A Semi-Empirical Retrieval Method of Above-Ground Live Forest Fuel Loads by Combining SAR and Optical Data

Yanxi Li and Binbin He *

School of Resources and Environment, University of Electronic Science and Technology of China, Chengdu 611731, China
* Correspondence: binbinhe@uestc.edu.cn

Abstract: Forest fuel load is the key factor for fire risk assessment, firefighting, and carbon emissions estimation. Remote sensing technology has distinct advantages in fuel load estimation due to its sensitivity to biomass and adequate spatiotemporal observations for large scales. Many related works applied empirical methods with individual satellite observation data to estimate fuel load, which is highly conditioned on local data and limited by saturation problems. Here, we combined optical data (i.e., Landsat 7 ETM+) and spaceborne Synthetic Aperture Radar (SAR) data (i.e., ALOS PALSAR) in a proposed semi-empirical retrieval model to estimate above-ground live forest fuel loads ($FL_{AGL}$). Specifically, optical data was introduced into water cloud model (WCM) to compensate for vegetation coverage information. For comparison, we also evaluated the performance of single spaceborne L-band SAR data (i.e., ALOS PALSAR) in fuel load estimation with common WCM. The above two comparison experiments were both validated by field measurements (i.e., BioSAR-2008) and leave-one-out cross-validation (LOOCV) method. WCM with single SAR data could achieve reasonable performance ($R^2 = 0.64$ or higher and RMSE$ = 35.3\%$ or lower) but occurred an underestimation problem especially in dense forests. The proposed method performed better with $R^2$ increased by $0.05-0.13$ and RMSE$ decreased by $5.8-12.9\%$. We also found that the underestimation problem (i.e., saturation problem) was alleviated even when vegetation coverage reached $65\%$ or the total $FL_{AGL}$ reached about $183$ Tons/ha. We demonstrated our $FL_{AGL}$ estimation method by validation in an open-access dataset in Sweden.

Keywords: fire risk; above-ground forest fuel load; water cloud model; SAR; optical data; vegetation coverage; inversion

1. Introduction

Wildfires have a vital impact on the formation and succession of ecosystems [1,2]. They can enrich biodiversity and vegetation vertical structure thereby promoting the nutrient cycle and enhancing vegetation resistance to diseases or insect pests [3,4]. However, an uncontrolled wildfire may destroy the soil and water conservation ability of vegetation, release atmospheric greenhouse gases, and even threaten the safety of human life and property [5–10]. In most forest ecosystems, wildfires are mainly divided into surface fires and crown fires, in which the above-ground live fuel load ($FL_{AGL}$) is the main energy resource. Prior knowledge of $FL_{AGL}$ (refer to the canopy and stem fuel load) is important for fire risk assessment, firefighting, and carbon emissions estimation.

The canopy fuel load commonly includes foliage (including needles and leaves), branch wood, and other suspended biomass such as lichens and mosses [11]. The foliage fuel load (FFL, referring to the dry weight of needles and leaves per unit area) and branch fuel load (BFL, referring to the dry weight of branches per unit area) are the main energy source supporting the spread of crown fire [12–17]. Stem fuel load (SFL, referring to the dry weight of stems per unit area) can interact with fires and help carry fires although it...
is not a substantial loss of biomass [18–22]. Hence, the $F_{LAGL}$ (i.e., FFL, BFL and SFL) are generally affected by fire [18] and play an important role in fire-related factors assessment, including flame length, fire intensity, fire severity and fuel loss [23–26]. $F_{LAGL}$ is also an indispensable input parameter for fire behavior models [27–29].

Precise grasp of the spatial distribution of $F_{LAGL}$ (i.e., fuel accumulation in two-dimensional plane space) is helpful for fire managers to assess forest fire risk and manage fuel load in advance (e.g., planning prescribed burning) and avoid uncontrollable fire disasters. Traditionally, $F_{LAGL}$ is usually obtained by combining field inventory data (e.g., diameter at breast height, height, tree species, etc.) and allometric equations. This kind of method derives the relationship between stand parameters and biomass fractions by cutting down trees, which is neither a desirable nor effective way [30]. The rapid development of remote sensing (RS) technology has brought new opportunities for the quantitative derivation of $F_{LAGL}$, since RS can provide adequate spatiotemporal earth observations from regional to global scale and implement fuel loads assessment even in rugged topography [11,31,32].

Previous studies mainly utilize single optical data to estimate fuel load. Those studies focused on the vegetation type classification from multispectral information first, and then applied the corresponding fuel model [33,34]. For high spatial resolution optical images (e.g., 0.6 m), the tree scale parameters are also derived to estimate fuel load by image segmentation technology [32,35]. Although optical data has advantage over vegetation classification and foliage biomass indication [36], it is not a good indicator of woody biomass. Because optical data can hardly characterize the vertical structure of forest, and the saturation point of spectral bands is low (between 15 and 100 Tons/ha) [11,37,38]. On the contrary, the exploration of SAR data in fuel load estimation is far less than that of optical data. While there are a number of studies focused on aboveground biomass estimation using SAR sensors [39–41], there are few specifically targeted at fuel load assessments, and even fewer that examine components of fuel loads (e.g., foliage, branch and stem).

In this study, we first used spaceborne SAR data in $F_{LAGL}$ estimation via common Water Cloud Model (WCM). Next, we introduced the optical data as vegetation information into common WCM. These two models were both validated by the BioSAR-2008 dataset and compared with each other. The spatial distribution of $F_{LAGL}$ was mapped using the optimal model. The overarching objective is to provide accurate information of $F_{LAGL}$ for fire managers, which can effectively guide fuel management work and thus achieve early prevention, suppression, and response of forest fire by planning prescribed burning, increasing the distance between fuels, etc.

