectively, and the determination coefficient (R^2) values are 0.654, 0.807 and 0.801, also corresponding to these three sites.

As shown in Fig. 5, among the three sites, the nonlinearity of the retrieval function in Fundulea is the lowest, and the nonlinearity of that at the Haouz site is the highest.

IV. EXPERIMENT RESULTS AND DISCUSSION

A. Weights for the upper and lower envelopes at different scales

As described in Section 2, a method for obtaining the weight coefficients in Eq. (8) is key for the proposed upscaling model.

TABLE 1. THE VALUE OF THE WEIGHT COEFFICIENTS OF THE UPSCALING MODEL AT DIFFERENT SCALES

Study Site	Weight Coefficients	Weight Coefficient Values at Different Aggregation Scales						Mean value
		2×2	5×5	10×10	20×20	25×25	40×40	
Zhangbei	a	0.24	0.24	0.23	0.21	0.22	0.21	0.23
	b	0.76	0.76	0.77	0.79	0.78	0.79	0.77
Haouz	a	0.23	0.21	0.22	0.22	0.22	0.22	0.22
	b	0.77	0.79	0.78	0.78	0.78	0.78	0.78
Fundulea	a	0.23	0.19	0.18	0.17	0.18	0.18	0.19
	b	0.77	0.81	0.82	0.83	0.82	0.82	0.81

As shown in Table 1, the weight coefficients a are all smaller than the values of b at the six aggregation scales, indicating that the distributions of the exact retrieval values are all closer to the lower envelope than they are to the upper envelope, which is in

accordance with the characteristics of the concave function. As [1], statistical autocorrelation and self-similar characteristics could be used to determine the weights.

First, the images of the three sites at high spatial resolution are aggregated to obtain nested simulated multiscale images at different coarse spatial resolutions to study how the weight coefficients change with the aggregation scales. The LAI and NDVI values at 30 m resolution are aggregated to six different scales with 2×2 , 5×5 , 10×10 , 20×20 , 25×25 , and 40×40 pixels. The aggregated LAI and NDVI data at 60 m, 150 m, 300 m, 600 m, 750 m, and 1200 m are then obtained. As shown in aggregation Path I in Eq. (3), the aggregated LAI values are taken as exact values on the left side of Eq. (8). The aggregated NDVI values are used as the inputs. The \hat{f} and \check{f} values on the right side of Eq. (8) are determined directly by the retrieval function f in Eq. (9). Using a least square method, the optimal weight coefficients at different scales are obtained from the simulated images. The values of the weight coefficients a and b for the six aggregation scales mentioned above are shown in Table 1.

As shown in Table 1, the values of the weight coefficients obtained from each site are not that different from each other, mainly because the nonlinearity of the three retrieval functions is

almost the same. Spatial heterogeneity increases with the aggregation of scale, and the values of the weight coefficients change little with increasing aggregation scale. We can conclude that the values of the weight coefficients are not sensitive to changes of scale but may be closely related to the nonlinearity of the retrieval functions and the inherent heterogeneity within the study areas.

Therefore, the weight coefficients of the ICGM-based upscaling model can be estimated with a least square method from a nested multiscale dataset that is either aggregated from the nonsynchronous historical high spatial resolution data or from the coarse spatial resolution data. Access to synchronous high spatial resolution data is no longer required in the proposed ICGM-based upscaling model, which is its greatest advantage over other models.

