Quantitative forest canopy structure assessment using an inverted geometric-optical model and up-scaling

Yuan Zeng \(^{a, b}\), Michael E. Schaepman \(^{a}\), Bingfang Wu \(^{b}\), Jan G. P. W. Clevers \(^{a}\), Arnold K. Bregt \(^{a}\)

\(^{a}\) Centre for Geo-Information, Wageningen University, Wageningen, The Netherlands
\(^{b}\) Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing, China

Online Publication Date: 01 January 2009
Quantitative forest canopy structure assessment using an inverted geometric-optical model and up-scaling

YUAN ZENG*†‡, MICHAEL E. SCHAEPMAN†, BINGFANG WU‡, JAN G. P. W. CLEVERS† and ARNOLD K. BREGT†

†Centre for Geo-Information, Wageningen University, Wageningen, The Netherlands
‡Institute of Remote Sensing Applications, Chinese Academy of Sciences, Beijing, China

(Received 2 July 2007; in final form 12 November 2007)

The physical-based geometric-optical Li–Strahler model can be inverted to retrieve forest canopy structural variables. One of the main input variables of the inverted model is the fractional component of sunlit background ($K_g$). $K_g$ is calculated by using pure reflectance spectra (endmembers) of the viewed surface components. In this paper, the feasibility of up-scaling from high (Quickbird) to medium (Hyperion) spatial resolution data for extracting the required endmembers is demonstrated. Furthermore, the sensitivity of the endmembers used as input for inverting Li–Strahler model is evaluated. After validating the inverted model results, namely spatially explicit forest mean crown closure and crown diameter using field measurements, it can be concluded that the regional scaling-based endmembers derived from the linear unmixing model are the best ones to be used in combination with the inverted Li–Strahler model for quantitatively monitoring disturbance in forest canopy structure.

1. Introduction

At local to regional and global scales, remote sensing has facilitated extraordinary advances in the modelling, mapping and understanding of ecosystems and their functioning. One basic characteristic of remote sensing in the twenty-first century is the extensive use of quantitative methods for estimating earth surface variables (Liang 2004). Forests, being one of the most important natural resources worldwide, not only regulate the global atmospheric cycles, but are also increasingly being used in dynamic global vegetation models for terrestrial CO$_2$ estimations (Sitch et al. 2003). Forest parameters, such as leaf area index (LAI), species diversity, canopy structural attributes (such as crown closure (CC) and crown diameter (CD)), etc. can directly indicate the condition and change of a forest ecosystem. Additionally, forest structure influences hydrologic processes, biogeochemical cycles and many interactions between the land surface and the atmosphere (Ustin 2004). Changes in forest structure may also provide information related to forest vigour, harvests, burns, stocking level, disease and insect infestations (Gillis and Leckie 1996). Thus, monitoring forest structure by quantitatively deriving the canopy structural variables over large areas improves our understanding of several environmental processes and allows accurate estimates of relevant disturbance processes.

*Corresponding author. Email: yuanz@irsa.ac.cn
Various authors monitoring forest canopy variables using remote sensing have focussed on inverting physical-based canopy reflectance models (e.g. Goodenough et al. 2006). The use of geometric-optical models allows the assessment of one (major) component of the anisotropy (Wanner et al. 1995); however, the complete separation of isotropic and volumetric scattering remains challenging (Schaepman-Strub et al. 2006). Many approaches have been presented treating the surface as an assemblage of discrete geometric objects with the reflectance modelled as a linear combination of viewed sunlit and shaded components. These have been used successfully to estimate forest canopy attributes (Franklin and Strahler 1988, Woodcock 1994, Hall et al. 1995, Gerard and North 1997, Woodcock et al. 1997, Gemmell 1999, Peddle et al. 1999, Scarth and Phinn 2000, Scarth et al. 2001, Peddle et al. 2003, Chopping et al. 2006). In addition, Zeng et al. (2007) derived pixel-based forest canopy structural variables (CC and CD) from EO-1 Hyperion data by inverting the Li–Strahler geometric-optical model. One of the most important input variables of the inverted model is the fractional image of sunlit background, which is calculated based on pure reflectance spectra of the viewed surface components, approximated using endmembers. Consequently, the method used for endmember extraction remains the main factor influencing the accuracy of a geometric-optical model in general, and the inverted Li–Strahler geometric-optical model in particular.

Different approaches may be used to extract endmembers, such as deriving the pure spectrum from an image with experimental knowledge of field training samples, from the observation with a field spectrometer and from an existing spectral library. Over the past years, several methods have been proposed for the purpose of autonomous and supervised endmember selection from hyperspectral data (Plaza et al. 2004). Among them, the most common way used for extracting the endmembers of sunlit and shaded scene components is calculating the mean spectra of detected extreme pixels directly from the image. However, using this image-based endmember extraction method, the detected ‘pure’ pixels may still contain mixtures of components that may not fall in the ‘sunlit background’ category due to a non-appropriate sensor spatial resolution, or the inherent fragmentation of landscapes used in these applications. In such a case, the mean spectra of the endmember pixels cannot represent this component correctly, and it remains as mixed pixels.