2. Materials

2.1. Study Site and Ground Measurements

The study area was selected in the forest region of Krycklan which is located in the boreal zone of Västerbotten county, northern Sweden (64°14′N, 19°46′E). This site covers 6800 ha in total, which is covered mostly by forests with scattered cropland and lakes. Pine, spruce, and birch are dominant tree species. The terrain is rugged, with an elevation between 125 and 350 m. The annual mean precipitation and temperature are 600 mm and 1 °C, respectively, and about half of the rainfall happens in snow.

Forest inventory data were obtained from the BioSAR campaign which was carried out in October 2008. The BioSAR-2008 dataset is available in the European Space Agency (ESA) Earth Observation campaigns (https://www.esa.int last access 10 December 2022). In the BioSAR-2008 dataset, a total of 31 forest stands were investigated, whose sizes range from 2.4 to 26.3 ha. Each stand is arranged with 10 circular sampling plots (radius 10 m). In each sample plot, the diameter at breast height (DBH) of all trees with a DBH > 4 cm were measured, and randomly selected 1.5 sample trees in average (according to the proportion of the basal area) to record their height and age. These field survey data were combined with allometric equations [42] to obtain different parts of fuel load (i.e., biomass
fractions). The stand-level summaries of the data were obtained by averaging across all sampling plots. Detailed information about these field inventory data can be found in [43]. However, two of the 31 forest stands are located adjacent to the bare land which may distort the reflected signal. We eliminated these two stands since they were classified as non-forest in global forest/non-forest maps from ALOS PALSAR data [44], thus resulting in 29 forest stands shown in Figure 1.

![Figure 1. Geographical location of the study area. (a) The location of study area in Sweden. (b) The distribution of forest stands.](image)

2.2. Satellite Data

The surface reflectance product of Landsat 7 ETM+ with a spatial resolution of 30 m was applied as multispectral data in this study. To match the field inventory data, a pre-processed Landsat 7 ETM+ product (7 October 2008) was selected and downloaded from Google Earth Engine (GEE) [45]. However, the ETM+ data obtained after the 31 May 2003 has scan-line-off strips due to a problem in the scan line corrector system [46]. Part of the stands in this study were affected by the gaps. For these stands, all pixels away from data gaps were extracted and averaged. For those stands that are not affected, all pixels were extracted and averaged as the observed optical data of each stand. There was no stand wholly located in the gaps. We compared pixels for which data were missing and for which data were present by averaging coincident data in the Landsat VCF product. There was a high correlation between the scan-line gaps and surrounding pixels, and we concluded it was acceptable to average across stand polygons without the missing pixel data.

The ALOS (Advanced Land Observing Satellite) PALSAR (Phased Array type L-band Synthetic Aperture Radar) mosaic yearly product with slope-correction and ortho-rectification was used as spaceborne SAR data in this study [47]. We acquired the product of 2008 with a spatial resolution of 25 m from GEE. The working wavelength of 23.6 cm in the L-band makes the sensor signal more sensitive to forest structure and moisture characteristics thus more suitable for forest monitoring [48]. The digital numbers were converted to gamma naught ($\gamma^0$) as follows [44]:

$$\gamma^0(dB) = 10 \times \log_{10}(DN^2) - 83$$

(1)

where $\gamma^0$ is backscatter coefficient in $dB$ unit and $DN$ is the digital number. The gamma naught ($\gamma^0$) in $dB$ unit was converted to the linear unit:

$$\gamma^0(linear) = 10^{\gamma^0(dB)/10}$$

(2)

The average of the converted SAR data within each stand is taken as the observed SAR data of each stand.
Auxiliary data includes vegetation coverage and land cover data. The vegetation coverage data used in this study is Landsat Vegetation Continuous Fields (VCF). Sexton [49] generated this 30-m-resolution tree cover percentage dataset by rescaling the 250-m-resolution MODerate-resolution imaging spectroradiometer (MODIS) VCF product and combining Landsat optical data. The “tree_canopy_cover” band of this product refers to the percentage of woody vegetation with a height greater than 5 m above the horizontal ground. They provided four global products at the 30 m resolution in 2000, 2005, 2010 and 2015. Here, we acquired the 2010 epoch from GEE to match field data. The land cover data used in this study is GlobeLand30 (http://www.globallandcover.com/ last access 10 December 2022). GlobeLand30 is the first 30 m resolution global land cover data product with an overall accuracy of 83.5% [50,51], which includes 10 types of land cover, such as artificial surface, vegetation, etc. There are three products in 2000, 2010 and 2020, respectively. The product of 2010 was selected for this experiment and for more details about this product please refer to [51].

3. Methods

3.1. Water Cloud Model (WCM)

The methods for vegetation parameter inversion based on remote sensing include physically based models, empirical models (statistical relationships), and semi-empirical models. Physical models are robust and transferable but have limitations in application due to their complicated calculations [52]. Empirical models are straightforward and widely used, but they lack physical meaning and are highly conditioned on local data [53,54]. The semi-empirical models integrate the empirical presentations with the physical model which sidestep the non-transferability obstacles of empirical methods and the complexity of physical models [55].

