Estimating Forest fAPAR from Multispectral Landsat-8 Data Using the Invertible Forest Reflectance Model INFORM

Huili Yuan 1, Ronghua Ma 1,*, Clement Atzberger 2, Fei Li 1, Steven Arthur Loiselle 3 and Juhua Luo 1

1 Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences, Nanjing 210008, China; E-Mails: yuanhuili.student@sina.com (H.Y.); fli@niglas.ac.cn (F.L.); jhluo@niglas.ac.cn (J.L.)
2 Institute of Surveying, Remote Sensing & Land Information (IVFL), University of Natural Resources and Life Science (BOKU), Vienna 1190, Austria; E-Mail: element.atzberger@boku.ac.at
3 Dipartimento di Biotecnologia, Chimica e Farmacia, CSGI, Via Aldo Moro 2, University of Siena, Siena 53100, Italy; E-Mail: loiselle@unisi.it

* Author to whom correspondence should be addressed; E-Mail: rhma@niglas.ac.cn; Tel.: +86-25-8688-2168.

Academic Editors: Randolph H. Wynne and Prasad S. Thenkabail

Received: 5 April 2015 / Accepted: 29 May 2015 / Published: 5 June 2015

Abstract: The estimation of the Fraction of Absorbed Photosynthetically Active Radiation in forests (forest fAPAR) from multi-spectral Landsat-8 data is investigated in this paper using a physically based radiative transfer model (Invertible Forest Reflectance Model, INFORM) combined with an inversion strategy based on artificial neural nets (ANN). To derive the forest fAPAR for the Dabie mountain test site in China in 30 m spatial resolution (size approximately 3000 km²), a database of forest canopy spectral reflectances was simulated with INFORM taking into account structural variables such as leaf area index (LAI), crown coverage and stem density as well as leaf composition. To establish the relationship between forest fAPAR and the reflectance modeled by INFORM, a logarithmic relationship between LAI and fAPAR was used previously established using on-site field measurements. On this basis, predictive models between Landsat-8 reflectance and fAPAR were established using an artificial neural network. After calibrating INFORM for the test site, forty-two forest stands were used to validate the performance of the method. The results show that spectral signatures modeled by INFORM correspond reasonably well with the forest canopy reflectance spectra derived from Landsat data. Deviations increase with increasing angle between surface normal of the hilly terrain and
Remote Sens. 2015, 7

The comparison of estimated and measured fAPAR \( R^2 = 0.47, \text{RMSE} = 0.11 \) demonstrates that INFORM can be inverted using neural nets to provide acceptable estimates of forest fAPAR. The accuracy of the predictions increased significantly when excluding pixels located in very steep terrain. This demonstrates that the applied topographic correction was not sufficiently accurate and should be improved for making optimum use of radiative transfer models such as INFORM.

**Keywords:** INFORM; model inversion; neural network; fAPAR; LAI; Landsat-8; OLI

1. Introduction

The estimation of the Fraction of Absorbed Photosynthetically Active Radiation (fAPAR) provides key information for estimating vegetation net primary productivity (NPP) [1]. Consequently, fAPAR has been identified by the Global Climate Observing System (GCOS) as one of the Essential Climate Variables (ECVs). Forest ecosystems play a critical role in the global carbon cycle but the accurate retrieval of forest fAPAR is hindered by the complex interaction of electromagnetic radiation with forest canopies [2]. Additional problems arise in hilly terrain.

Most studies use empirical models to estimate forest fAPAR [3]. The functional relationships (linear or nonlinear) between satellite-derived vegetation indices and field-measured fAPAR established by regression analysis are easy to understand and computationally efficient to operate over large scales. They also constitute an efficient means to generate reference information from field samples useful for testing alternative approaches [4,5]. Empirical models, however, do not exploit the complex transfer of solar radiation within the forest stands and can only be applied to the conditions where they were developed.

A potentially more robust approach consists in using radiative transfer models. Physically-based models can handle many different types of conditions and account for the effect of biochemical, structural or observational variables on canopy reflectance [6]. After model construction and calibration, it is possible to invert canopy reflectance models (CRM) to retrieve variables of interest from the measured reflectance. Canopy reflectance models act thus as the interface between the remotely sensed images and forest variables of interest. A prerequisite for using CRMs is that the satellite observations have to be corrected for atmospheric and terrain effects so that the resulting image corresponds to the Bottom of Atmosphere (BoA) reflectance [7–9].

Since the 1970s, a variety of CRMs have been developed to simulate forest canopy reflectance. In early CRMs such as SUITS (a radiative transfer model named by the developer’ name) [10] and SAIL(Scattering by Arbitrarily Inclined Leaves) [11], vegetation canopies have been modeled assuming a homogenous canopy layer composed of infinite small scattering elements (the leaves). The models do not mimic the influence of clumpiness, open spots, and shadowing effects, but assume a homogenous distribution of scattering and absorbing particles (e.g., leaves) in the canopy. This over-simplifies forest canopies, where leaves are generally not homogeneously distributed, and where shadows play an important role in the scene reflectance [12,13]. The simplification of canopy layer in the mentioned models also neglects the effects of understory which is known to have a strong influence on canopy
reflectance [14]. Sometimes the sensitivity of some bands onto the overstory leaf area index (LAI) even disappears when the understory reaches an LAI higher than one [15].

To deal with shadowing effects resulting from tree crowns, geometrical-optical models such as the Li-Strahler model [16] describe the forest as a composite of opaque geometrical shapes. These models, however, mostly ignore the non-zero transmissivity of tree crowns.