Solving the above well-known mixture problem, an up-scaling method using linear unmixing has been used (Zhukov et al. 1999, Haertel and Shimabukuuro 2005, Zurita-Milla et al. 2006). When a concurrent high spatial resolution image is available, the linear unmixing model can be applied to unmix the low spatial resolution data and calculate the required endmembers. This method also provides an avenue to up-scale the information from local to regional scales. In addition, the high spatial resolution data can also be used to evaluate the purity of those extreme pixels selected by methods applied to lower spatial resolution images only. By up-scaling the abundance fractions of scene components from the high spatial resolution image to each extreme pixel, the purest single pixel may be determined for every component. We address this scaling-based endmember extraction method as the purest-pixel selection process.

The objective of this paper is to study the feasibility of up-scaling from high spatial resolution data to medium spatial resolution hyperspectral data for extracting the endmembers of the viewed surface components. Subsequently, how
these endmembers influence and improve the results of the inverted Li–Strahler geometric-optical model will be evaluated. Finally, the advantages and disadvantages of the image-based and scaling-based endmember extraction methods used for monitoring regional scaled forest canopy structural variables will be discussed.

2. Methods

2.1 Study area

The study area selected is the Longmenhe forest natural reserve (centred at 31°20' N, 110°29' E), in Xingshan county, Hubei province (China, see figure 1). It is located towards the northeast of the Three Gorges region and 80 km away from the Three Gorges dam. The total reserve size is about 4644 ha and the altitude is around 1300 m above sea level. The study area belongs to the temperate climate zone (Köppen: Cwa-subtropical monsoon (McKnight and Hess 2000)). The average precipitation is about 100–150 mm per month. During spring and summer (April to September), this area can even receive an average precipitation of 200–300 mm per month, rendering timely satellite data acquisition very difficult. The Longmenhe forest reserve is mainly occupied by natural subtropical evergreen broadleaved forest (dominant species are *Quercus spinosa* and *Cyclobalanopsis oxyodon*), deciduous broadleaved forest (dominant species are *Platycarya strobilacea*, *Quercus glandulifera var. brevipetiolata* and *Betula luminifera*) and coniferous forest (dominant species are *Pinus tabulaeformis* and *Larix kaempferi*).

2.2 Field data

The field data were collected in June 2003. With support of 1:50 000 topographic maps, a total of 40 sample sites (100 × 100 m) within the study area were selected based on different plant distribution patterns and topographic strata (figure 2(a)). The central location of each ground sample site was determined by using differential global positioning system (GPS), with an accuracy of ±5 m. At every sample site, five sample plots (20 × 20 m) were collected by overlaying each sample site with a grid of 20 m cells and then randomly selecting five grid cells. At each sample plot, measurements of forest structural variables were carried out. Those include forest CC, CD, stem diameter at breast height (DBH), tree height (H), trunk height (TH) and stem density (SD).

Crown closure (CC), also known as crown cover, is the percentage of ground covered by a vertical projection of the outermost perimeter of the crowns in a stand (Brack 1999). Crown diameter (CD), also called crown width, is the span of crown and used to describe the crown size (About 2007). The positions of all trees with DBH ≥5 cm in each sample plot were recorded. The value of CD for each tree was measured in the field by taking twice the distance between the projected outer boundary of the crown. Then, all crowns were drawn on a crown map and from this map, the CC was determined. Tree height (H) and trunk height (TH) were estimated by performing angular measurements using a compass clinometer. Stem density (SD) was collected by counting the number of trees in each plot. Besides these forest canopy variables, information of forest type, plant species and vertical distribution were all recorded by experts. Figures 2(b)–(d) show a canopy photograph within the sample plot, the crown map and the plant profile from three different sample sites,
representing three dominant forest communities (deciduous broadleaved forest, evergreen broadleaved forest and coniferous forest) in the study area.
2.3 Image data and processing

2.3.1 EO-1 Hyperion. The EO-1 Hyperion scene used in this study was acquired on 10 June 2004, around 11:00 a.m. local time. Hyperion, one of the three sensors on the NASA EO-1 platform, was launched on November 2000. Being a push-broom imaging instrument, Hyperion provides imagery with 242 spectral bands (of which 196 are non-zero and not overlapping), at a spectral resolution of 10 nm and 30 m spatial resolution. The conversion of digital numbers (DNs) to radiances is performed using a scaling approach proposed by Beck (2003) (radiance at visible and near-infrared bands (VNIR) = DN/40; radiance at short-wavelength infrared bands (SWIR) = DN/80). Several stripes (data columns of poor quality) in the Hyperion data contain no information and/or unusual low radiance. Those pixels are detected and replaced by the average radiance value of their immediate left and right neighbouring pixels using the method proposed by Han et al. (2002) and Zeng et al. (2007). To obtain the hemispherical directional reflectance factor (HDFR) approximated by surface reflectance (Schaepman-Strub et al. 2006) of this image, we use ACORN version 4.0, a commercially available atmospheric correction program based on MODTRAN 4 radiative transfer code (AIG 2002). ACORN uses two water absorption channels (940 and 1140 nm) to evaluate the amount of water vapour in combination with the visibility at the moment of data acquisition. Due to the low signal-to-noise ratio at the beginning and the end of the spectral range (<436 nm and ≥2385 nm) and significant water absorption influence in several spectral bands, a total of 64 bands are dropped from the initial 196 bands. Geometric correction is performed using 26 ground control points (GCPs) relative to 1:50 000 topographic maps and the geometric co-registration error is less than one 30 m pixel. After image pre-processing, the geolocated Hyperion data expressed as surface reflectance in UTM Zone 49 N (WGS-84 datum) and 30 m spatial resolution is finally depicted as in figure 3(a). The dimension of the image is 208 (columns) × 173 (lines) × 132 (spectral bands).