Therefore, the semi-empirical model (i.e., WCM) was applied in this study to simulate the scattering mechanisms of the forest. The L-band SAR data used in this experiment has a larger wavelength than the size of leaves, thus it can completely penetrate the forest canopy and receive signals below the canopy [56,57]. Since the L-band SAR data has strong penetration, an extended water cloud model which considering the canopy cover was used to supplement the contribution of gaps in forests to total backscatter [58]. In this model, the total backscatter is the sum of three parts:

\[
\sigma_t = (1 - f_{ve}) \sigma_{gr}^0 + f_{ve} \sigma_{veg}^0 e^{-\alpha h} + f_{ve} \sigma_{veg}^0 (1 - e^{-\alpha h})
\]

(3)

The first part represents the direct ground scattering through forest gaps. The second part represents ground scattering attenuated by forest coverage. The last part represents direct scattering from vegetation. \( f_{ve}, \sigma_{gr}^0 \) and \( \sigma_{veg}^0 \) are fractional vegetation coverage (the percentage of horizontal ground covered by vegetation), ground backscatter coefficient, and vegetation backscatter coefficient, respectively. \( e^{-\alpha h} \) describes the two-way tree transmissivity which is exponentially related to \( \alpha \) (the two-way attenuation per meter of tree canopy) and \( h \) (the thickness of vegetation layer).

By combining the similar items of Equation (3), it can be expressed by a simpler form:

\[
\sigma_t = \sigma_{gr}^0 T_{for} + \sigma_{veg}^0 (1 - T_{for})
\]

(4)

\[
T_{for} = (1 - f_{ve}) + f_{ve} e^{-\alpha h}
\]

(5)

where \( T_{for} \) describes the forest transmissivity. In biomass estimation, \( T_{for} \) can be replaced with a biomass related function [52]:

\[
\sigma_t = \sigma_{gr}^0 e^{-\delta B} + \sigma_{veg}^0 (1 - e^{-\delta B})
\]

(6)

where \( \delta \) is the forest transmissivity parameter and \( B \) is aboveground biomass. In the forest with gaps, additional compensation for \( \sigma_{veg}^0 \) is necessary since the backscatter from the forest still contains a certain contribution of ground. Equation (6) can be inverted to
get $\sigma_{\text{veg}}^0$ from those so-called “dense-forest” pixels backscatter $\sigma_{df}^0$ and biomass value of “dense-forest” class $B_{df}$:

$$\sigma_{\text{veg}}^0 = \frac{\sigma_{df}^0 - \sigma_{gr}^0 e^{-\delta_{df}}} {1 - e^{-\delta_{df}}} \quad (7)$$

3.2. Semi-Empirical Modelling for Combining Optical and SAR Data

3.2.1. Limitation of WCM

In previous applications of WCM presented in Section 3.1, SAR is the only remote sensing data source used to calibrate the model and retrieve the target parameters. The single parameter ($\delta$) that needs calibration in WCM can be expressed as the following equation according to Equations (4)–(6):

$$\delta = -\frac{1}{B} \ln (1 - f_{vc} + f_{vc} e^{-a h}) \quad (8)$$

For each determined biomass value ($B$), the forest transmissivity ($\delta$) can be further abstracted as the following functional relationship:

$$\delta = F(f_{vc}, a, h) \quad (9)$$

That is, the forest transmissivity coefficient which describes the trend of backscatter with the increase of biomass is determined by three physical variables: fractional vegetation coverage ($f_{vc}$), two-way signal attenuation ($a$), and vegetation layer thickness ($h$). The last two variables can be indicated by SAR well. However, SAR shows a relatively weak representative ability of $f_{vc}$ compared to optical data. On contrary, many studies have used optical data to derive vegetation coverage since optical information (including spectral reflectance and vegetation indices (VIs)) is sensitive to vegetation coverage [59–66].

3.2.2. Vegetation Coverage Information Expressed by Optical Data

The optical reflectance signal from vegetation canopy is composed of both soil information and vegetation information which can be divided by vegetation coverage [67,68].

$$R_{\text{total}} = f_{vc} R_{\text{veg}} + (1 - f_{vc}) R_{\text{soil}} \quad (10)$$

where $R_{\text{total}}$, $R_{\text{veg}}$ and $R_{\text{soil}}$ represent the total reflectance, the pure vegetation reflectance, and the pure soil reflectance, respectively. Hence, the vegetation coverage can be derived inverting Equation (10) [66,69]:

$$f_{vc} = \frac{R_{\text{total}} - R_{\text{soil}}} {R_{\text{veg}} - R_{\text{soil}}} \quad (11)$$

It should be noted that the pure soil reflectance and the pure vegetation reflectance can be considered as constant in a certain study area [61]. Therefore, Equation (11) can be simplified as:

$$f_{vc} = a R + b \quad (12)$$

where $a$ and $b$ are constants in the study area and $R$ is the optical information of pixels including spectral bands and VIs.

In addition, according to Equation (10) and WCM, the forest transmissivity of optical data can be expressed as:

$$T_{\text{for}} = 1 - f_{vc} \quad (13)$$

Similar to WCM described above, here the forest transmissivity of optical data can be replaced by an exponential function of biomass:

$$T_{\text{for}} = 1 - f_{vc} = e^{-tB} \quad (14)$$
where \( \tau \) expresses the forest transmissivity of optical data and \( B \) is aboveground biomass. Then, the vegetation coverage can also be expressed as:

\[
\hat{f}_{vc} = 1 - e^{-\tau B}
\]  

(15)

3.2.3. Introducing Optical Data as Vegetation Coverage Information into WCM

This study introduces optical data as vegetation coverage information into WCM. Assuming that can enhance the model’s ability to simulate ground reflection signals in dense forest areas by supplementing vegetation coverage information.