In recent years, the spectral invariants theory ($p$-theory) [17] was applied in forests and a PARAS (a simple parameterization model using photon recollision probability, a spectral invariant which describes the structure of a canopy, for calculating forest stand reflectance.) model family was developed [18–21]. PARAS models require only a limited set of input parameters and the canopy structure is described using only one parameter, the photon recollision probability $p$. Therefore, the PARAS models are also called semi-physical forest reflectance models, neglecting crown level clumping and the associated shading patterns on the background, as well as multiple interactions of photons between the tree layer and the understory layer [20].

Among existing CRMs for forests (for an overview see [22]), INFORM [23,24] provides a simple while still realistic description of the interaction of the electromagnetic radiation with forest canopies. This computationally simple model takes into account the effects of shadowing and crown transmittance as well as the effects of leaf geometry and leaf optical properties on the crown transmittance. The impact of leaf geometry and leaf optical properties were neglected by the forest-light interaction model (FLIM) [25] on which INFORM builds.

Whatever CRM is used, it is worth noting that fAPAR usually cannot be derived directly by inverting CRMs as fAPAR is usually not a build-in variable of CRMs. The common way to retrieve fAPAR from the modeled reflectance is utilizing the modeled spectral reflectance or its derivatives (e.g., VIs) to establish a linear or non-linear relationship with fAPAR [14,26,27]. This leads to so-called semi-empirical models. A close relationship between canopy reflectance and fAPAR can be expected because both are functionally linked to leaf area index (LAI) [28]. For example, based on the continuous observations of fAPAR and LAI during the growing period, linear or logarithmic relationships were identified between the two biophysical variables [29–32]. Equally well known is also the fact that LAI has a close relationship with the reflectance modeled from CRM, which is the foundation for estimates of LAI by inverting the CRM. Therefore, the relationship between LAI and fAPAR, as well as between LAI and canopy reflectance, presents a possible link between the modeled canopy reflectance and fAPAR. This link can be exploited using various techniques such as look-up-tables (LUT) and numerical optimization procedures [33–35]. Compared with numerical optimization techniques and LUT, a well calibrated artificial neural network (ANN) can be applied to estimate variables with higher computational efficiency and more accurate approximations of complex non-linear functions [36].

In this context, this paper will investigate the estimation of forest fAPAR from multispectral Landsat-8 data using the INFORM model. Important sub-tasks include:

(a) to calibrate the INFORM parameters for adapting the CRM to local conditions,
(b) to evaluate its performance in modeling the forest canopy reflectance by sensitivity analysis and comparison with observed spectral signatures, and
(c) to estimate and evaluate forest fAPAR by inverting INFORM based on ANN.

As the study is conducted in a hilly environment, special attention will be given to possible terrain effects.
2. Study Site and Data

2.1. Study Area and Field Measurements

The Dabie study area (115°24′–116°3′E, 31°13′–31°44′N) is located on the northern slope of the Dabie mountain with steep topography and complex terrain. Dabie is located in southern China and covers an area of about 3814 km², including a forest area of 2940 km². Twenty-one relatively homogenous broad-leaved forest stands with poplar, chestnut and beech as main species were identified, located between 125 m to 1210 m above mean sea level (Figure 1). Within each of these stands, two quadratic plots of 30 m length were established leading to a total number of forty-two reference plots.

![Figure 1](image)

Figure 1. Location of the study area in China (Left) and the field-measured sampling plots (red dots) shown in a digital elevation model (DEM) (Right). At each of the twenty-one field sampling plots, two reference plots were established. The total number of reference plots thus equals 42.

Terrain effects represent one of the most important perturbations in remote sensing over mountainous terrain [37]. To study terrain effects, several reference plots with low cosine of the incidence angle (angle $i$ between the direction of the sun and the local surface normal) [38] were chosen. Information regarding elevation, slope, aspect and $\cos i$ is reported in Table 1 for each of the 42 reference plots. As shown in Figure 2, according to the location of the reference plots and the sun, two groups are distinguished: (1) plots with ‘high’ $\cos i > 0.51$ ($n = 33$; shown in black), and (2) plots with “low” $\cos i \leq 0.51$ ($n = 9$; shown in bold).

The central location of each reference plot was determined with an accuracy of about ±5 m using a differential GPS device (Trimble Pro XR). At each plot, measurements of forest structural variables were conducted in October 2012 including canopy fAPAR, LAI, stem density (SD), crown diameter (CD) and stand height (H) (Table 2).
Table 1. Information regarding elevation, slope and aspect of the 42 individual field plots. The variable $\cos \theta$ is the angle between the direction of the sun and the local surface normal. The 42 reference plots are split into two groups: (1) reference plots with high $\cos \theta$ (e.g., $\cos \theta > 0.51$) in black, and (2) reference plots with low $\cos \theta$ (e.g., $\cos \theta \leq 0.51$) in bold. The latter reference plots represent non-favorable observations conditions (orientation of the surface) and will therefore be analyzed separately.