To be able to use the Hyperion image for mapping forest canopy structural variables, a forest classification needs to be first performed. This classification can be used not only for identifying the forested/non-forested regions, but also for distinguishing three dominant forest communities, as previously mentioned. The spectral angle mapping (SAM) (Kruse et al. 1993) algorithm is used to classify the Hyperion data. Since every homogeneous forest region is much larger than one pixel of 30 m, seven pixels, indicating the spectra of typical forest classes, are detected from the Hyperion image based on the field measurements. We separately recode the deciduous broadleaved forest (DBF), evergreen broadleaved forest (EBF) and coniferous forest (CF) into classes 1, 2 and 3. The unclassified regions include both the non-forested area, as well as the pixels that were affected by clouds and shadows. The processing and accuracy statement of this Hyperion forest classification have been discussed in Zeng et al. (2007).

2.3.2 QuickBird. The QuickBird data consists of a panchromatic image at 0.61 m spatial resolution and a multi-spectral image with four spectral bands (blue/green/red/near-infrared) at 2.44 m spatial resolution. It was collected on 23 August 2003 at 11:08 a.m. local time. This QuickBird image serves as high spatial resolution data and is used together with the Hyperion data for deriving the endmembers of the various scene components. Although the two images were acquired at different dates, the solar zenith and azimuth angles are similar (23.5° and 104.5° for Hyperion; 29.2° and 127.9° for
QuickBird) and the sensor viewing directions are almost the same (nadir looking). Hence, only negligible variation in view/illumination geometry is expected to exist.

We first merge the panchromatic and multi-spectral QuickBird images using a principal component method and a cubic convolution re-sampling technique. The output image includes a 0.6 m spatial resolution and four spectral bands. Then, an object-oriented classification (eCognition) is carried out to classify the sunlit and shaded scene components. eCognition is designed to segment the image into units of similar spectral and spatial patterns and to classify those segments according to a pre-defined rule base (Baatz et al. 2004). This object-based image processing technique is particularly suited for the analysis of very high spatial resolution remote sensing data. The image segmentation groups pixels in each object or segment based on three parameters: scale, colour and shape (smoothness and compactness) (Luscier et al. 2006).

In this case, we apply the following rules for the class assignments: scale parameter 5, 10 and 25 based 80% on colour and 20% on shape, having 90% smoothness and 10% compactness. The object-classified QuickBird image is shown in figure 3(b). Through geometric correction with 20 GCPs and 1:10 000 topographic maps, the QuickBird image is used as a base image to spatially co-register the Hyperion data. Hence, a Hyperion subset image with 86 (columns) × 68 (lines) is masked for matching the QuickBird classification image, such that each Hyperion pixel corresponds to 50 × 50 QuickBird pixels, the latter including 4300 (columns) × 3400 (lines).

2.4 Model building

The inverted Li–Strahler geometric-optical model is used to retrieve the forest canopy structural variables CC and CD. The methodology and approach selected is presented.
Figure 4. Main methodological approaches used in this study, illustrated as a flow chart. Three methods have been tested and compared with ground samples.

In figure 4. The model mainly requires three scene components: the sunlit background \((G)\), the sunlit canopy \((C)\) and the shadow \((T)\). The endmember abundances of these three components are extracted from the Hyperion data using three approaches:

1. Image-based pixel purity index (PPI) method: derivation of the extreme pixels from the Hyperion image automatically, with clustering of the pixels representing the three components by their mean spectra;
2. Scaling-based purest-pixel selection: evaluation of the purity of the PPI extreme pixels based on the classification of the QuickBird components. After this comparison, the spectrum of the purest extreme (singular) pixel is used to represent the three components, respectively; and
3. Scaling-based linear unmixing model: inversion of the linear spectral unmixing approach in the overlapping region of QuickBird and Hyperion. For each pixel of the Hyperion image, the proportions of the three components are provided by the QuickBird classification.
CC and CD are estimated from the inverted Li–Strahler model by integrating: (1) the fractional images of sunlit background \((K_g)\) obtained by forward linear spectral unmixing of the Hyperion data using the extracted endmembers; (2) slope and aspect images derived from a digital elevation model (DEM); and (3) mean canopy structural parameters for each dominant forest class, obtained from field measurements. Finally, we use the field-measured CC and CD from the 40 sample sites to validate the model estimation.