B. Correcting the results for scale effects

To compare the influence of different weight coefficients on the accuracy of the proposed upscaling model, different values for weight coefficient a are used for the three sites according to the maximum values (group A), the minimum values (group B) and the mean values (group C) in Table 1. Using Eq. (8), the ICGM-based upscaling models with site-dependent weight coefficients are established. Meanwhile, to evaluate the performances of the upscaling models, we make a comprehensive comparison among the ICGM-based, TSEM-based, and CGM-based upscaling models. LAI_{cor}^{TSE} and LAI_{cor}^{CGM} represent the corrected LAI in the TSEM-based and CGM-based upscaling models, and LAI_{cor}^{ICGM} represents the corrected LAI in the ICGM-based upscaling model, which is also subdivided into three classes according to the different groups of weight coefficients. $LAI_{cor}^{ICGM}(A)$, $LAI_{cor}^{ICGM}(B)$ and $LAI_{cor}^{ICGM}(C)$ represent the ICGM-based upscaling models with the maximum, minimum and mean values of coefficient a at six different aggregation scales, respectively.

Here, we use the relative error (RE) and the RMSE to evaluate the accuracies of the upscaling models.

$$RE = \frac{1}{N} \sum_{k=1}^N \frac{|LAI_{COR,k} - LAI_{EXA,k}|}{LAI_{COR,k}} \quad (10)$$

$$RMSE = \sqrt{\frac{\sum_{k=1}^N (LAI_{COR,k} - LAI_{EXA,k})^2}{N}} \quad (11)$$

where N represents the number of pixels at high spatial resolution within the corresponding coarse spatial resolution pixel; $LAI_{COR,k}$ represents the corrected LAI with the corresponding upscaling model; and $LAI_{EXA,k}$ represents the exact value of LAI at coarse spatial resolution.

As shown in Fig. 6, with increasing aggregation scale, the RE and RMSE of the scale effects increase substantially. Comparing the three different sites, the scale effects increase with increasing spatial heterogeneity.

At the Zhangbei and Fundulea sites, the correction accuracy of the ICGM-based upscaling model is slightly lower than that of the TSEM-based upscaling model, and the results of the ICGM-based upscaling model with different coefficients are different from each other. At the Zhangbei site, although the correction accuracy of the ICGM-based upscaling model is lower than that of the TSEM-based upscaling model at six aggregation scales, the ICGM-based upscaling model can still reduce the original RE of 3.35% to 0.25% at the 40×40 aggregation scale. However, at the Haouz site, the performances of the three groups of ICGM upscaling models with different coefficients are almost the same as that of the TSEM-based upscaling model. At the 20×20 aggregation scale and above, the corrected RE of the ICGM-based upscaling model is lower than that of the TSEM-based upscaling model, and the corrected RMSE is equal to or better than that of the TSEM-based upscaling model. The traditional CGM-based upscaling model with constant weight coefficients yields overcorrection results at all three sites.

To show the correction effect more intuitively, scatter plots of the correction results at 5×5 and 40×40 aggregation scales at the Haouz site are used as examples of the correction effect's moderate scale effects. The distribution of scatters at other sites and at different scales are very similar; however, these sites have different RE and RMSE values and are not shown here. The performances of the above upscaling models are compared systematically.