Here, the vegetation coverage is introduced into WCM first:

\[
\sigma^0_t \otimes f_{vc} = \left[ \sigma^0_{gr} e^{-\delta R} + \frac{\sigma^0_{df} - \sigma^0_{gr} e^{-\delta R}}{1 - e^{-\delta R}} (1 - e^{-\delta B}) \right] \otimes f_{vc} 
\]  

where \( \otimes \) represent any one of the four arithmetic (i.e., addition, subtraction, multiplication, and division) and \( f_{vc} \) is vegetation coverage. Then, the new model is obtained by express \( f_{vc} \) using optical data:

\[
\sigma^0_t \otimes R = \left[ \sigma^0_{gr} e^{-\delta R} + \frac{\sigma^0_{df} - \sigma^0_{gr} e^{-\delta R}}{1 - e^{-\delta R}} (1 - e^{-\delta F}) \right] \otimes \frac{1 - b - e^{-\tau F}}{a} 
\]  

(17)

where \( \sigma^0_t \), \( \sigma^0_{gr} \) and \( \sigma^0_{df} \) represent the total backscatter, the ground backscatter, and the dense vegetation backscatter, respectively; \( R \) is optical information including spectral bands and Vi's; \( F_{df} \) and \( F \) are “dense forest” fuel load and fuel load, respectively; \( a, b, \delta \), and \( \tau \) are the four empirical coefficients in the new model.

3.3. Model Calibration and Fuel Load Retrieval

For the calibration of the new model expressed in Equation (17), there are seven unknowns to be determined: \( \sigma^0_{gr}, \sigma^0_{df}, a, b, \tau, \delta \) and \( F_{df} \). \( \sigma^0_{gr} \) represents the backscatter coefficient of the surface without vegetation cover (i.e., ground) and \( \sigma^0_{df} \) is the backscatter coefficient of dense vegetation cover. The specific calibration will be described in detail later. The other parameters (i.e., \( a, b, \tau, \delta \) related to the forest structure characteristics were optimized by the non-linear Levenberg–Marquardt algorithm [70]. \( F_{df} \) is the fuel load value of the “dense forest” class. Here, we selected the 90 percentile of fuel load value in our study area as \( F_{df} \) according to [52,71].

The acquisition of \( \sigma^0_{gr} \) needs auxiliary data including Landsat VCF product and GlobeLand30 to ensure the reliability of the selected candidate “ground” pixels, since non-soil or non-vegetation pixels (such as water body, impervious surface, etc.) may cause fluctuation of backscatter coefficient, resulting in distorted \( \sigma^0_{gr} \). Considering there are very few pixels with vegetation coverage of zero after excluding non-soil or non-vegetation pixels, the pixels with vegetation coverage less than 25% were also included in candidate “ground” pixels. Taking the study area as the center, all pixels within the radius of 250 pixels that match the above conditions were selected. The median value of the backscatter coefficient of all selected pixels was taken as \( \sigma^0_{gr} \) since the median value can represent the centralized trend of data.

The acquisition of \( \sigma^0_{df} \) is similar to \( \sigma^0_{gr} \). The corresponding Landsat VCF product of the study area should be used as auxiliary data. By selecting the pixels with vegetation coverage greater than 70%, the median value of all pixels was taken to represent the reflection signal under dense vegetation coverage in the study area.

Since the data were under constraints of limited size, the leave-one-out cross-validation (LOOCV) method [72] was used to validate the new model. In the LOOCV method, each sample was excluded one by one, and the model is calibrated with the remaining samples to predict the excluded sample.

The new model has no corresponding numerical solution. Thus, the look-up table (LUT) methodology was applied to acquire the optimal fuel load from the observed RS
Remote Sens. 2023, 16, 5

To build the LUT, the calibrated model was executed forward with one input (i.e., fuel load value) to generate simulated RS data. The range of fuel load value in LUT was set according to field FL<sub>AGL</sub> measurements.

After the completion of LUT, the fuel load was retrieved through a cost function that quantized the similarity between the simulated and observed signals to search out the best-fitted simulated RS signal in LUT with the observed one. In this experiment, we used the absolute difference as the cost function:

$$MAE_r = |\mu_{sim} - \mu_{obs}|$$ (18)

where $\mu_{sim}$ and $\mu_{obs}$ represent the simulated and observed RS data, respectively.

The determine coefficient ($R^2$, Equation (19)), root mean square error (RMSE, Equation (20)) and relative RMSE (RMSE<sub>r</sub>, Equation (21)) were used as accuracy indicators for models:

$$R^2 = \frac{\sum_{i=1}^{n}(P_i - \bar{P})(O_i - \bar{O})}{\sqrt{\sum_{i=1}^{n}(P_i - \bar{P})^2 \sum_{i=1}^{n}(O_i - \bar{O})^2}}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n}(P_i - O_i)^2}$$

$$RMSE_r = \frac{RMSE}{\bar{O}}$$

where $P_i$ and $O_i$ represent predicted and observed fuel load, respectively, $\bar{P}$ and $\bar{O}$ represent the mean of predicted and observed fuel load, respectively, and $n$ is the sample size.

3.4. Comparison Experiments

In this study, the WCM with L-band spaceborne SAR data was first applied to FL<sub>AGL</sub> (above-ground forest fuel loads, indicating FFL, BFL and SFL) inversion. Further, the new model was developed to retrieve FL<sub>AGL</sub> by introducing optical data as vegetation coverage information into WCM. Performances of WCM and new model were adequately verified using the leave-one-out cross-validation (LOOCV) method. The flowchart of the new model for FL<sub>AGL</sub> estimation from SAR and optical data is illustrated in Figure 2.