2.4.1 Inverted geometric-optical model. The Li–Strahler geometric-optical model (Li and Strahler 1985, 1992) is based on the assumption that the bidirectional reflectance distribution function (BRDF) is modelled as a purely geometric phenomenon, resulting from a scene of discrete three-dimensional objects being illuminated and viewed from different positions in the hemisphere. The reflectance associated with a given viewpoint is treated as an area-weighted sum of four fixed reflectance components: sunlit canopy, sunlit background, shaded canopy and shaded background. Moreover, in most cases, these four components can be simplified to three: sunlit background, \(G\), sunlit canopy, \(C\), and shadow, \(T\) (Hall et al. 1995, Li and Wang 1995, Peddle et al. 1999, Peddle et al. 2003). This model also assumes that the resolution of the remote sensing image is much larger than the size of individual crowns, but smaller than the size of forest stands, and that the individual trees are randomly (Poisson) distributed within the pixel (Woodcock et al. 1994). Based on the principle of three-dimensional geometry of a spherical crown on a flat background, each proportion of components can be expressed by a combination of the forest canopy structural parameters. For inverting this model, the fractional abundance of sunlit background \((K_g)\) can be used for deriving the expected forest mean CC and CD, see equations (1) to (6) (Strahler and Jupp 1990, Li and Strahler 1992, Woodcock et al. 1997):

\[
K_g = e^{-\pi M [\sec \theta_i + \sec \theta_v - O(\theta_i, \theta_v, \varphi)]},
\]

\[
O(\theta_i, \theta_v, \varphi) = 1/\pi (\sec \theta_i + \sec \theta_v) (t - \sin t \cos t),
\]

\[
\cos t = \frac{h \tan \theta_i - \tan \theta_v \cos \varphi}{r (\sec \theta_i + \sec \theta_v)},
\]

\[
M = \frac{-\ln(K_g)}{(\sec \theta_i + \sec \theta_v)(\pi - t + \cos t \sin t)},
\]

\[
CC = 1 - e^{-\pi M}
\]

and

\[
(CD/2)^2 = \sqrt{(1 + \omega)^2 M^2 + 4V(m)\omega - (1 + \omega)M} / 2\omega,
\]

where \(\theta_i\) and \(\theta_v\) are the zenith angles of illumination and viewing direction, \(O(\theta_i, \theta_v, \varphi)\) is the average of the overlap function between illumination and viewing shadows of individual crowns as projected onto the background, \(\varphi\) is the difference in azimuth angle between illumination \((\varphi_i)\) and viewing \((\varphi_v)\), \(M\) is the mean of the
‘treeness’ parameter $m$, $V(m)$ is the variance of the calculated pixel-based $m$ and $\omega$ is the coefficient of variation of the crown size, which is defined as the ratio of the mean to the variance of the squared crown radius.

Since the study area is a mountainous region, the crown shapes of the broadleaved forest need to be modelled, representing an ellipsoid, with tree height from ground to mid-crown ($h$), crown radius in vertical direction ($b$) and crown radius in horizontal direction ($r$). In addition, double transformations are required to allow crowns to be treated as spheres and accommodating the sloping surfaces (Schaaf et al. 1994). The equations for all transformations of $\theta_i$, $\phi_i$, $\theta_v$, $\phi_v$ and $h$ are explained in detail in Zeng et al. (2007).

Resulting from the above, the input of the inverted Li–Strahler model to determine pixel-based CC and CD are the fractional image of $K_g$, the solar zenith and azimuth angles, the view zenith and azimuth angles, the local slope and aspect images and the mean measured parameters for different kinds of forest crown shapes: $h$, $b$ and $r$. Among them, deriving the proportional image of sunlit background is the most crucial process, which requires accurately extracting the image endmembers of the three scene components $G$, $C$ and $T$.

2.4.2 Image-based endmember extraction. The PPI (Boardman 1993) method is a widely used approach to find the best ‘spectrally pure’ (extreme) pixels in multispectral and hyperspectral images (Plaza et al. 2004, Chang and Plaza 2006). The PPI is computed by repeatedly projecting $n$-dimensional scatter-plots onto a random unit vector, and then the extreme pixels in each projection (those pixels that fall onto the ends of the unit vector) are recorded and the total number of times each pixel is marked as extreme is noted. To implement PPI, we need to provide two parameters, $k$ and $l$. After $k$ (the number of iterations) times repeated projections to different random lines, those pixels that count above $l$ (a certain cut-off threshold) are declared ‘pure’. These extreme pixels are loaded into an interactive $n$-dimensional visualization tool and rotated in real time until a desired number of endmembers are visually identified as clusters. Then, the mean spectrum of the clustered extreme pixels is the final endmember.

2.4.3 Scaling-based purest-pixel selection. This method is based on the theory of selecting the purest single pixel from a hyperspectral image to indicate each endmember of interest. The potential pure pixels are coming from the derived extreme pixels through the PPI method. When the pixel size of the hyperspectral image is too wide to represent a 100% pure component in real life, we need to upscale the information of component fractions within each extreme pixel from a co-registered high spatial resolution image to evaluate the real purity of these extreme pixels. After overlaying and locating every extreme pixel onto the high spatial resolution image, we list the mean proportions, calculated from the high spatial resolution data, and select the purest pixel among these extreme pixels to be the endmember for each component.

2.4.4 Scaling-based linear unmixing model. Traditionally, the linear spectral unmixing approach has been widely used to calculate the percentages of several individual surface components contained in each pixel of a remote sensing image (Adams et al. 1986, Settle and Drake 1993, Peddle et al. 1999, Goodwin et al. 2005). The model assumes that the reflectance ($S$) of each pixel is a linear combination of endmembers ($E$), which are the pure reflectance spectra for each component. The general equations are:
\[ S_j = \sum_{i=1}^{q} K_i E_{i,j} + \epsilon_j \quad j = 1, 2, ..., p \]  

(7)

and

\[ 1 = \sum_{i=1}^{q} K_i, \quad K_i \geq 0, \]  

(8)

where \( q \) is the number of components, \( p \) is the number of image bands, \( K \) is the fractional abundance of each component within the pixel and \( \epsilon \) is the residual for each band.