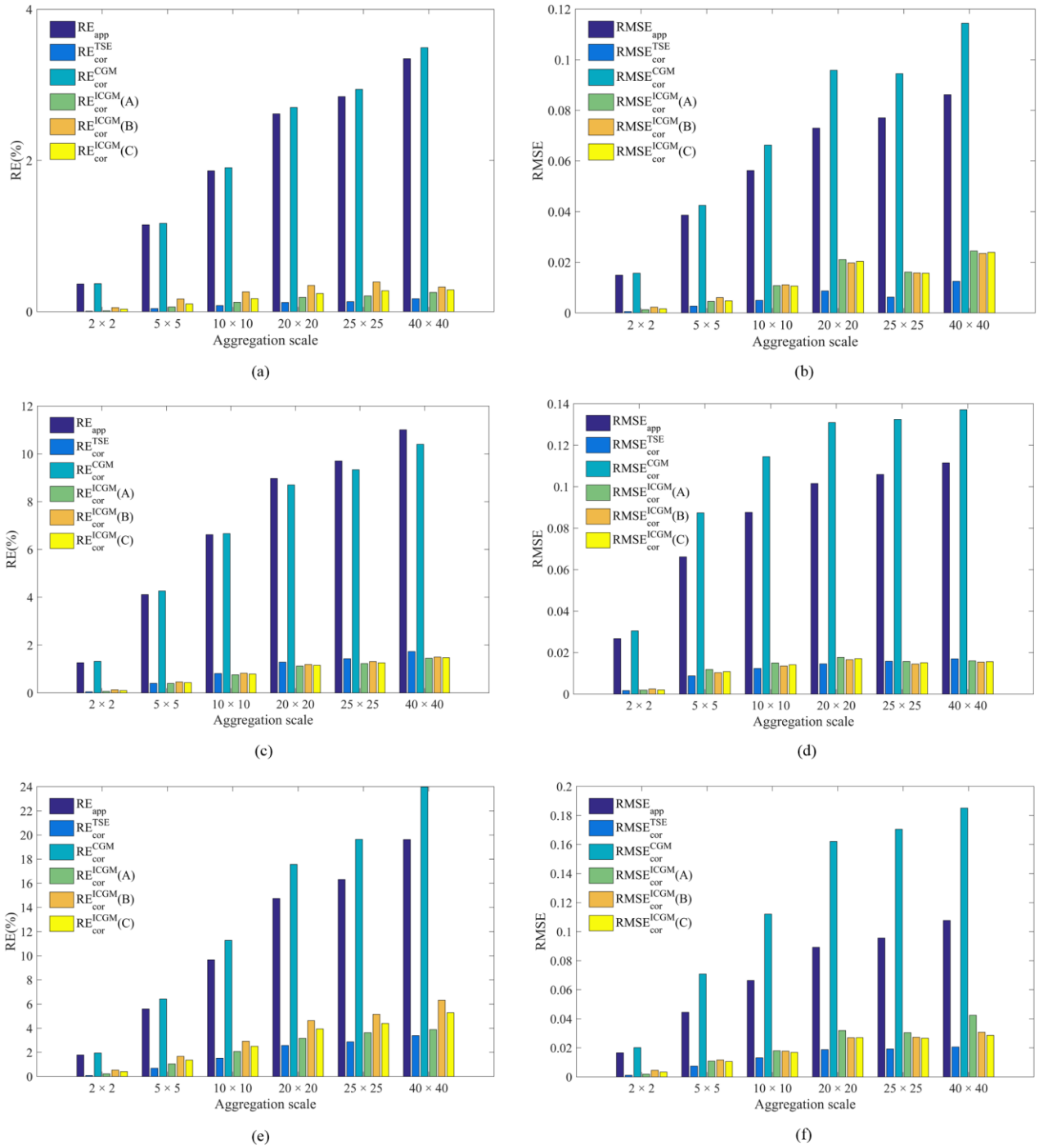


Fig. 6. Comparisons among the correction effects of upscaling models based on TSEM, CGM and ICGM with different coefficients at six different scales. (a) and (b) are RE and RMSE for Zhangbei, respectively; (c) and (d) are RE and RMSE for Haouz, respectively; (e) and (f) are RE and RMSE for Fundulea, respectively.

