forests

Article

Forest Canopy Water Content Monitoring Using Radiative Transfer Models and Machine Learning

Liang Liu [1], Shaoda Li [1,*], Wunian Yang [1,*], Xiao Wang [2], Xinrui Luo [1], Peilian Ran [1] and Helin Zhang [3,4]

[1] College of Earth Science, Chengdu University of Technology, Chengdu 610059, China; liuliangcdut@gmail.com (L.L.); 2019010040@stu.ceut.edu.cn (X.L.); ranpeilian@stu.cdut.edu.cn (P.R.)
[2] School of Architecture and Civil Engineering, Chengdu University, Chengdu 610106, China; wangxiao@cdu.edu.cn
[3] State Key Laboratory of Remote Sensing Science, Jointly Sponsored by Beijing Normal University and Institute of Remote Sensing and Digital Earth of Chinese Academy of Sciences, Beijing 100875, China; zhanghelinrs@gmail.com
[4] Beijing Engineering Research Center for Global Land Remote Sensing Products, Institute of Remote Sensing Science and Engineering, Faculty of Geographical Science, Beijing Normal University, Beijing 100875, China
* Correspondence: lisd@cdut.edu.cn (S.L.); ywn@cdut.edu.cn (W.Y.)

Abstract: Forests are facing various threats, such as drought, in the context of global climate change. Canopy water content (CWC) is a crucial indicator of forest water stress, mortality, and fire monitoring. However, previous studies on CWC have not adequately simulated forests with heterogeneous and discontinuous canopy structures. At the same time, there is a lack of field validation. This study retrieved the forest CWC across the contiguous U.S. (CONUS) with coupled radiative transfer models (RTMs) and the random forest (RF) algorithm. A Gaussian copula and prior knowledge were used for model parameterization. The results indicated that more accurate simulations of leaf trait dependencies and canopy structure characteristics lead to better CWC inversion. In addition, GeoSail, coupled with PROSPECT-5B, showed good performance (R^2 = 0.68, *RMSE* = 0.15 kg m^{-2}, *MAE* = 0.12 kg m^{-2}, *rRMSE* = 12.78%, *Bias* = $-$0.036 kg m^{-2}) for forest CWC retrieval. Large variation existed in forest CWC, spatiotemporally, and evergreen needle forest (ENF) showed strong CWC capacity. This study underscores the suitability of 3D RTMs for inversing the parameters of forest canopies.

Keywords: canopy water content; MODIS; random forest; copula; Google Earth Engine

check for updates

Citation: Liu, L.; Li, S.; Yang, W.; Wang, X.; Luo, X.; Ran, P.; Zhang, H. Forest Canopy Water Content Monitoring Using Radiative Transfer Models and Machine Learning. *Forests* **2023**, *14*, 1418. https://doi.org/10.3390/f14071418

Academic Editor: Giorgos Mallinis

Received: 7 June 2023
Revised: 4 July 2023
Accepted: 10 July 2023
Published: 11 July 2023

1. Introduction

Canopy water content (CWC), the amount of water stored in the vegetation canopy, is typically determined by multiplying the leaf water content by the canopy leaf area index (LAI). This calculation incorporates information about the leaf water condition and the canopy structure [1]. CWC is a critical parameter for assessing vegetation growth and monitoring drought stress. It is influenced by soil water supply and atmospheric demand [2]. Because of global climate change, drought events are becoming more frequent and severe, and future projections suggest they may become even more extreme [3,4]. In this context, the forest, as a critical component of terrestrial ecosystems, is facing severe threats from drought, leading to increased mortality rates [5–7]. Therefore, monitoring forest CWC using remote sensing technology can provide valuable insights into forest growth, water stress status, and drought response [8].

Optical and microwave sensors can effectively detect forest CWC [9,10]. Vegetation optical depth (VOD), retrieved from microwave emissions, is favored by many researchers for large-scale studies related to water or biomass, as it is unaffected by weather conditions and offers a short revisit period [11,12]. However, VOD provides vegetation information on both water and biomass with low spatial resolution [2]. In contrast, optical sensors offer a higher spatial resolution. In optical remote sensing, water absorbs radiation across the spectrum [13]. Additionally, the estimation of CWC generally relies on water-sensitive bands, such as 970 nm, 1200 nm, 1450 nm, 1600 nm, 1940 nm, and 2500 nm [1,14].