However, in this study, we use an inverted approach of this model. The fractions of each component (\( K \)) within the pixel of a hyperspectral image are obtained from an overlapping high spatial resolution image, and the endmembers (\( E \)) of each component are the final requirements. In practice, for deriving the endmembers, we need the overlapped region of the two images with at least \( f \) pixels, and the number \( f \) must be more than the number of components. Thus, equation (7) can be more conveniently expressed in matrix notation (9):

\[ \epsilon_j = S_j - K E_j, \]  

(9)

where \( \epsilon_j \) is the \( f \)-dimensional vector of the residuals in band \( j \), \( S_j \) represents the \( f \)-dimensional vector of the pixels’ reflectance in band \( j \), \( K \) is the \( f \times q \) matrix of the fractions, and \( E_j \) is the \( q \)-dimensional vector of the reflectance of the components in band \( j \).

We then seek a set of numerical values for the unknowns in \( E_j \), such that the sum of the residual squares becomes minimum (see equation (10)). Then the least squares solution for \( E_j \) can be calculated using equation (11) (Haertel and Shimabukuro 2005). After applying equation (11) to every spectral band of the hyperspectral image, the calculated reflectance spectra of the components are finally aggregated to become the required endmembers:

\[ \frac{\partial (\epsilon_j \cdot \epsilon_j)}{\partial E_j} = 0 \]  

(10)

and

\[ E_j = (K^T K)^{-1} K^T S_j. \]  

(11)

2.4.5 Validation. For validating the inverted model results of CC and CD, we calculate the mean value of 3 x 3 pixels (90 x 90 m) from the output image for comparison to one field sample site of 100 x 100 m. In total, 40 sample sites with field-measured CC and CD can be used to validate the model estimation. The obtained model estimated values are plotted against the corresponding ground measurements, and their agreement is assessed by the coefficient of determination (\( R^2 \)) and the root mean square error (RMSE):

\[ \text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \hat{x}_i)^2}, \]  

(12)
where \( x_i \) is the model estimated CC and CD, \( \hat{x}_i \) is the ground-measured CC and CD, \( n \) is the number of sample sites, which could be interpreted from the inverted Li–Strahler model depending on the different input endmembers.

For evaluating the suitability of the three endmember extraction methods, we use \( t \)-statistics as the statistical indicator. This allows several methods to be compared based on the sample size \( (n) \) and indicates whether or not the model estimations are statistically significant at a particular confidence level (Stone 1993, Tadros 2000). It is calculated using both the RMSE and the mean bias error (MBE). The smaller the value of \( t \), the better the correspondence between model estimate and field measurement, thus the model performance. We determine a critical \( t \) value as \( t_{a/2} \) at \( \alpha \) level of significance and \( n - 1 \) degrees of freedom:

\[
t = \sqrt{\frac{(n-1)(\text{MBE})^2}{\text{RMSE}^2 - (\text{MBE})^2}}.
\] (13)

3. Results

First, the PPI method is applied to extract endmembers of the three components \( G \), \( C \) and \( T \) from the Hyperion image. The parameter \( k \) is initially set to \( 10^4 \) and the threshold value of \( l \) is determined as 500 by iteration, which results into 10 extreme pixels, including all kinds of components. After auto-clustering, these derived extreme pixels are grouped into three clusters (\( G \): four extreme pixels; \( C \): two extreme pixels and \( T \): four extreme pixels). The mean spectra of each cluster, as the image-based endmembers, are shown in figure 5(a).

The 10 extreme pixels are all located in the overlapping region of the Hyperion and QuickBird image (figure 3(a)). Consequently, the scaling-based method of purest-pixel selection can be used for estimating the purity of each extreme pixel. From the co-registered classified QuickBird image, 10 windows of \( 50 \times 50 \) pixels are matched to every extreme pixel from the Hyperion image. The calculated mean proportions of \( G \), \( C \) and \( T \) components are listed in table 1. Not every extreme pixel automatically detected by the PPI method is a real ‘pure’ pixel, most of them are mixtures. The purity of the PPI endmembers is shown in table 2. It indicates that the endmembers of \( G \) and \( C \), containing 0.7616 and 0.8102 fractions of the components \( G \) and \( C \), respectively, are more accurate than the endmember \( T \), which only involves 0.5203 proportion of \( T \) mixed with 0.4775 proportion of \( C \). After comparing the \( G \), \( C \) and \( T \) components of these 10 extreme pixels, three purest pixels (table 1: \( G3 \), \( C1 \) and \( T3 \)) are selected. Subsequently, the Hyperion reflectance spectra of these three pixels are used to express the endmembers of \( G \), \( C \) and \( T \) separately (see figure 5(b)). Those are called the pixel scaling-based endmembers.