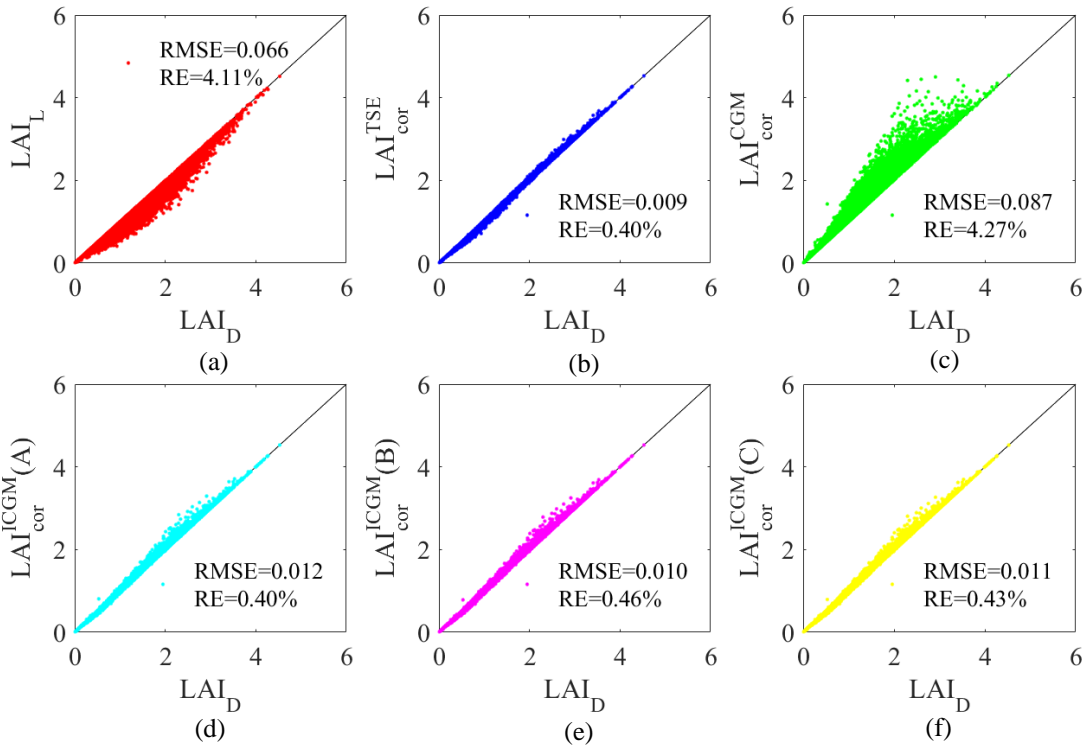


Fig. 7. Scatter plots showing the relationship between the exact LAI values and the estimated LAI values before and after correction at the 5×5 aggregation scale at the Haouz site. (a): The LAI values before correction; (b): Correction with a TSEM-based model; (c): Correction with a CGM-based model; (d): Correction with an ICGM-based model with maximum weight coefficients (A); (e): Correction with an ICGM-based model with minimum weight coefficients (B); (f): Correction with an ICGM-based model with mean weight coefficients (C).