![Figure 2. Methodological flowchart of FL<sub>AGL</sub> estimation from integrated optical and SAR data with new model.](image)

In addition to the six spectral bands of Landsat, four VIs were also selected for new model exploration in this study. The VIs including Normalized Difference Infrared Index
(NDII$_{SWIR2}$) [73], Normalized Difference Tillage Index (NDTI) [74], Normalized Difference Vegetation Index (NDVI) [75], and Ratio Vegetation Index (RVI) [76] are listed in Table 1. The composition of NDII$_{SWIR2}$ and NDTI both contain short wave infrared band, which indicates the vegetation water content and dry matter situation well [54,77–79]. NDVI composed of Red and NIR bands with minimal water absorption is also sensitive to vegetation coverage thus a good indicator of vegetation biophysical variable [54,80]. RVI that defined as the ratio of Green band to Blue band here since chlorophyll absorbs red and blue light, reflects green light and carotenoids also absorb blue light [81].

Table 1. VIs introduced into WCM for FLA Gr retrieval.

<table>
<thead>
<tr>
<th>Vegetation Indices</th>
<th>Equations</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normalized Difference Infrared Index</td>
<td>$NDII = \frac{NIR - SWIR2}{NIR + SWIR2}$</td>
<td>[73]</td>
</tr>
<tr>
<td>Normalized Difference Tillage Index</td>
<td>$NDTI = \frac{SWIR1 - SWIR2}{SWIR1 + SWIR2}$</td>
<td>[74]</td>
</tr>
<tr>
<td>Normalized Difference Vegetation Index</td>
<td>$NDVI = \frac{NIR - RED}{NIR + RED}$</td>
<td>[75]</td>
</tr>
<tr>
<td>Ratio Vegetation Index</td>
<td>$RVI = \frac{Green}{Blue}$</td>
<td>[76]</td>
</tr>
</tbody>
</table>

Based on the 10 optical variables (spectral bands and vegetation indices) in this paper, we explored all the possible combinations. To verify the effectiveness of the proposed algorithm, the experimental results were compared with those before the introduction of optical data. A total of 41 experiments were conducted in this study (Table 2).

Table 2. Experimental setup in modelling FLA Gr.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Data Source</th>
<th>Input Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>WCM</td>
<td>HV</td>
<td>HV</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HV $\otimes$ Blue</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HV $\otimes$ Green</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HV $\otimes$ Red</td>
</tr>
<tr>
<td>New model</td>
<td>HV, Blue, Green, Red, NIR, SWIR1, SWIR2, NDII, NDTI, NDVI, RVI</td>
<td>HV $\otimes$ NIR</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HV $\otimes$ SWIR1</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HV $\otimes$ SWIR2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HV $\otimes$ NDII</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HV $\otimes$ NDTI</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HV $\otimes$ NDVI</td>
</tr>
<tr>
<td></td>
<td></td>
<td>HV $\otimes$ RVI</td>
</tr>
</tbody>
</table>

$\otimes$ represents any one of the four arithmetic (i.e., addition, subtraction, multiplication, and division).

4. Results

The performance of FLA Gr inversion using WCM with SAR is shown in Table 3. It can be found that the WCM with HV polarization of L-band spaceborne SAR data has potential in FLA Gr inversion with reasonable performance ($R^2 = 0.64$ or higher).
Table 3. Retrieval results of FL_{AGL} (FFL, BFL and SFL) by HV polarization backscatter alone using WCM.

<table>
<thead>
<tr>
<th>Fuel Load</th>
<th>R²</th>
<th>RMSE</th>
<th>RMSEr</th>
</tr>
</thead>
<tbody>
<tr>
<td>FFL</td>
<td>0.70</td>
<td>1.85</td>
<td>29.9%</td>
</tr>
<tr>
<td>BFL</td>
<td>0.70</td>
<td>4.23</td>
<td>27.4%</td>
</tr>
<tr>
<td>SFL</td>
<td>0.64</td>
<td>25.80</td>
<td>35.3%</td>
</tr>
</tbody>
</table>

The quantitative estimation of FL_{AGL} can be further improved by adding optical information with the new model. The performances of new models that performed better than WCM are listed in Table 4. It can be found that there are many new models that perform better than WCM through introducing optical data as vegetation coverage information into WCM (R² increased by 0.05–0.13 and RMSEr decreased by 5.8–12.9%).

Table 4. Retrieval results of FL_{AGL} (FFL, BFL and SFL) by the combination of HV backscatter and optical information (band or vegetation index) with new model. The bold underlined items are the best retrieval results.