Subsequently, the linear unmixing model is applied in the overlapping area of the Hyperion and classified QuickBird images. The model parameter \( q \) is 3 with \( G \), \( C \)

Figure 5. Extracted \( G \) (sunlit background), \( C \) (sunlit canopy) and \( T \) (shadow) endmembers: (a) image-based endmembers extracted using the PPI method on Hyperion (Hy) data; (b) scaling-based endmembers extracted by selecting the purest pixel from the PPI extreme pixels based on the QuickBird (QB) classification; and (c) regional scaling-based endmembers extracted using the linear unmixing approach based on the overlapping region of QuickBird and Hyperion.
and $T$ components, $p$ is 132 bands of corrected Hyperion data and $f$ is 5848 pixels, which are located in the central region of the Hyperion image (figure 3(a)). Therefore, the final aggregated spectra from the linear unmixing model based on both the QuickBird fractional information and the Hyperion reflectance data can provide the endmembers for the whole Hyperion image. We call them the regional scaling-based endmembers, and they are shown in figure 5(c).

According to the extracted endmembers above, a forward linear spectral unmixing (equations (7) and (8)) is performed to derive the proportional image of sunlit background ($K_g$) from the whole Hyperion image that is one of the required inputs for the inverted Li–Strahler model. There are some areas where the model produces a negative $K_g$. It implies that these areas just consist of a combination of sunlit canopy and shadow, and the linear spectral unmixing cannot achieve a logical output of sunlit background in these areas based on the provided endmembers. Hereby, we recode these areas as infeasible region in the final output images. The other required inputs for the inverted Li–Strahler model, slope and aspect images with 30 m spatial resolution, are created from DEM data using a general topographic analysis model. The mean canopy structural parameters $h$, $b$, $r$ and $\omega$ for each dominant forest class obtained from the field measurements are shown in table 3.

The inversion of the Li–Strahler model is implemented with the pixel-based inputs. Figure 6 illustrates the final mapping results of forest CC and CD distributed in the Longmenhe study area. The region in white shows the non-forested and cloud-influenced area determined by the forest classification. The black region presents the infeasible area. Generally, the proportion of the infeasible area can point out how suitable the extracted endmembers are for this Hyperion image. The output images show 41.2% (figures 6(a)–(b)), 13.9% (figures 6(c)–(d)) and 35.8%

### Table 1. Fractions of $G$, $C$ and $T$ calculated from the QuickBird classification for each extreme pixel derived using the PPI method.

<table>
<thead>
<tr>
<th>Extreme pixels</th>
<th>$G$</th>
<th>$C$</th>
<th>$T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$G_1$</td>
<td>0.5092</td>
<td>0.4576</td>
<td>0.0332</td>
</tr>
<tr>
<td>$G_2$</td>
<td>0.7884</td>
<td>0.1656</td>
<td>0.0460</td>
</tr>
<tr>
<td>$G_3$</td>
<td>0.9976</td>
<td>0.0024</td>
<td>0.0000</td>
</tr>
<tr>
<td>$G_4$</td>
<td>0.7512</td>
<td>0.2236</td>
<td>0.0252</td>
</tr>
<tr>
<td>$C_1$</td>
<td>0.0000</td>
<td>0.9804</td>
<td>0.0196</td>
</tr>
<tr>
<td>$C_2$</td>
<td>0.3528</td>
<td>0.6400</td>
<td>0.0072</td>
</tr>
<tr>
<td>$T_1$</td>
<td>0.0036</td>
<td>0.6428</td>
<td>0.3536</td>
</tr>
<tr>
<td>$T_2$</td>
<td>0.0000</td>
<td>0.2928</td>
<td>0.7072</td>
</tr>
<tr>
<td>$T_3$</td>
<td>0.0000</td>
<td>0.2060</td>
<td>0.7940</td>
</tr>
<tr>
<td>$T_4$</td>
<td>0.0052</td>
<td>0.7684</td>
<td>0.2264</td>
</tr>
</tbody>
</table>

### Table 2. Calculated average purity (fractions of $G$, $C$ and $T$ from QuickBird classification) of the endmembers derived using the PPI method.

<table>
<thead>
<tr>
<th>PPI clusters</th>
<th>$G$</th>
<th>$C$</th>
<th>$T$</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPI endmember $G$</td>
<td>0.7616</td>
<td>0.2123</td>
<td>0.0261</td>
</tr>
<tr>
<td>PPI endmember $C$</td>
<td>0.1764</td>
<td>0.8102</td>
<td>0.0134</td>
</tr>
<tr>
<td>PPI endmember $T$</td>
<td>0.0022</td>
<td>0.4775</td>
<td>0.5203</td>
</tr>
</tbody>
</table>
Table 3. Inverted Li–Strahler model input parameters (mean canopy structure) for each forest class derived from the field measurements.

<table>
<thead>
<tr>
<th>Dominant forests</th>
<th>$h$ (m)</th>
<th>$b$ (m)</th>
<th>$r$ (m)</th>
<th>$\omega$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deciduous broadleaved forest</td>
<td>9.79</td>
<td>3.97</td>
<td>1.79</td>
<td>1.25</td>
</tr>
<tr>
<td>Evergreen broadleaved forest</td>
<td>8.86</td>
<td>3.36</td>
<td>1.61</td>
<td>3.02</td>
</tr>
<tr>
<td>Coniferous forest</td>
<td>8.41</td>
<td>4.63</td>
<td>1.51</td>
<td>1.87</td>
</tr>
</tbody>
</table>

(figures 6(e)–(f)) proportion of infeasible region for the image-based PPI endmembers, the pixel scaling-based endmembers and the regional scaling-based endmembers, respectively.