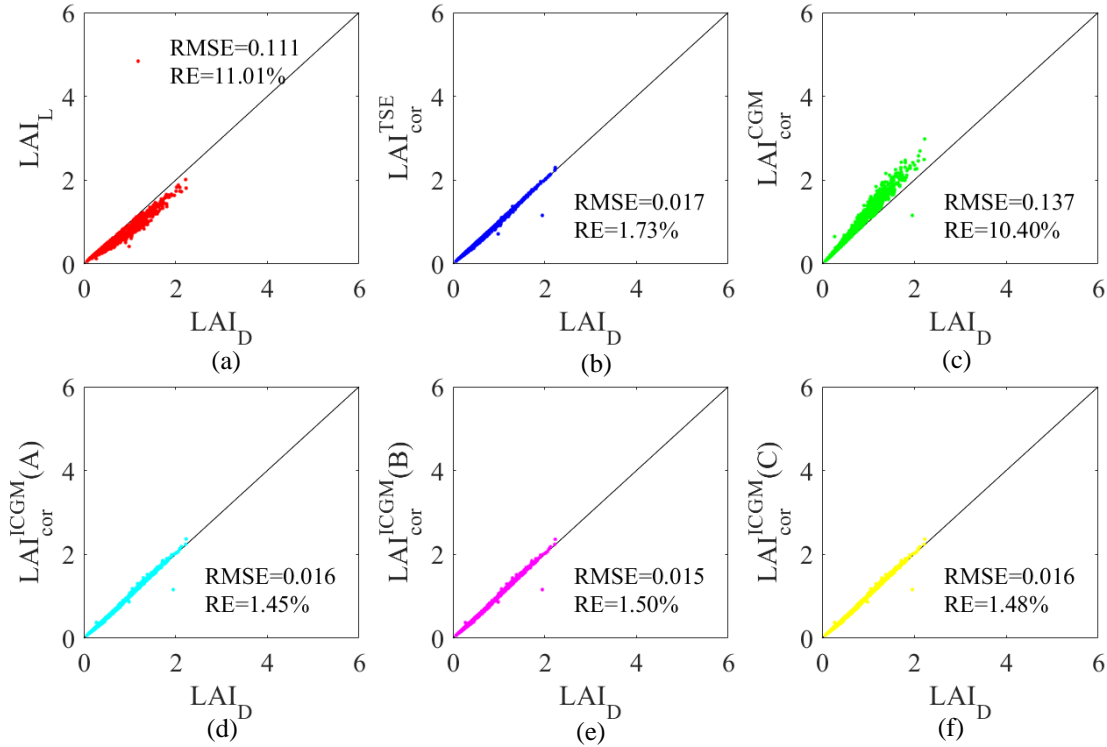


Fig. 8. Scatter plots showing the relationships between the exact LAI values and the estimated LAI values before and after correction at the 40×40 aggregation scale at the Haouz site. (a): LAI values before correction; (b): Correction with the TSEM-based model; (c): Correction with the CGM-based model; (d): Correction with the ICGM-based model with maximum weight coefficients (A); (e): Correction with the ICGM-based model with minimum weight coefficients (B); (f): Correction with the ICGM-based model with mean weight coefficients (C).

Fig. 7 and Fig. 8 show that the concave function of the exponential retrieval function will result in a LAI estimate that

is lower than the exact value, which is consistent with Raffy's research [2]. The results of the ICGM-based upscaling model with three group weight coefficients indicate that the scale

effects are significantly reduced. At the 5×5 aggregation scale, the RE are reduced from 4.11% to 0.4%, 0.46% and 0.43% within the three groups of ICGM-based upscaling models, and the RMSE are reduced from 0.066 to an average of 0.011, which is slightly lower than or equal to that of the TSEM-based upscaling model. At the 40×40 aggregation scale, all three ICGM-based upscaling models perform better than the TSEM-based upscaling model and yield a RE of approximately 1.5%. These comparisons reveal that the effect of the maximum, mean or minimum values of the weight coefficients on the accuracy of upscaling is not obvious, which implies that determining the weight coefficients either from the nonsynchronous historical high spatial resolution data or from the coarse spatial resolution data themselves is feasible.

The TSEM-based upscaling model may achieve high accuracy with careful mathematical deduction; however, the acquisition of correction values would be strictly limited by the availability of high spatial resolution data. However, the weight coefficients of the ICGM-based upscaling model can be directly derived from the aggregated results of coarse spatial resolution data. In most cases, the ICGM-based upscaling model can achieve almost the same effect as the TSEM-based upscaling model. As shown in Fig. 7(d)-(f) and Fig. 8(d)-(f), with the ICGM-based upscaling model, the LAI values after correction concentrate on both sides of the line 1:1, and both the RE and RMSE decrease significantly after correction. In the traditional CGM-based upscaling model, which uses a weight coefficient of 1/2, the results appear to be excessively corrected at the three study sites.

V. CONCLUSIONS

In accordance with the convex hull of computational geometry theory, this paper proposed an ICGM-based upscaling model to correct for the scale effects of the LAI; in this model, the actual distribution of the NDVI is used to acquire weight coefficients for the proposed upscaling model.