<table>
<thead>
<tr>
<th>ID</th>
<th>Elevation</th>
<th>Slope</th>
<th>Aspect</th>
<th>$\cos \theta$</th>
<th>ID</th>
<th>Elevation</th>
<th>Slope</th>
<th>Aspect</th>
<th>$\cos \theta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>149</td>
<td>39.26</td>
<td>293.43</td>
<td>0.1237</td>
<td>22</td>
<td>149</td>
<td>17.46</td>
<td>238</td>
<td>0.6295</td>
</tr>
<tr>
<td>2</td>
<td>124</td>
<td>28.4</td>
<td>294.59</td>
<td>0.2707</td>
<td>23</td>
<td>117</td>
<td>3.81</td>
<td>90</td>
<td>0.6331</td>
</tr>
<tr>
<td>3</td>
<td>25</td>
<td>15.95</td>
<td>329.32</td>
<td>0.3776</td>
<td>24</td>
<td>117</td>
<td>3.81</td>
<td>90</td>
<td>0.6331</td>
</tr>
<tr>
<td>4</td>
<td>38</td>
<td>15.05</td>
<td>343.81</td>
<td>0.3902</td>
<td>25</td>
<td>162</td>
<td>2.26</td>
<td>108.44</td>
<td>0.6347</td>
</tr>
<tr>
<td>5</td>
<td>195</td>
<td>12.26</td>
<td>355.6</td>
<td>0.4412</td>
<td>26</td>
<td>136</td>
<td>4.39</td>
<td>220.6</td>
<td>0.6414</td>
</tr>
<tr>
<td>6</td>
<td>68</td>
<td>18.91</td>
<td>41.05</td>
<td>0.4650</td>
<td>27</td>
<td>136</td>
<td>4.17</td>
<td>210.96</td>
<td>0.6483</td>
</tr>
<tr>
<td>7</td>
<td>22</td>
<td>10.64</td>
<td>356.18</td>
<td>0.4659</td>
<td>28</td>
<td>165</td>
<td>3.44</td>
<td>123.69</td>
<td>0.6529</td>
</tr>
<tr>
<td>8</td>
<td>27</td>
<td>8.71</td>
<td>337.62</td>
<td>0.4887</td>
<td>29</td>
<td>27</td>
<td>3.72</td>
<td>140.19</td>
<td>0.6624</td>
</tr>
<tr>
<td>9</td>
<td>59</td>
<td>17.37</td>
<td>48.24</td>
<td>0.5064</td>
<td>30</td>
<td>57</td>
<td>20.26</td>
<td>208.3</td>
<td>0.7528</td>
</tr>
<tr>
<td>10</td>
<td>158</td>
<td>15.8</td>
<td>260.68</td>
<td>0.5463</td>
<td>31</td>
<td>131</td>
<td>16.44</td>
<td>198.95</td>
<td>0.7594</td>
</tr>
<tr>
<td>11</td>
<td>151</td>
<td>13.64</td>
<td>254.06</td>
<td>0.5795</td>
<td>32</td>
<td>168</td>
<td>25.65</td>
<td>105.6</td>
<td>0.7611</td>
</tr>
<tr>
<td>12</td>
<td>66</td>
<td>9.78</td>
<td>60.89</td>
<td>0.5889</td>
<td>33</td>
<td>40</td>
<td>20.02</td>
<td>202.17</td>
<td>0.7727</td>
</tr>
<tr>
<td>13</td>
<td>51</td>
<td>15.48</td>
<td>248.84</td>
<td>0.5912</td>
<td>34</td>
<td>53</td>
<td>19.19</td>
<td>196.7</td>
<td>0.7842</td>
</tr>
<tr>
<td>14</td>
<td>56</td>
<td>7.58</td>
<td>58.09</td>
<td>0.5913</td>
<td>35</td>
<td>106</td>
<td>21.62</td>
<td>123.86</td>
<td>0.8112</td>
</tr>
<tr>
<td>15</td>
<td>137</td>
<td>13.59</td>
<td>250</td>
<td>0.5927</td>
<td>36</td>
<td>115</td>
<td>21.58</td>
<td>191.55</td>
<td>0.8146</td>
</tr>
<tr>
<td>16</td>
<td>163</td>
<td>2.49</td>
<td>287.18</td>
<td>0.5931</td>
<td>37</td>
<td>62</td>
<td>25.65</td>
<td>195.6</td>
<td>0.8262</td>
</tr>
<tr>
<td>17</td>
<td>146</td>
<td>0.75</td>
<td>251.57</td>
<td>0.6145</td>
<td>38</td>
<td>34</td>
<td>26.84</td>
<td>120.16</td>
<td>0.8286</td>
</tr>
<tr>
<td>18</td>
<td>146</td>
<td>0.75</td>
<td>251.57</td>
<td>0.6145</td>
<td>39</td>
<td>69</td>
<td>24.83</td>
<td>128.42</td>
<td>0.8451</td>
</tr>
<tr>
<td>19</td>
<td>146</td>
<td>0.75</td>
<td>251.57</td>
<td>0.6145</td>
<td>40</td>
<td>98</td>
<td>33.6</td>
<td>196.39</td>
<td>0.8560</td>
</tr>
<tr>
<td>20</td>
<td>146</td>
<td>0.75</td>
<td>251.57</td>
<td>0.6145</td>
<td>41</td>
<td>121</td>
<td>26.06</td>
<td>156.93</td>
<td>0.8988</td>
</tr>
<tr>
<td>21</td>
<td>7</td>
<td>0.95</td>
<td>90</td>
<td>0.6198</td>
<td>42</td>
<td>145</td>
<td>30.82</td>
<td>151.65</td>
<td>0.9293</td>
</tr>
</tbody>
</table>

Figure 2. Location relationship between the reference plots and the sun.
On clear days, fAPAR measurements were taken using a Tracing Radiation and Architecture of Canopies (TRAC) instrument which was developed at Canada Centre for Remote Sensing [39]. A fish-eye lens instrument was used under overcast condition (Regent Instruments Inc., Tokyo, Japan). Compared to TRAC showing fAPAR directly, the fish-eye lens provided images requiring further processing by CAN-EYE software [40]. LAI was estimated using a Li-Cor LAI-2200 Plant Canopy Analyzer under overcast sky conditions (Lincoln Inc., Nebraska)[41]. As shown in Figure 3, the red dots are the locations of LAI-2200 measurements and TRAC measurements, and the box respects one plot of 30 m length, which corresponds to one Landsat pixel. LAI and fAPAR measurements were taken at eight regularly spaced points within each plot, from which the average was calculated. Stem density (SD) was obtained by counting the number of trees in a plot. Stand height (H) was calculated from the mean height of all trees. The height of each tree (from the ground to the top) was estimated from angular measurements. Crown diameter (CD) was the average of the crown diameter of the trees. Stem density, stand height and crown diameter were all measured from the trees whose diameter at breast height (DBH) were greater than 5 cm within a plot. In a forest stand, the average soil reflectance was determined from three measurements using a FieldSpec 4 Hi-Res spectroradiometer (Analytical Spectra Devices., Inc., Boulder, CO, USA). All canopy variables were measured between 10:00 and 16:00 to minimize sun angle effects.