Figure 6. Mapping results of forest crown closure (CC, left column) and crown diameter (CD, right column) in the Longmenhe study area: (a) CC map using image-based PPI endmembers; (b) CD map using image-based PPI endmembers; (c) CC map using pixel scaling-based endmembers; (d) CD map using pixel scaling-based endmembers; (e) CC map using regional scaling-based unmixing endmembers; and (f) CD map using regional scaling-based unmixing endmembers.
The scatter plots in figure 7 show the reliability between the model-interpreted values and the corresponding ground measurements. The closer the points are to the 1:1 line (diagonal line), the less the RMSE is. For the image-based PPI endmembers, the model provided no output for two sample sites because they are located in the infeasible region (figures 7(a)–(b)). The values of $R^2$ and RMSE based on 38 sample sites are $R^2_{CC}=0.51$, RMSE$_{CC}=0.055$, $R^2_{CD}=0.26$ and RMSE$_{CD}=1.07$ (table 4). For the pixel scaling-based endmembers, estimated output images include all 40 sample sites since none are in the infeasible region. The results are demonstrated in figure 7(c)–(d) and the results for the 40 sample sites are

![Figure 7. Corresponding scatter plots of the results presented in figure 6. Values of root mean square error (RMSE) between the model-estimated CC/CD and the ground measured CC/CD are listed. Dotted line represents a 1:1 relationship.](image-url)

Y. Zeng et al.
For comparing the three endmember extraction methods, we calculate the $t$ values at 0.01 level of significance ($\alpha=0.01$) using the same samples ($n=38$) that are common in all three methods (table 5). The endmembers from the regional scaling-based linear unmixing model, combining QuickBird and EO-1 Hyperion images, are the most suitable endmembers for deriving CC by inverting the Li–Strahler model. The $t$ value (0.301) of this method is much less than the critical $t$ (2.715). That indicates the model-derived CC by this method is not significantly different from the field measurements. However, the measured CC and the model-estimated CC by the endmembers extracted through the PPI method and the purest-pixel selection approach do have significant differences ($t>2.715$). Thus, the model performs worse based on these two methods. For deriving CD, the $t$ values show that the model produced estimates using the endmembers by the three methods are all statistically different from the measured CD. Therefore, the model does not perform well for CD.

4. Discussion

Our results show that the Li–Strahler model can be inverted to retrieve the pixel-based forest canopy structural variables CC and CD. No matter which endmember extraction method is used, the model predictions of CC (figures 7(a), (c) and (e)) are more reliable than those for CD. The values of CD are overestimated by the model, yielding less accurate results (figures 7(b), (d) and (f)). Better results for CC

### Table 4. Coefficient of determination ($R^2$) for linear regression and root mean square error (RMSE) between the model-estimated CC/CD and the ground measured CC/CD.

<table>
<thead>
<tr>
<th>Model estimated CC/CD by 3 methods</th>
<th>$R^2$</th>
<th>RMSE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>40 samples</td>
<td>39 samples</td>
</tr>
<tr>
<td>CC (PPI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CC (pixel scaling)</td>
<td>0.66</td>
<td>0.60</td>
</tr>
<tr>
<td>CC (regional unmixing)</td>
<td>0.62</td>
<td>0.56</td>
</tr>
<tr>
<td>CD (PPI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CD (pixel scaling)</td>
<td>0.39</td>
<td>0.38</td>
</tr>
<tr>
<td>CD (regional unmixing)</td>
<td>0.40</td>
<td>0.40</td>
</tr>
</tbody>
</table>

### Table 5. Comparison of the three endmember extraction methods for deriving CC and CD by the inverted Li–Strahler model using the $t$-statistic for 38 samples ($n=38$) at 0.01 level of significance ($\alpha=0.01$).

<table>
<thead>
<tr>
<th>Model estimation</th>
<th>PPI method</th>
<th>Purest-pixel selection</th>
<th>Regional unmixing</th>
<th>Critical $t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC</td>
<td>3.215</td>
<td>3.222</td>
<td>0.301</td>
<td>2.715</td>
</tr>
<tr>
<td>CD</td>
<td>4.461</td>
<td>3.288</td>
<td>4.056</td>
<td>2.715</td>
</tr>
</tbody>
</table>
estimation compared to CD from the Li–Strahler inverted model were also observed by other researchers (Franklin and Strahler 1988, Woodcock et al. 1997). The information in the remote sensing image (reflectance spectra) is more related to the CC variable in the forest region than to the CD variable. It is difficult to find a reliable relationship between image intra-stand variance and tree size (CD). Therefore, CC estimation is more sensitive to the geometric-optical model of Li–Strahler, and it can be estimated more accurately than CD.

The mapping results of CC and CD are directly related to the value of $K_g$, the abundance fraction of sunlit background. According to the principle of the Li–Strahler model, the surface consists of three components $G$, $C$ and $T$. Regions, with no or even a calculated negative fraction of $G$ indicate that, in these regions, the surface is just composed of sunlit canopy and shadow, and there are no estimates of CC and CD from the inverted model. Consequently, for very dense forest, it is difficult to derive the proportion of sunlit background and the inverted Li–Strahler model is not able to map CC and CD. Although the assumption of the Li–Strahler model states that every pixel of the image should contain several crowns, the model can not work either if the number of crowns is too much. That is also one of the reasons for some infeasible areas predicted by the inverted Li–Strahler model.