Approaches to CWC derived from remote sensing can be categorized into the empirical statistical method, physical model inversion, and hybrid methods [15]. The empirical statistical method estimates CWC by establishing regression relationships between field observations and band reflectance or vegetation indices [16]. This method is easy to understand and convenient to operate, but it relies on observational data and lacks generalization [1]. The physical model determines the response of CWC to canopy reflectance by simplifying the simulation of solar radiation and the transfer process in the soil–leaf–canopy–atmosphere [17]. Therefore, the accuracy of the physical process description and parameterization directly determines the result [18]. The radiative transfer model (RTM) is a commonly used physical model for CWC inversion. Both SAIL (Scattering by Arbitrarily Inclined Leaves) and GeoSail (a combination of SAIL and the Jasinski geometric model) belong to the RTM, and they generally simulate canopy spectrum characteristics by coupling with the leaf PROSPECT model [13,19–21]. In nature, remote sensing inversion is typically ill-posed. The ill-posed problem is characterized by using limited observations to estimate the current state of a complex surface system, and the amount of data information is less than the number of model parameters [22]. Therefore, challenges remain in the inversion of RTMs given the complex physical process description, the large number of parameters to be acquired, and the inherent ill-posed problem [23–25]. Hybrid-based approaches retrieve CWC by combining statistical methods with RTM inversion [26]. Practical inversion algorithms include iterative optimization [23], machine learning regression [26,27], and look-up tables [28].

In CWC inversion, previous studies commonly employed PROSAIL (leaf PROSPEPCT coupled with canopy SAIL) with unconstrained inputs, which may exacerbate the ill-posed problem [29]. Additionally, validation is typically conducted within the crop, grass, or shrub-covered area [1,15,30]. The development of plant trait datasets, such as Leaf Optical Properties Experiment 93 (LOPEX93) [31], the ANGERS Leaf Optical Properties Database (ANGERS) [32], the TRY database [33], and the National Ecological Observatory Network (NEON) [34], enables the further alleviation of ill-posed problems by integrating field observations with copula or ecological rule-based approaches. However, knowledge gaps persist in forest CWC inversion because of the complex canopy structure and the scarcity of field observations [21].

In recent years, Google Earth Engine (GEE) has experienced rapid growth and has been widely utilized for vegetation, agricultural, urban, hydrological, and land cover applications [35]. GEE encompasses remote sensing datasets, such as the MODerate-resolution Imaging Spectroradiometer (MODIS), Landsat, and Sentinel series, as well as meteorological and geophysical datasets [36,37]. Given its powerful computing performance and its purpose in geospatial data analysis, GEE serves as an excellent tool for monitoring long-term dynamic changes at various scales, from local to global [26,38].

In this study, our focus was on (1) parameterizing the model inputs with a Gaussian copula and prior knowledge, (2) building a CWC-retrieving model via simulations and RF regression, (3) evaluating the accuracy of CWC inversion based on field observations, and (4) capturing spatiotemporal variation using the GEE and MODIS products. We strive to provide a more efficient strategy for forest CWC inversion.

2. Materials

2.1. Study Area

We chose the contiguous U.S. (CONUS) as the study area because of the abundant field data collected by NEON. As shown in Figure 1, the eastern region of the CONUS is dominated by deciduous broad forests (DBFs), while the western region is characterized by evergreen needle forests (ENFs). The total ENF, EBF, and DBF areas across the CONUS are 6.67×10^5 km^2, 1.68×10^5 km^2, and 2.11×10^6 km^2, respectively. In evergreen forests, the dominant species include *Abies balsamea*, *Abies lasiocarpa*, *Juniperus virginiana*, *Picea glauca*, *Picea mariana*, *Pinus contorta*, *Pseudotsuga menziesii*, *Tsuga canadensis*, etc. On the other hand, *Acer rubrum*, *Betula alleghaniensis*, *Carya tomentosa*, *Fraxinus americana*, *Liquidambar styraciflua*, *Liriodendron tulipifera*, *Populus tremuloides*, *Quercus alba*, and *Ulmus americana*, among others, dominate deciduous forests [39].