<table>
<thead>
<tr>
<th>Fuel Load</th>
<th>R²</th>
<th>RMSE</th>
<th>RMSEr</th>
<th>Combination Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>FFL</td>
<td>0.79</td>
<td>1.38</td>
<td>22.3%</td>
<td>HV/Green</td>
</tr>
<tr>
<td></td>
<td>0.77</td>
<td>1.49</td>
<td>24.0%</td>
<td>HV/RVI</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>1.56</td>
<td>25.2%</td>
<td>HV-SWIR1</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>1.59</td>
<td>25.7%</td>
<td>HV-SWIR1</td>
</tr>
<tr>
<td></td>
<td>0.75</td>
<td>1.60</td>
<td>25.8%</td>
<td>HV-SWIR2</td>
</tr>
<tr>
<td></td>
<td>0.74</td>
<td>1.64</td>
<td>26.5%</td>
<td>HV-Blue</td>
</tr>
<tr>
<td></td>
<td>0.72</td>
<td>1.64</td>
<td>26.5%</td>
<td>HV-Blue</td>
</tr>
<tr>
<td></td>
<td>0.72</td>
<td>1.69</td>
<td>27.3%</td>
<td>HV/Red</td>
</tr>
<tr>
<td></td>
<td>0.71</td>
<td>1.73</td>
<td>27.9%</td>
<td>HV-SWIR2</td>
</tr>
<tr>
<td></td>
<td>0.71</td>
<td>1.74</td>
<td>28.1%</td>
<td>HV × NDII</td>
</tr>
<tr>
<td></td>
<td>0.70</td>
<td>1.62</td>
<td>26.1%</td>
<td>HV-Green</td>
</tr>
<tr>
<td></td>
<td>0.70</td>
<td>1.71</td>
<td>27.6%</td>
<td>HV/NIR</td>
</tr>
<tr>
<td>BFL</td>
<td>0.75</td>
<td>3.34</td>
<td>21.6%</td>
<td>HV/RVI</td>
</tr>
<tr>
<td></td>
<td>0.71</td>
<td>3.44</td>
<td>22.3%</td>
<td>HV-Green</td>
</tr>
<tr>
<td></td>
<td>0.71</td>
<td>3.56</td>
<td>23.1%</td>
<td>HV/Green</td>
</tr>
<tr>
<td></td>
<td>0.71</td>
<td>3.56</td>
<td>23.1%</td>
<td>HV/NIR</td>
</tr>
<tr>
<td></td>
<td>0.77</td>
<td>16.38</td>
<td>22.4%</td>
<td>HV/NDII</td>
</tr>
<tr>
<td></td>
<td>0.77</td>
<td>16.56</td>
<td>22.7%</td>
<td>HV/NDTI</td>
</tr>
<tr>
<td></td>
<td>0.74</td>
<td>17.35</td>
<td>23.7%</td>
<td>HV/NDVI</td>
</tr>
<tr>
<td></td>
<td>0.69</td>
<td>19.13</td>
<td>26.2%</td>
<td>HV-Blue</td>
</tr>
<tr>
<td></td>
<td>0.69</td>
<td>19.68</td>
<td>26.9%</td>
<td>HV + Blue</td>
</tr>
<tr>
<td></td>
<td>0.68</td>
<td>19.90</td>
<td>27.2%</td>
<td>HV-Green</td>
</tr>
<tr>
<td>SFL</td>
<td>0.67</td>
<td>19.68</td>
<td>26.9%</td>
<td>HV × RVI</td>
</tr>
<tr>
<td></td>
<td>0.67</td>
<td>19.90</td>
<td>27.2%</td>
<td>HV-Red</td>
</tr>
<tr>
<td></td>
<td>0.67</td>
<td>20.33</td>
<td>27.8%</td>
<td>HV + Green</td>
</tr>
<tr>
<td></td>
<td>0.66</td>
<td>19.89</td>
<td>27.2%</td>
<td>HV/Blue</td>
</tr>
<tr>
<td></td>
<td>0.66</td>
<td>21.10</td>
<td>28.9%</td>
<td>HV/NIR</td>
</tr>
<tr>
<td></td>
<td>0.65</td>
<td>20.89</td>
<td>28.6%</td>
<td>HV-SWIR2</td>
</tr>
<tr>
<td></td>
<td>0.64</td>
<td>20.83</td>
<td>28.5%</td>
<td>HV + Red</td>
</tr>
</tbody>
</table>

For foliage fuel load, when SAR data were used alone for inversion, R² is 0.70 and RMSE is 1.85 Tons/ha (Table 3). Almost all optical data can be fused with HV backscattering coefficient to improve accuracy (Table 4). The R² was increased from 0.70 to 0.79 and the RMSEr was decreased from 29.9% to 22.3%. Taking the Blue band as an example,
among all the four combination ways of HV and Blue band (addition, subtraction, multiplication, and division), both the inversion performance of ‘HV-Blue’ and ‘HV/Blue’ are better than that of single HV, so as to other optical variables. In other words, the addition of any optical information in this study will certainly improve the retrieval results of FFL since leaves are distributed in the forest canopy, and the multispectral signals can reflect the density of green vegetation and biochemical characteristics of leaves such as leaf type, chlorophyll, etc. Among all the inversion results of FFL by new model, ‘HV/Green’ is the best (R²: 0.79; RMSE: 1.38 Tons/ha; RMSEr: 22.3%), as shown in bold underlined of Table 4.

For branch fuel load, when SAR data were used alone for inversion, the performance was similar to FFL with R² of 0.70 Tons/ha and RMSE of 4.23 Tons/ha (Table 3). As for the new model, the most optical information can be suitably integrated with HV backscatter which reduced RMSE especially the Green band and RVI, indicating that the addition of optical information improves the performance of WCM for forest backscatter simulation. The reason is that optical data is sensitive to vegetation coverage, as described in Section 3.2. Among all the results obtained by the new model, ‘HV/RVI’ is the best (R²: 0.75; RMSE: 3.34 Tons/ha; RMSEr: 21.6%), as shown in bold and underlined in Table 4.

For stem fuel load, the performance of WCM with HV backscatter alone was worse than BFL or FFL (Table 3), while two kinds of sensor information compensate each other well especially SAR data and optical vegetation index (Table 4). Specifically, a single multispectral band has little effect on SFL inversion, while VIs selected in this study have a more obvious effect. Among all new model inversion results of SFL, ‘HV/NDII’ is the best (R²: 0.77; RMSE: 16.38 Tons/ha; RMSEr: 22.4%), as shown in bold and underlined in Table 4.

More detailed information about the best results of FL_AGL inversion using WCM and the new model is shown in Figure 3. For FL_AGL retrieved by HV backscatter alone using WCM (the first row in Figure 3), the dispersion of observed fuel load and predicted by WCM is scattered, and there is a distinct underestimation at the high value. The R² ranges from 0.64 to 0.70 while the RMSEr ranges from 27.4% to 35.3%. The fusion of optical information and HV polarization using the new model effectively alleviated the underestimation of high value and improved the model performance (R² ranges from 0.75 to 0.79 while the RMSEr ranges from 21.6% to 22.4%).