### Table 2. Summary statistics for field measured fraction of absorbed photosynthetically active radiation in forests (fAPAR) and forest stand variables (n = 42).

<table>
<thead>
<tr>
<th>Variables</th>
<th>Mean</th>
<th>Std. dev.</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>fAPAR</td>
<td>0.7944</td>
<td>0.1035</td>
<td>0.5754</td>
<td>0.9370</td>
</tr>
<tr>
<td>LAI</td>
<td>3.54</td>
<td>1.59</td>
<td>1.09</td>
<td>6.37</td>
</tr>
<tr>
<td>SD [ha⁻¹]</td>
<td>1695</td>
<td>1767</td>
<td>244</td>
<td>5155</td>
</tr>
<tr>
<td>CD [m]</td>
<td>5.16</td>
<td>1.46</td>
<td>1.20</td>
<td>5.64</td>
</tr>
<tr>
<td>H [m]</td>
<td>10.19</td>
<td>3.05</td>
<td>5.50</td>
<td>13.31</td>
</tr>
</tbody>
</table>

**Figure 3.** Sampling scheme for TRAC transects and LAI-2200 measurement points. The distance between two LAI measurement points was ten meters in each cardinal direction.
2.2. Image Data and Image Pre-Processing

For modelling and mapping LAI, a Landsat-8 image acquired on 12 October 2013 was used. The Landsat-8 image was radiometrically corrected by utilizing the Radiometric Calibration module and the FLAASH Atmospheric Correction module integrated in ENVI 5.0 with the spectral response function [9].

As most of the existing CRMs are not suitable for areas with steep topography and complex terrains, the Landsat-8 image was topographically corrected with the C-correction model [42,43] (Figure 4). Topographic corrections with this simple model remove major illumination induced brightness changes. However, residual effects are to be expected [38,44] and were evident in our data set (Figure 5). No further attempts were made to remove the remaining terrain effects, as such data quality is to be expected in typical remote sensing application studies.

![Figure 4. False color composite image (bands 4-3-2) of the study area after atmospheric and terrain related preprocessing.](image)

**Figure 4.** False color composite image (bands 4-3-2) of the study area after atmospheric and terrain related preprocessing.

![Figure 5. Observed impact of $\cos i$ (local terrain orientation with respect to the sun) on Landsat-8 reflectance in the nIR band. The linear regression line between near-infrared band reflectance and $\cos i$ ($n = 42$) is also shown.](image)

**Figure 5.** Observed impact of $\cos i$ (local terrain orientation with respect to the sun) on Landsat-8 reflectance in the nIR band. The linear regression line between near-infrared band reflectance and $\cos i$ ($n = 42$) is also shown.
3. Method

The estimation of forest fAPAR from multispectral remote sensing data and the INFORM canopy reflectance model consisted of five steps (Figure 6):

(a) establishing an empirical relationships between field measured LAI and in situ fAPAR,
(b) modeling visible and near-infrared bi-directional spectral reflectance of forest canopies with variable structure, leaf bio-chemistry and canopy background reflectance using INFORM,
(c) using INFORM to generate a synthetic fAPAR dataset through application of the empirical relationship between LAI and fAPAR to the LAI values used as input to the synthetic reflectance data base,
(d) determining the predictive relationship between spectral reflectance and fAPAR of the synthetic data base based on back-propagation (BP) neural networks, and
(e) applying the trained ANN to remote sensing data and analyzing the simulated results.

**Figure 6.** General procedure for estimating forest fAPAR from Landsat data using a forest reflectance model.

3.1. INFORM

INFORM was used to simulate the canopy reflectance of forest stands between 400 and 2500 nm [23,24]. The model is based on FLIM [25], combined with SAILH (SAIL model with hot) [11,45], and PROSPECT (a radiative transfer model based on Alle’s generalized “plate model” that represents the optical properties of plant leaves from 400 nm to 2500 nm) [46].

In INFORM, a forest is considered as a discontinuous canopy layer with tree crowns and gaps [25]. The average crown transmittance in sun direction and in observation direction ($T_s$ and $T_o$) is calculated from SAIL. The required leaf transmittance and reflectance ($\tau$ and $\rho$) are calculated from PROSPECT. The bi-directional reflectance of the forest canopy in a given spectral band is expressed as:

\[
\rho_{\lambda} = f_{\text{INFORM}}(\text{LAI}_0, SD, CD, ...) 
\]
Remote Sens. 2015, 7

\[ R = R_C \cdot C + R_G \cdot G \]  \hspace{1cm} (1)

where \( R_C \) is the crown reflectance at infinite crown depth and \( R_G \) is the background reflectance. \( C \) is the “crown factor”:

\[ C = (1 - T_s \cdot T_o) \cdot c_s \cdot c_o \]  \hspace{1cm} (2)

where \( T_s \) and \( T_o \) are the average crown transmittance in the sun direction and in observation direction, respectively; \( c_o \) and \( c_s \) are the ground coverage by crowns in observation direction and by shadow in sun direction, respectively.