Determining an accurate abundance of $K_g$ depends on extracting suitable endmembers of $G$, $C$ and $T$ for the image. The PPI method is an image-based endmember extraction approach. Extreme pixels are the primary outputs from this method. However, no criteria are given on how to select the parameter $k$ and the threshold $l$, which determine the number of extreme pixels. Those need to be detected by repeated trainings for different images. Another problem of the PPI method is the purity of the extracted extreme pixel. The results of this study (tables 1 and 2) show that the PPI endmembers exhibit errors because the extreme pixels are still mixed. Comparing figure 7(a) and (c), it is illustrated that two sample sites, which have a ground-measured CC of 0.9 and 0.95, cannot be mapped by the PPI endmembers. This indicates that the extracted endmembers from the PPI method are not suitable for this Hyperion image.

The method of pixel scaling-based endmember extraction determines the purest pixel directly from the Hyperion image. The endmembers come from the reflectance spectra of single pixels. Therefore, the mapping results of this method produce the least infeasible areas (figures 6(c)–(d)) and all 40 sample sites can be predicted by the inverted model (figures 7(c)–(d)). Nevertheless, there are several limitations for applying this method. Firstly, for each extreme pixel, the corresponding high spatial resolution data should be available. Secondly, the spatial resolution of the image should allow the possibility of finding a real pure pixel to represent each endmember. Thirdly, the certainty of the classification of the high spatial resolution image and the quality of the co-registration between the up-scaled calculated proportions and the extreme pixels mainly influence the accuracy of the purest pixel selection.

The scaling-based endmember extraction method using the linear unmixing model can be used to solve the mixture problem of the medium or low spatial resolution data and automatically derive the endmembers, when a matched high spatial resolution image is available. Because the extracted endmembers are the aggregated results from the total overlapping region of the QuickBird and the Hyperion data, the endmembers cannot match the Hyperion image like the real spectra. So, the output images of figures 6(e)–(f) show more infeasible areas than figures 6(c)–(d) and one sample site with a value of 0.95 for CC is missed (figures 7(c) and (e)). However,
the $t$ value (table 5) shows that it is the best performing method for deriving CC. In addition, since the QuickBird image of this case is smaller than the Hyperion image, when the overlapping region is including all scene components and it is located in a representative area of the Hyperion image, then the calculated regional scaling-based endmembers can be used for the whole Hyperion image. It indicates that the method of linear unmixing is appropriate for up-scaling the information from high spatial resolution data, and it can also expand the information from local to regional scale.

5. Conclusions

In this study, we use an inverted Li–Strahler geometric-optical model to derive the forest canopy structural variables CC and CD from a hyperspectral EO-1 Hyperion image collected in the Longmenhe broadleaved forest natural reserve, located in the Three Gorges region of China. The accuracy of this inverted model is directly related to the quality of image endmembers of three scene components, namely sunlit background ($G$), sunlit canopy ($C$) and shadow ($T$). For evaluating, and possibly improving, the method of endmember extraction, we compare three approaches: the image-based PPI method, the pixel scaling-based method of purest-pixel selection and the regional scaling-based linear unmixing model. After validating the inverted model mapping results of CC and CD by ground measurements, it can be concluded that the regional scaling-based endmembers derived from the linear unmixing model are the best ones to be used in combination with the inverted Li–Strahler model, and these scaling-based aggregated endmembers improve the Li–Strahler model inversion for monitoring the forest canopy structure.

Scales and scaling currently are one of the central issues in quantitative remote sensing (Liang 2004, Schaepman 2007). Although the scaling-based endmember extraction method using the linear unmixing model is applied using QuickBird as high spatial resolution images and Hyperion as medium spatial resolution images in this study, it can be implemented for any two sets of remote sensing images with different spatial and spectral resolution. Potential combinations may include scaling from QuickBird to Landsat TM, from Landsat TM to MODIS, and even from Hyperion to MODIS. Therefore, using the presented up-scaling-based method to explore the information from images at larger scale seems to be feasible.

In future work, we will continue quantifying forest structural variables in the whole Three Gorges region of China using multi-spatial, multi-temporal and multi-spectral resolution data through this inverted Li–Strahler geometric-optical model and the described up-scaling method. Even though the inverted model and the scaling method will need further analysis and calibration, the presented results show confidence in the approach selected. Moreover, combining physical-based canopy reflectance models used for quantitatively monitoring the change of the forest structure over time in combination with statistical models will allow assessing ecosystem disturbance over extended periods of time.

Acknowledgements

We acknowledge the support from the ‘Knowledge Innovation Program of the Chinese Academy of Sciences’ (no. KZCX1-YW-08-02-03 and no. KZCX3-SW-334). We appreciate Xu Wenting, Huang Jianxi, Tian Yichen and the College of Life Sciences, Wuhan University for participation in the field campaign. Pepijn van Oort is thanked for statistical support. We also thank the anonymous reviewers for constructive comments on our manuscript.
References


McKnight, T.L. and Hess, D., 2000, *Climate Zones and Types: Dry Humid Subtropical Climate (Cfa, Cwa)* (New York: Prentice Hall).