By using dynamic weight coefficients for the upper and lower envelopes instead of constant coefficients, the accuracy of the upscaling model improved considerably. Three sites were selected for study using *in situ* observed data from the VELARI database. The results were evaluated by comparing the corrected value of the LAI to the exact value; good agreement was observed between the corrected LAI and the exact LAI with the ICGM-based upscaling model. At the Zhangbei, Haouz and Fundulea sites, the values of RE can be reduced from 3.35%, 11.01% and 19.62% to an average of 0.28%, 1.48% and 5.16%, respectively, and the value of RMSE can be reduced from 0.09, 0.11 and 0.11 to 0.024, 0.016 and 0.03, respectively, at the kilometer scale (40×40 aggregation pixel). In most cases, with the ICGM-based upscaling model, LAI products at coarse spatial resolution and with accuracies comparable to those of the TSEM-based upscaling model can be obtained.

Compared with the TSEM-based and CGM-based upscaling models, the advantages of the proposed ICGM-based upscaling model are readily apparent. The ICGM-based model can not only yield correction effects with high accuracy, but also be applied to all types of retrieval functions, regardless of whether

the functions are continuous or derivable. Furthermore, synchronous high spatial resolution data are not required as they are in the TSEM-based upscaling model because the weight coefficients of the ICGM-based upscaling model can be obtained from the aggregated results of nonsynchronous historical high spatial resolution data or the aggregated results of coarse spatial resolution data. Because the weight coefficients of the ICGM-based upscaling model change little at different scales, those coefficients are not sensitive to different scales and can be used as constants for a given study site. Consequently, this study will enhance the ability to obtain accurate values of LAI products at coarse spatial resolutions.

REFERENCES

- [1] H. Wu, and Z. -L. Li, "Scale Issues in Remote Sensing: A Review on Analysis, Processing and Modeling," *Sensors*, vol.9, no.3, pp. 1768-1793, Mar, 2009.
- [2] M. Raffy, "Change of Scale in Models of Remote Sensing: A General Method for Spatialization of Models," *Remote Sens. Environ.*, vol.40, no.2, pp. 101-112, May, 1992.
- [3] G. J. Yan, R. H. Hu, Y. T. Wang, H. Z. Ren, W. J. Song, J. B. Qi, and L. Chen, "Scale Effect in Indirect Measurement of Leaf Area Index," *IEEE Trans. Geosci. Remote Sens.*, vol.54, no.6, pp.3475-3484, Jun, 2016.
- [4] N. S. -N Lam, and D. A. Quattrochi, "On the issues of scale, resolution, and fractal analysis in the mapping sciences," *Prof Geog*, vol.44, no.1, pp.88-98, Feb, 1992.