\( G \) is the “ground factor”:

\[ G = F_{cd} \cdot T_s \cdot T_o + F_{cs} \cdot T_o + F_{od} \cdot T_s + F_{os} \]  \hspace{1cm} (3)

where \( F_{cd}, F_{cs}, F_{od} \) and \( F_{os} \) are the ground fractions, representing crowns with shadowed background, crowns with sunlit background, shadowed open space and sunlit open space respectively. The four fractions are calculated as follows:

\[ F_{cd} = c_o c_s + \text{rho}[c_o(1-c_o)c_s(1-c_s)]^{0.5} \]

\[ F_{cs} = c_o(1-c_o) - \text{rho}[c_o(1-c_o)c_s(1-c_s)]^{0.5} \]  \hspace{1cm} (4)

\[ F_{od} = c_s(1-c_o) - \text{rho}[c_o(1-c_o)c_s(1-c_s)]^{0.5} \]

\[ F_{os} = (1-c_o)(1-c_s) - \text{rho}[c_o(1-c_o)c_s(1-c_s)]^{0.5} \]

where \( \text{rho} \) represents the positive correlation between \( c_o \) and \( c_s \) as:

\[ \text{rho} = e^{(-\frac{\theta_o}{\Psi})} \]  \hspace{1cm} (5)

\[ g = (\tan^2 \theta_o + \tan^2 \theta_s - 2 \tan \theta_o \tan \theta_s \cos \Psi)^{0.5} \]

where \( g \) is a geometrical factor depending on the illumination and observing geometry, \( \theta_o \) and \( \theta_s \) are observation and solar zenith angles, and \( \Psi \) is the relative azimuth between sun and observer. \( H \) is the tree height. \( CD \) is the tree crown diameter.

To derive \( C \) and \( G \), several variables, such as \( c_o \) and \( c_s \), are defined.

\[ c_o = 1 - e^{-\frac{k \cdot S0}{\cos \theta_o}} \]  \hspace{1cm} (6)

\[ c_s = 1 - e^{-\frac{k \cdot S0}{\cos \theta_s}} \]  \hspace{1cm} (7)

\[ k = \frac{\pi \cdot (0.5 \cdot CD)^2}{10000} \]  \hspace{1cm} (5)

where \( SD \) is the stem density and \( k \) is the average crown horizontal area depending on \( CD \).

Regarding the average crown transmittance, the SAILH model is used to estimate \( T_s \) and \( T_o \) with the leaf area index (LAI), average leaf inclination angle (ALA), leaf reflectance \( (\rho_{leaf}) \), leaf transmittance
(\tau_{\text{leaf}}), observation zenith angle (\theta_o), solar zenith angle (\theta_s), the relative azimuth angle between sun and observer (\psi), the fraction of diffuse radiation (skyl) and the hot spot (hot) [45] as input parameters.

\[ T_0 = f_{\text{SAIL}}(\text{LAI}, \text{ALA}, \tau_{\text{leaf}}, \rho_{\text{leaf}}, \theta_o, \psi, \text{skyl}, \text{hot}) \]  
\[ T'_0 = f_{\text{SAIL}}(\text{LAI}, \text{ALA}, \tau_{\text{leaf}}, \rho_{\text{leaf}}, \theta_o, \psi, \text{skyl}, \text{hot}) \]

In addition, \( R_G \) and \( R_C \) are also derived from the SAILH model.

\[ R_G = f_{\text{SAIL}}(\text{LAI}_U, \text{ALA}_U, \tau_{\text{leaf}}, \rho_{\text{leaf}}, \rho_{\text{soil}}, \theta_o, \theta_s, \psi, \text{skyl}, \text{hot}) \]  
\[ R_C = f_{\text{SAIL}}(\text{LAI}_{\text{inf}}, \text{ALA}, \tau_{\text{leaf}}, \rho_{\text{leaf}}, \rho_{\text{soil}}, \theta_o, \theta_s, \psi, \text{skyl}, \text{hot}) \]

where \( \text{LAI}_U \) and \( \text{ALA}_U \) are \( \text{LAI} \) and \( \text{ALA} \) of the understory, \( \text{LAI}_{\text{inf}} \) is the LAI at infinite crown depth and \( \rho_{\text{soil}} \) is the soil reflectance (from field measurements). The leaf transmittance and reflectance (\( \tau_{\text{leaf}} \) and \( \rho_{\text{leaf}} \)) in Equation (12) are calculated using the PROSPECT model [46] with the leaf structure parameter \( (N) \), chlorophyll content \( (a+b) \) \( (C_{ab}) \), cellulose and lignin content \( (C_m) \) and equivalent water thickness \( (C_w) \) as input parameters.

\[ \tau_{\text{leaf}} = f_{\text{prospect}}(N, C_{ab}, C_m, C_w) \]  
\[ \rho_{\text{leaf}} = f_{\text{prospect}}(N, C_{ab}, C_m, C_w) \]

As shown in Table 3, INFORM is parameterized based on twelve synthetic variables. Consequently, the reflectance of forest stands canopy can be expressed as:

\[ R = f_{\text{INFORM}}(\text{SD}, \text{CD}, \text{H}, \text{LAI}, \text{LAI}_{\text{inf}}, \text{ALA}_U, \text{ALA}_U, \rho_{\text{soil}}, N, C_{ab}, C_m, \theta_o, \theta_s, \psi, \text{skyl}, \text{hot}) \]  

**Table 3.** Summary of main invertible forest reflectance model (INFORM) parameters. Variable meaning is given in the main text.
Except the forest stand variables (SD, CD, LAI and H, Table 2), the remaining input parameters (geometrical information and the leaf-level parameters) were fixed (Table 4). The geometrical information (observation and sun zenith angles; azimuth angle) was taken from the corresponding Landsat-8 orbit characteristics, location and date. The leaf-level parameters including leaf chlorophyll content \((a+b)\), dry matter content and equivalent water thickness were taken from LOPEX’93 experiment [47] representing poplar, chestnut and beech leaves.