Figure 3. Comparison of FL_AGL (FFL, BFL and SFL) inversion results using WCM (top row) and the new model (bottom row) on BioSAR-2008.

Based on the optimum models, we derived the spatial distribution of FL_AGL in study area and shown in Figure 4. Considering the scan-line-off strips in Landsat 7, we filled
these gaps using ENVI 5.3 software before mapping to make the final fuel load map spatially continuous. The product was masked by the GlobeLand30 to exclude those non-forest areas. It can be found that the distribution of these three types of fuel loads in space presents a good consistency.

![Image](image.png)

**Figure 4.** The spatial distribution of $F_{LAGL}$ in the study area.

The performances of WCM and new model with different vegetation coverage are shown in Figure 5. The observed and predicted fuel loads by WCM and new model were sorted by fractional vegetation coverage. The black dots, blue dots, and red dots are observed fuel loads, predicted fuel loads by WCM, and predicted fuel loads by the new model, respectively. The dotted lines of corresponding colors are the trend lines fitted by these points. It can be found that with the increase in vegetation coverage, the fuel load shows an upward trend. Particularly, the trend of the red line (predicted fuel loads by new model) is more similar to that of the black line (observed fuel load) than the blue line (predicted fuel loads by WCM). Moreover, the WCM using HV alone shows an underestimation problem in the case of dense fuel loads especially when $F_{FL} > 4$ Tons/ha, $B_{FL} > 10$ Tons/ha, and $S_{FL} > 30$ Tons/ha. However, the new model can effectively alleviate this problem by introducing optical data as vegetation coverage information into WCM. Specifically, the new model did not exhibit any underestimation even when vegetation coverage reached 65% and the total $F_{LAGL}$ reached about 183 Tons/ha.

![Image](image.png)

**Figure 5.** Comparison of model performance with different vegetation coverage.

5. Discussion

5.1. Significance of this Study and Analysis with Previous Studies

$F_{LAGL}$ are generally affected by fire [18] and play an important role in the assessment of fire-related factors [23–26]. However, there are few previous studies on the quantitative inversion of fuel load using spaceborne SAR data. In fact, previous studies have shown high sensitivity of SAR on forest biomass components and stand structure [82–86]. In particular, the long band electromagnetic waves are sensitive to forest biomass and water
content thus can be used as a direct measurement of forest biomass and structure [87], thus derived many successful applications of SAR in above-ground biomass estimation [52,88–91]. The results in this study are consistent with previous studies that SAR can be a good indicator of FL_{LAG} (i.e., biomass components) with R^2 is 0.64 or higher and RMSEr is 35.3% or lower. The reason of this result may due to the strong penetration ability of SAR signals and its ability to characterize forest vertical structure [87].

However, single SAR data-based biomass estimation studies exhibited different saturation levels limited to forest types, sensor characteristics, and other factors [82,92]. That is, the sensitivity of backscatter signal to biomass decreases when the vegetation coverage reaches a certain biomass density [93]. To overcome such limitations, SAR and optical data were synergized to estimate biomass using empirical methods [91,94,95]. Indeed, multispectral signals can reflect the density of green vegetation and biochemical characteristics of leaves such as leaf type, chlorophyll, etc. Thus, the combination of two types of data performs better than single data, which can effectively enhance the saturation point of dense vegetation area [96,97].

Following previous studies, this study developed a semi-empirical retrieval method of above-ground forest fuel load (FL_{LAG}, indicating FFL, BFL, and SFL) estimation by introducing optical data as vegetation coverage information into WCM. Results showed that the new model which combines optical information and SAR data performed better than WCM using spaceborne SAR data alone (R^2 increased by 0.05–0.13 and RMSEr decreased by 5.8–12.9%). In addition, the new model did not exhibit any underestimation even when vegetation coverage reaches 65% and the total FL_{LAG} reaches about 183 Tons/ha. These results are also consistent with previous studies that the integration of SAR and optical data can alleviate the saturation problems [96,97]. The optical data is helpful to fitting forest transmissivity parameter because of its sensitivity to vegetation coverage, thus simulating remote sensing signals realistically. It is expected that this method can be applied in forest fire prevention and control in the future by providing spatiotemporal continuous distribution information of fuel load.

5.2. References for Subsequent Fuel Load Research

For FFL, it can be found that all optical variables selected in this study combined with HV polarization appropriately can effectively improve the inversion (e.g., increased R^2 or decreased RMSE), especially the Green band. Since the vegetation cover type of the study area is evergreen coniferous forest, even in autumn or winter, the leaves still contain chlorophyll, which has strong reflection characteristics to green light. For BFL, most fused data in proper combination mode contribute to inversion by decreasing RMSE, especially the SWIR band and RVI. For SFL, the combination of VIs and HV is better than that of spectral bands. Most VIs selected in this experiment contribute to SFL inversion since they are composed of near-infrared and short wave infrared which correlated to biomass [98]. The best performance of SFL inversion is “HV/NDII” since NDII is composed of NIR and SWIR2 bands which can indicate the vegetation water content and dry matter status [54,77–79].