- [5] C. Y. Cao, N. Lam, "Understanding the scale and resolution effects in remote sensing and GIS," In *Scale in remote sensing and GIS*: pp.57-72, 1997.
- [6] H. Wu, B. -H. Tang and Z. -L. Li, "Impact of nonlinearity and discontinuity on the spatial scaling effects of the leaf area index retrieved from remotely sensed data," *Int J Remote Sens*, vol.34, no.9-10, pp.3503-3519, 2013.
- [7] J. M. Chen, "Spatial Scaling of a Remotely Sensed Surface Parameter by Contexture," *Remote Sens. Environ.*, vol.69, no.1, pp.30-42, Jul, 1999.
- [8] Y. Tian, C. E. Woodcock, Y. Wang, J. L. Privette, N. V. Shabanov, L. M. Zhou, Y. Zhang, W. Buermann, J. R. Dong, B. Veikkanen, T. Ha'me, K. Andersson, M. Ozdogan, Y. Knyazikhin, and R. B. Myneni, "Multiscale analysis and validation of the MODIS LAI product: II. Sampling strategy," *Remote Sens. Environ.*, vol.83, no.3, pp.431-441, Dec, 2002.
- [9] Z. L. Hu, and S. Islam, "Effects of Spatial Variability on the Scaling of Land Surface Parameterizations," *Bound.-Layer Meteor.*, vol.83, no.3, pp. 441-461, Jun, 1997.
- [10] W. J. Fan, Y. Y. Gai, X. R. Xu, and B. Y. Yan, "The spatial scaling effect of the discrete-canopy effective leaf area index retrieved by remote sensing," *SCI China Earth SCI*, vol.56, no.9, pp.1548-1554, Sep, 2012.
- [11] Z. L. Hu, and S. Islam, "A Framework for Analyzing and Designing Scale Invariant Remote Sensing Algorithms," *IEEE Trans. Geosci. Remote Sens.*, vol.35, no.3, pp. 747-755, May, 1997.
- [12] S. Garrigues, D. Allard, F. Baret, and M. Weiss, "Influence of landscape spatial heterogeneity on the non-linear estimation of leaf area index from moderate spatial resolution remote sensing data," *Remote Sens. Environ.*, vol. 105, no.4, pp.286-298, Dec, 2006.
- [13] H. Pelgrum, "Spatial aggregation of land surface characteristics," Wageningen: Univ. of Wageningen, 2000.
- [14] L. Wang, W. J. Fan, X. R. Xu, and Y. Liu, "Scaling transform method for remotely sensed FAPAR based on FAPAR-P model," *IEEE Geosci. Remote Sens Lett.*, vol.12, no.4, pp:706-710, Apr, 2017.
- [15] T. Hu, Q. H. Liu, Y. M. Du, H. Li, H. S. Wang, and B. Cao, "Analysis of the Land Surface Temperature Scaling Problem: A Case Study of Airborne and Satellite Data over the Heihe Basin," *Remote Sens.*, vol.7, no.5, pp.6489-6509, May, 2015.
- [16] L. Wang, W. J. Fan, X. R. Xu, and Y. Liu, "Scaling Transform Method for Remotely Sensed FAPAR Based on FAPAR-P Model," *IEEE Geosci. Remote Sens Lett.*, vol.12, no.4, pp. 706-710, Apr, 2015.
- [17] A. Simic, J. M. Chen, J. Liu, and F. Csillag, "Spatial scaling of net primary productivity using subpixel information," *Remote Sens. Environ.*, vol.93, no.1-2, pp:246-258, Jul, 2004.
- [18] G. F. Yin, J. Li, Q. H. Liu, L. H. Li, Y. L. Zeng, B. D. Xu, L. Yang, and J. Zhao, "Improving leaf area index retrieval over heterogeneous surface by