### Table 4. INFORM parameters kept constant during model simulations.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Designation</th>
<th>Unit</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sun zenith angle</td>
<td>(\theta_s)</td>
<td>deg</td>
<td>42.6133</td>
</tr>
<tr>
<td>Observation zenith angle</td>
<td>(\theta_o)</td>
<td>deg</td>
<td>0</td>
</tr>
<tr>
<td>Azimuth angle</td>
<td>(\Psi)</td>
<td>deg</td>
<td>180</td>
</tr>
<tr>
<td>Fraction of diffuse radiation</td>
<td>skyl</td>
<td>fraction</td>
<td>0.1</td>
</tr>
<tr>
<td>Hot spot parameter</td>
<td>hot</td>
<td>ratio</td>
<td>1.4</td>
</tr>
<tr>
<td>Average leaf angle of tree canopy</td>
<td>ALA</td>
<td>deg</td>
<td>55</td>
</tr>
<tr>
<td>Leaf area index at infinite crown depth</td>
<td>LAI(_{inf})</td>
<td>m(^2)·m(^{-2})</td>
<td>15</td>
</tr>
<tr>
<td>Leaf area index of understory</td>
<td>LAI(_U)</td>
<td>m(^2)·m(^{-2})</td>
<td>0.5</td>
</tr>
<tr>
<td>Average leaf angle of understory</td>
<td>ALA(_U)</td>
<td>deg</td>
<td>45</td>
</tr>
<tr>
<td>Chlorophyll content ((a+b))</td>
<td>Cab</td>
<td>(\mu g)·cm(^{-2})</td>
<td>44</td>
</tr>
<tr>
<td>Cellulose and lignin content</td>
<td>Cm</td>
<td>g·cm(^{-2})</td>
<td>0.003493</td>
</tr>
<tr>
<td>Equivalent water thickness</td>
<td>Cw</td>
<td>g·cm(^{-2})</td>
<td>0.009</td>
</tr>
<tr>
<td>Mesophyll structure parameter</td>
<td>N</td>
<td>/</td>
<td>1.7</td>
</tr>
</tbody>
</table>

During the processing of the field-measured data, it was found that stem density (SD) and crown diameter (CD) were closely related. This permitted the constraining of the INFORM parameterization by using their fitted relation: \(SD = 8388CD^{1.486}\) \((R^2 of 0.68, n = 42)\) (Figure 7).

![Figure 7. Empirical relationship between crown diameter and stem density used in INFORM to link stem density (SD) to crown size (CD) \((n = 42)\).](image)

### 3.2. Training the Artificial Neural Network

To invert the CRM, Artificial Neural Networks (ANN) were employed. ANN “learn” the relation between a set of input variables (here the INFORM simulated forest reflectances) and the target variable
(here: fAPAR) [48]. In the present work, a simple three layer feed-forward backpropagation network was used with tan-sigmoidal transfer functions in hidden and output layers. The ANN was trained using a training database generated from INFORM. To assess forest fAPAR, a previously established relation between LAI and fAPAR was used (see next section and Figure 8). For successful training, the distribution of inputs and outputs was important as well as a compact network architecture and efficient training algorithms.

3.2.1. Generating the Training Database

The training database consisted of the input variables (*i.e.*, the modeled spectra from INFORM) and the corresponding target variables (*i.e.*, the canopy structural and bio-chemical variables used to simulate the spectra and with fAPAR linked to LAI).

To ensure that the training data was consistent with the real Landsat-8 data, the continuous reflectance data modeled from CRM were spectrally resampled to multispectral wavebands using the appropriate spectral response functions, following:

\[
\rho = \frac{\int_{t_1}^{t_2} \rho_s(\lambda) \cdot B(\lambda) \, d\lambda}{\int_{t_1}^{t_2} B(\lambda) \, d\lambda}
\]  

where \(\rho\) is the resampled reflectance, \(\rho_s\) is the modeled reflectance, and \(B(\lambda)\) is the spectral response function. To permit using alternative satellite sensors, only the first four Landsat channels were simulated: blue (band 1), green (band 2), red (band 3) and near infrared (band 4).

Several studies found that fAPAR had a significant correlation with LAI at the canopy-level. Canopy LAI (LAIc) can be expressed as the product of single tree LAI (LAI_s) and crown closure (Co). The crown closure Co was defined as the ground coverage by crowns in nadir direction and was calculated as a function of stem density (SD) and crown diameter (CD) (Equations (6) and (8)). Consequently, the canopy LAI (LAIc) was computed as:

\[
LAI_c = LAI_s \times \left(1 - e^{\frac{-\pi (0.5 CD^2 SD)}{10000 \cos \theta_s}} \right)
\]  

Using field data (\(n = 42\)), the logarithmic function relating fAPAR with LAI was determined (Figure 8):

\[
fAPAR = \min(0.1896 \times \ln(LAI) + 0.5502, 0.95)
\]  

On this basis, a fAPAR dataset was generated corresponding with the simulated spectral reflectances. Both the values of fAPAR and spectral reflectance fall between 0 and 1, so no data normalization was applied. Our empirical relation corresponds well with similar models reported in the literature [49].
3.2.2. Establishing the Network Architecture and Training

The backpropagation ANN was realized within MATLAB software (www.mathworks.com). The chosen neural network architecture consists of one input layer with four linear neurons for input of the three visible bands and the near-infrared reflectance of Landsat-8 plus one hidden layer with 20 neurons and one output layer with one neuron (fAPAR). The study was restricted to the four mentioned bands as these bands are included in most sensors, such as Quickbird, SPOT and two small Chinese optical satellites (HJ-1A and HJ-1B). Hence, the developed methodology can potentially be applied to map fAPAR from other sensors as well. Both the hidden layer and output layer use the sigmoid transfer function. The train epochs were set as 100.