On the whole, the WCM with L-band spaceborne SAR data has great potential in FL_{LAG} estimation. Although WCM exhibits an underestimation problem in dense mature forests, it performed well in sparse young forests. The addition of optical information effectively alleviates this problem, and has the most obvious effect on improving FFL, followed by BFL and SFL. This may be because leaves are distributed in the canopy and affect spectral reflectance by the physical scattering of light and the radiation energy absorbed by chemical absorption. Scattering is caused by the optical inhomogeneity of the plant surface and cell internal structure. Absorption is caused by pigments in visible light, water, and other biochemical substances such as cellulose, lignin, starch, and protein in the infrared band [99]. Although it is difficult for optical signals to penetrate the forest canopy to detect stem due to their limited penetration ability, the addition of optical data still improved the accuracy of SFL inversion. This may be attributed to the optical data
supplemented the vegetation coverage information in WCM. Moreover, the optical remote sensing data has a high temporal resolution which could describe the phenological character of vegetation and indicate the vegetation coverage well. Therefore, the addition of optical data can enhance the ability of new models in simulating the remote sensing signal of landscapes and effectively improve the estimation accuracy of FL_{AGL}.

5.3. Sources of Error and Uncertainty

There are some sources of uncertainty in this experiment. The field data calculated through forest structure parameters and allometric equations may introduce some error, since the allometric equations were derived from statistical relationships of limited field data which cannot describe all trees characteristics. The selection of “ground” candidate pixels may also introduce errors, since those pixels with less vegetation covered may be also selected as an approximation of pure “ground” pixel. In addition, the field data used in this study did not provide any soil moisture or information during sampling, which are determinants for radar backscatter. So, considering the uncertainties may be caused by soil moisture, we only selected the HV polarization to perform models in this study. In the HV polarimetric channel, volume scattering is the dominant scattering mechanism, and its scattering center of the interferometric phase is closer to the vegetation canopy top than that of the HH channel. The contribution of the ground layer is less important than that of the vegetation layers in HV polarimetric channel [100–105]. In future work, we will try to quantitatively consider the soil moisture in fuel load estimation by SAR (including cross and co-polarizations). The application of SAR data still needs a lot of work. In addition to the soil moisture, there is also the interference of rainfall and wind speed on SAR observation signals, especially in the analysis of time series SAR data [106,107]. In the future fuel load estimation work, we will further take these factors into account.

5.4. Model Operation and Optimization

The models in this study are semi-empirical models, which means that the model operation and optimization were all dependent on field data. Therefore, we divided the field measurements into calibrate data and validate data first. Then, the parameters in the models were calibrated by the calibration data. The model performance was validated by the validation data. During the LOOCV process, the parameters were decided by 28 groups of field data each time. That is, these parameters were determined by the calibration data without any artificial influence. Specifically, these parameters were fitted by the non-linear Levenberg–Marquardt algorithm which provides a unique solution for certain calibration data. In fact, the calibration of these parameters is definitely a process of optimization. Because the non-linear Levenberg–Marquardt algorithm is actually a parameter optimization algorithm, it can provide the most optimized parameters by fitting the model with the calibration data. That is, the parameters fitted by non-linear Levenberg–Marquardt algorithm were indeed the optimal one under the given calibration data. To sum up, the calibration and optimization of parameters is the same process, which is completed and determined by the non-linear Levenberg–Marquardt algorithm and the calibration data.

The inversion results of the forest fuel load also show that the calibration of the parameters meets expectations. In addition, the combination of SAR and optical data has better correlation performance than single SAR data. Take the Green band and FFL as an example, the $R^2$ is higher when combined with the HV and Green band (Figure 6).
6. Conclusions

Fuel load is a critical factor in fire ignition, spread, and intensity. Few previous studies used spaceborne SAR data to estimate fuel load. This study explored the potential of spaceborne L-band SAR data in live fuel load estimation using WCM and achieved reasonable performance ($R^2 = 0.64$ or higher and RMSEr = 35.3% or lower). However, the results of using WCM exhibited the underestimation phenomenon in dense mature forest. We presented a new model by introducing the optical data as vegetation coverage information into WCM, since the multispectral signals can reflect the density of green vegetation and biochemical characteristics of leaves such as leaf type, chlorophyll, etc. The new model performs better than WCM in FL$_{AGL}$ inversion with $R^2$ increased by 0.05–0.13 and RMSEr decreased by 5.8–12.9%. Besides, the new model did not show any underestimation problem even when vegetation coverage reaches 65% and the total FL$_{AGL}$ reaches about 183 Tons/ha. Specifically, the addition of most optical parameters (e.g., bands or VIs) especially the Green band in this study could certainly improve the inversion results of FFL since leaves are distributed in the forest canopy. As for BFL, the Green band, SWIR band and the RVI contributed to its estimation. As for SFL, the VIs (e.g., NDII, NDTI, and NDVI) have a more obvious effect than optical bands on SFL inversion. These results show the feasibility of using the proposed new model to estimate FL$_{AGL}$ based on optical and SAR data, which provides a good application prospect for forest fire risk warning, prevention, and extinguishing.

**Author Contributions:** Conceptualization, Y.L. and B.H.; methodology, Y.L.; validation Y.L. and B.H.; formal analysis, Y.L.; investigation, Y.L.; resources, B.H.; data curation, Y.L.; writing—original draft preparation, Y.L.; writing—review and editing, Y.L. and B.H.; visualization, Y.L.; supervision, B.H.; project administration, B.H.; funding acquisition, B.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by National Natural Science Foundation of China (Contract No. U20A2090).

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The BioSAR-2008 is an open-access dataset which is available in European Space Agency (http://eopi.esa.int).

**Acknowledgments:** The authors would like to appreciate the ONERA, ESA, and BioSAR team for providing the BioSAR 2008 datasets.

**Conflicts of Interest:** The authors declare no conflict of interest.
References


Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual authors and/or contributors and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.