integrating textural and contextual information: a case study in the heihe river basin,” *IEEE Geosci. Remote Sens Lett.*, vol.12, no.2, pp.359-363, Feb, 2015.

- [19] N. A. Brunsell, R. R. Gillies, “Scale issues in land-atmosphere interactions: implications for remote sensing of the surface energy balance,” *Agric. for. meteorol.*, vol.117, pp.203-221, Jan,2003.
- [20] R. H. Zhang, J. Tian, Z.-L. Li, H. B. Su, S. H. Chen, and X. Z. Tang, “Principles and methods for the validation of quantitative remote sensing products,” *SCI China Earth SCI*, vol.53, no.5, pp.741-751, Mar, 2010.
- [21] H. J. Luan, Q. J. Tian, T. Yu, X. L. Hu, Y. Huang, L. T. Du, L. M. Zhao, X. Wei, J. Han, Z. W. Zhang, and S. P. Li, “Modeling continuous scaling of NDVI based on fractal theory,” *Spectrosc Spect Anal*, vol.33, no.7, pp.1857-1862, Jul, 2013.
- [22] L. Wu, X. N. Liu, X. P. Zheng, Q. M. Qin, H. Z. Ren, and Y. J. Sun, “Spatial scaling transformation modeling based on fractal theory for the leaf area index retrieved from remote sensing imagery,” *J Appl Remote Sens*, vol.9, pp.096015, Aug, 2015.
- [23] L. Wu, Q. M. Qin, X. N. Liu, H. Z. Ren, J. H. Wang, X. P. Zheng, X. Ye, and Y. J. Sun, “Spatial up-scaling correction for leaf area index based on the fractal theory,” *Remote Sens.*, vol.8, no.3, pp.197, Feb, 2016.
- [24] J. L. Jiang, X. N. Liu, C. H. Liu, L. Wu, X. P. Xia, M. L. Liu, and Z. H. Du, “Analyzing the spatial scaling bias of rice leaf area index from hyperspectral data using wavelet-fractal technique,” *IEEE J. Sel. Top. Appl. EarthObs. Remote Sens.*, vol.8, no.6, pp.3068-3080, Jun, 2015.
- [25] H. Chen, H. Wu, Z.-L. Li, B.-H. Tang, R. L. Tang and G. J. Yan, “Spatial upscaling of remotely sensed leaf area index based on discrete wavelet transform,” *Int J Remote Sens*, vol.40, no.5-6, pp.2343-2358, Apr, 2018.
- [26] J. L. Jiang, X. S. Ji, X. Yao, Y. C. Tian, Y. Zhu, W. X. Cao and T. Cheng, “Evaluation of Three Techniques for Correcting the Spatial Scaling Bias of Leaf Area Index,” *Remote Sens.*, vol.10, no.2, pp.221, Feb, 2018.
- [27] Y. Huang, Q. J. Tian, J. Geng, L. Wang, and H. J. Luan, “Review of spectral and spatial scale effects of remotely sensed biophysical and biochemical vegetation parameters,” *Acta Ecologica Sinica*, vol.36, no.3, pp.883-891, Feb, 2016.
- [28] J. W. Rouse, R. H. Haas, J. A. Schell, and D.W. Deering, “Monitoring vegetation systems in the great plain with ERTS,” *Proceedings of the 3rd ERTS Symposium*, pp.48-62, 1973.
- [29] M. T. VanWijk, M. Williams, and G. R. Shaver, “Tight coupling between leaf area index and foliage N content in arctic plant communities,” *Oecologia*, vol.142, no.3, pp.421-427, 2005.
- [30] X. R. Xu, W. J. Fan, and X. Tao, “Spatial scale effect in retrieval of leaf area index of continuous vegetation,” *SCI China Earth SCI*, vol.39, no.1, pp.79-87, 2009.



Hong Chen received the B.E. degree in Central South University, Changsha, China, in 2006, the M.S. degree in Beijing Normal University, Beijing, China, in 2009 and the Ph.D. degree in Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, China, in 2018.

She is currently an Engineer with China Aero Geophysical Survey & Remote Sensing Center for Nature Resources. Her research interest mainly focuses on scaling of remotely sensed products.



Hua Wu received the Ph.D. degree in cartography and geographical information system from the Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing, in 2010.

He is currently an Associate Research Fellow with the Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences. His research mainly includes the

retrieval and validation of surface temperature and emissivities, and scaling of remotely sensed products.

Zhao-Liang Li received the Ph.D. degree from the University of Strasbourg, Strasbourg, France, in 1990.

Since 1992, he has been a Research Scientist with ICube, Illkirch, France. He joined the Institute of Agricultural Resources and Regional Planning, Beijing, China, in 2012. He has participated in many national and international projects such as the National Aeronautics and Space Administration (NASA)-funded Moderate Resolution Imaging Spectroradiometer, European Community (EC)-funded program Exploitation of angular effects in land surface observations from satellites (EAGLE), and ESA-funded programs SPECTRA. He has published more than 100 papers in international refereed journals. His research interests include thermal infrared radiometry, parameterization of land surface processes at coarse spatial resolution, and assimilation of satellite data to land surface models.



Jienan Tu received the B.E. and M.S. degree in Xi'an University of Science and Technology, Xi'an, China, in 2008 and the Ph.D. degree in Institute of Geology and Geophysics, Chinese Academy of Sciences, Beijing, China, in 2012.

He is currently a Senior Engineer working at China Aero Geophysical Survey & Remote Sensing Center for Nature Resources. His research interest mainly focuses on remote sensing of environment and geology disasters.