Once the ANN was established, the training data were divided into three subsets [36]. Half of the simulated datasets were randomly selected for training, and the remaining half were divided into two parts: validation datasets and testing datasets. The validation datasets were used during the learning process of the network to prevent over-fitting. The testing datasets serve as internal evaluation of the ANN estimation performance.

The training process to calibrate ANN coefficients relied on the minimization of a misfit function by a back-propagation algorithm [50]. During this process, the values of the weights and biases in the network were iteratively adjusted to minimize the error between the network outputs and the measured outputs (fAPAR). The coefficient of determination ($R^2$) and the root mean squared error (RMSE) between the network outputs and the measured outputs were calculated to measure accuracy of the model inversion. The optimal network architecture was determined after several trials to provide the best estimation performance. Once calibrated, a neural net can be considered as an approximation of the surface response of the inverse radiative transfer model by means of a simple parametric model with known coefficients. A calibrated ANN can be applied easily to real image data. In case the retrieved fAPAR value fell outside the definition range (0–1), the value of the closest bound was used.

![Figure 8. fAPAR as a function of LAI for the forest stands at three test sites ($n = 42$). Both, fAPAR and LAI were measured in the field.](image-url)
4. Results and Discussion

4.1. Sensitivity Analysis of INFORM Parameters

Sensitivity analysis of model parameters is an important step in the process of model adaptation to regional conditions for later generation of synthetic reflectance data using CRMs. It is also an efficient way to investigate the effects of forest canopy variables on reflectance signatures and helps understanding how canopy structure and leaf bio-chemicals influence the model simulations.

In the sensitivity analysis, the variable parameters to be analyzed were classified into five ranks and the corresponding reflectance spectra were simulated from INFORM. During the simulation, other parameters in INFORM were kept constant (shown in Table 4). The analyzed variables were varied within the range of measured data (Table 2) providing information on the sensitivity of INFORM generated forest canopy reflectance for different stem densities, crown diameters, tree heights, dry matter contents and chlorophyll contents (a+b), respectively, in the wavelength range between 400 and 2500 nm (Figure 9). When one variable (e.g., LAI) was analyzed, other variables (e.g., SD, CD, H, Cab and Cm) were set as the average of the minimum and maximum.

Stem density (SD), crown diameter (CD) and LAI had very strong influence in all bands, in particular in the near-infrared. With increasing stem density (SD), the modeled reflectance gradually became saturated. The chlorophyll contents (a+b) and dry matter content had little influence in all bands due to very small parameter ranges. As a result, the chlorophyll contents (a+b) and dry matter content were kept constant for the remaining study (at 0.0035 and 44, respectively) (see Table 4).

![Figure 9. Cont.](image)
Figure 9. INFORM generated forest reflectance spectra (0.4–2.5 µm) for different values of stem density, crown diameter, LAI, height, dry matter content and leaf chlorophyll content (a+b), respectively. The ranges of simulated structural parameter were chosen according to observed ranges. Ranges of bio-chemical parameters were chosen according to LOPEX experiment.

4.2. Validation of the Modelled Canopy Reflectance

For the forty-two reference plots of our study area, forest reflectance spectra from 400 to 2500 nm were simulated with INFORM and resampled to Landsat-8 wavebands (visible and near-infrared bands). For the simulations, the observed forest structural and bio-chemical parameters were used. Modelled reflectances in the visible and near-infrared wavelength regions of samples with relatively high cosi (e.g., above 0.51—see Table 1) mostly compared well to the measured values (Figure 10). Most points were relatively closely distributed along the 1:1 line (e.g., within ±0.01 reflectance units in the visible channels and ±0.1 in the nIR) thus confirming other experimental results by [24].

On the contrary, for the nine samples taken in steep terrain (e.g., with low cosi—see Table 1), it was not possible to achieve a good forward simulation accuracy with INFORM. RMSE between measured and simulated reflectances were substantially higher compared to the samples with high cosi (Table 5). This finding clearly shows that the topographic correction was not well performing in very steep terrain. For this reason, it was decided to clearly distinguish between these two samples (e.g., high/low cosi) when analysing the fAPAR modeling results, as one cannot expect good CRM inversion results if the observed reflectance spectra are perturbed by terrain effects, etc.

As the improvement of existing topographic correction algorithms was out of the scope of the present study, we preferred to work with the imperfect Landsat data and instead distinguished the two groups of reference samples (see Table 1). Obviously, the imperfections related to the terrain correction would need to be resolved if similar methods are to be applied in hilly terrain.
Figure 10. Modeled and measured forest canopy reflectance (33 plots from ‘normal’ terrain—high cos¡) in different wavebands: (a) blue band; (b) green band; (c) red band and (d) near-infrared band of Landsat-8. The 1:1 line is shown as continuous line. The dotted lines in (a), (b) and (c), are described as y = x ± 0.01, and in (d), y = x ± 0.1.

Table 5. Accuracy of INFORM reflectance simulations with field measured canopy characteristics as inputs. The RMSE is between INFORM simulations and Landsat-8 observations. We distinguish steep terrain (e.g., low cos¡) and normal terrain (e.g., high cos¡).

<table>
<thead>
<tr>
<th></th>
<th>“Steep” Terrain (Low cos¡)</th>
<th>“Normal” Terrain (High cos¡)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>RMSE</td>
</tr>
<tr>
<td>Blue</td>
<td>9</td>
<td>0.013</td>
</tr>
<tr>
<td>Green</td>
<td>9</td>
<td>0.019</td>
</tr>
<tr>
<td>Red</td>
<td>9</td>
<td>0.021</td>
</tr>
<tr>
<td>nIR</td>
<td>9</td>
<td>0.200</td>
</tr>
</tbody>
</table>

4.3. Validation of fAPAR Inversion Results

A scatterplot (n = 33) of field measured and Landsat-8 based forest fAPAR values is shown in Figure 11 excluding the nine samples from very steep terrain. Most points are slightly below or scatter around the 1:1 line with a RMSE of 0.11. The explained variance (R²) amounts to 0.47. As stems are
“misinterpreted” as leaves in optical devices used for taking reference measurements, a slight overestimation of field measurements can be expected.

We noticed again a strong impact of topography (e.g., imperfect topographic correction) on modelling results. Regarding the fAPAR estimations obtained from the nine samples with low \(\cos i\), only 3% of the variance was explained with RMSE of 0.16 (Table 6). The main reason for the strong errors are related to the topography of the study area.

**Figure 11.** Validation of estimated fAPAR against *in situ* measured fAPAR (33 plots with high \(\cos i\)).

**Table 6.** Accuracy of forest fAPAR from INFORM against field measured values. We distinguish steep terrain (e.g., low \(\cos i\)) and normal terrain (e.g., high \(\cos i\)).

<table>
<thead>
<tr>
<th>“Steep” Terrain (Low (\cos i))</th>
<th>“Normal” Terrain (High (\cos i))</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td>RMSE</td>
</tr>
<tr>
<td>9</td>
<td>0.16</td>
</tr>
</tbody>
</table>

4.4. Estimation of Forest fAPAR in the Study Area

The trained ANN was applied to the Landsat-8 imagery of the entire study test site. No pixels were excluded for terrain orientation. As shown in Figures 12 and 13, estimated values of fAPAR are generally in a reasonable range. fAPAR values are stretched in Figure 12 along a color ramp for cartographic reasons. Although most of the forest stands in the fAPAR map appear rather homogenous, certain variations can be observed. Gradual variations in forest fAPAR are expected according to changing growing conditions. On the other hand, extreme values appearing in some regions with low \(\cos i\) (e.g., high uncertainty) are possibly artefacts related to terrain effects.
Figure 12. Distribution of fAPAR in the study area Dabie mountain in (a) and distribution of cosi in (b). Note that non-forest pixels are shown in white.

Figure 13. Frequency distribution of the forest fAPAR map of the study area. Only pixels covered with forest are taken into account. No pixel is excluded for terrain orientation.

5. Conclusions

In this paper, the forest fAPAR was estimated from Landsat-8 data through inversion of the physically-based INFORM radiative transfer model based on ANN. Three main conclusions can be drawn from our study:

1) INFORM seems moderately well suited for modeling the bi-directional spectral reflectance of forest canopies in the visible and near-infrared bands as a function of structural and bio-chemical forest characteristics. The parameters most strongly influencing the simulated reflectances of the model are LAI, crown diameter, and stem density. The modelled reflectance showed a consistent underestimation in red band and overestimation in near-infrared band. Probably, part of the differences could be reduced by using higher precision field measurement data. For example, in-situ leaf characteristics and other species-related factors were neglected in this study or
unavailable. Obviously, being a relatively simple CRM, INFORM only captures the main factors of variability. More advanced CRMs need to be tested to reach a closer agreement between forward simulations and Landsat observations.

(2) The forest fAPAR was modeled for the Dabie mountain test site by inverting INFORM through an artificial neural network approach. Results suggest that this method can successfully estimate forest fAPAR from multispectral images and achieve an acceptable accuracy with $RMSE$ of 0.11 (14% of average) and $R^2$ of 0.47. In further studies, efforts will be made to estimate and evaluate time series of forest fAPAR based on the method.

(3) The mountainous terrain posed the biggest challenge for successful retrieval of fAPAR. The employed simple topographic C-correction method created several overcorrections and artifacts negatively influencing the fAPAR retrieval. The insufficient topographic correction already became visible when running INFORM in forward mode (e.g., with field measured canopy characteristics entered into the CRM). In doing so, it was noted that the simulated canopy reflectance spectra strongly deviated from the observed (Landsat-8) spectral profiles.

Acknowledgments

This study was financially supported by the “Strategic Priority Research Program-Climate Change: Carbon Budget and Related Issues” of the Chinese Academy of Sciences (Grant No. XDA05050106). The data collection of this study was supported by the Lake and Watershed Data Center, Scientific Data Sharing Platform for Lake and Watershed (SDSPLW), Nanjing Institute of Geography and Limnology, Chinese Academy of Sciences.

Author Contributions

Huili Yuan conceived and designed the experiments, performed the experiments, analyzed the data and wrote the paper; Clement Atzberger provided the transfer radiative model, analyzed the data and supervised the experiments; Fei Li conceived the experiments and provided the inverse tool; Ronghua Ma, Juhua Luo and Steven Arthur Loiselle contributed in the acquisition of field data.

Conflicts of Interest

The authors declare no conflict of interest.

References


© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/).