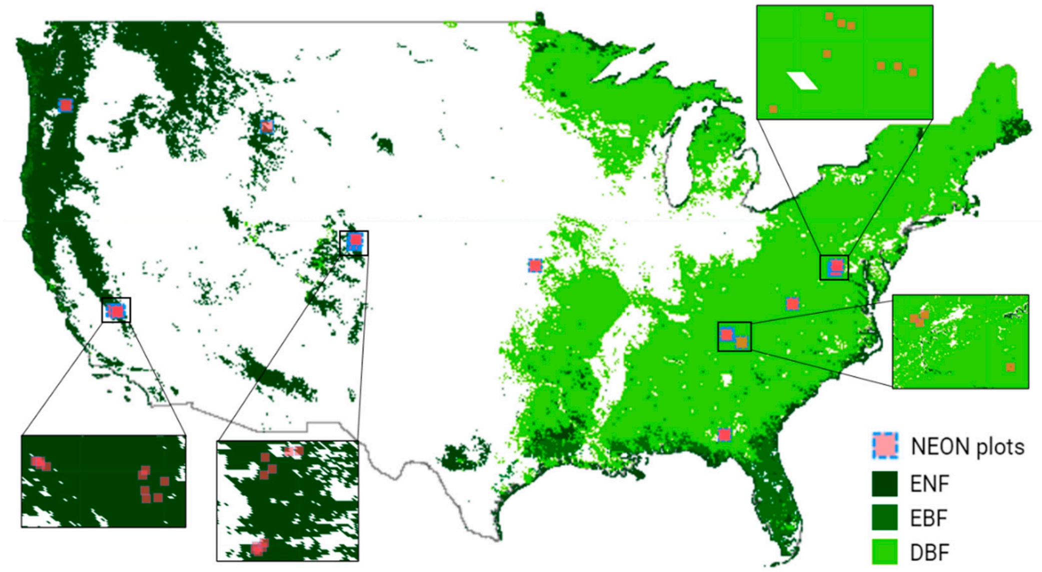

Figure 1. Forest distribution across the CONUS based on plant functional type classifications from the MODIS MCD12Q1 product. ENF: evergreen needle forest, EBF: evergreen broad forest, DBF: deciduous broad forest.

2.2. Field Observation Datasets

2.2.1. LOPEX93 Dataset

In 1993, a field campaign called the Leaf Optical Properties Experiment was implemented by the Joint Research Centre (JRC) of the European Commission in Ispra, Italy [31]. During the experiment, participants collected approximately 70 leaf samples representing more than 50 species from plants, crops, and trees at the JRC, and physical measurements and biochemical analyses were conducted [31]. During the experiment, various biochemical constituents such as leaf water, chlorophyll, lignin, cellulose, starch, and proteins (nitrogen) were measured. Directional–hemispherical reflectance and transmittance were recorded in the 400–2500 nm region [31]. The LOPEX93 database, created using this experiment's data, includes 330 records of biochemical constituents and leaf optics from 45 different species. It covers a wide range of leaf traits [40]. The LOPEX93 database provides variables such as leaf mass per area (LMA), equivalent water thickness (EWT), chlorophyll (Cab) concentration, and carotenoid (Car) concentration. As a result, it has been extensively utilized for parameterizing RTMs at the leaf scale [28,41–43].

2.2.2. ANGERS Dataset

ANGERS is a leaf optics dataset that connects leaf elements' visible/infrared spectral characteristics with biochemical analyses and physical measurements. It is part of an experiment conducted by the National Institute for Agricultural Research in Angers, France [32]. A total of 276 leaf samples, representing 39 different species, were collected for this database. Physical measurements and biochemical analyses were performed on these samples [32]. The ANGERS dataset contains directional–hemispherical reflectance and transmittance (400–2450 nm) obtained using an ASD FieldSpec spectrometer. It also includes measurements of biochemical constituents (water, carotenoid, chlorophyll, anthocyanin, etc.) obtained through biochemical analyses and physical measurements [32,43].

The ANGERS dataset, often used with LOPEX93, is utilized to parameterize leaf-level RTMs because both datasets provide the variables required by the models [28,41–43].

2.2.3. National Ecological Observatory Network (NEON)

The National Ecological Observatory Network (NEON), an observation facility funded by the National Science Foundation, aims to collect long-term ecological data to enhance understanding and forecast the effects of climate change and human activities on ecosystems across the United States [34]. The NEON established 47 terrestrial sites, ranging from deserts to tropical forests, equipped with the Terrestrial Instrument System, the Terrestrial Observation System, the Aquatic Observation System, the Aquatic Instrument System, and the Airborne Observation Platform [44].

The Terrestrial Observatory System determines the sampling frame at the site level based on the plot type and taxonomic group. Spatial sampling uses three plot types: tower, distributed, and gradient plots. Plots are selected using randomization, stratification, and sample number optimization, guided by the National Land Cover Database (NLCD) [45]. In forest ecosystems, tower base plots measure 40 m $\times$ 40 m and are divided into four 20 m $\times$ 20 m subplots. Vegetation structure, leaf area index (LAI), herbaceous clip harvest, plant diversity, and canopy foliar sampling were conducted in the core area. In contrast, soil sampling was implemented in the high-impact area [46]. Physical, chemical, and stable isotope data for plant foliage are measured through canopy sampling and analyses. Data such as LMA, EWT, Cab, and Car are provided through the Plant Foliar Traits (DP1.10026.001) data product [39].

The NEON provides Digital Hemispheric Photos (DHPs) for the Plot Vegetation data product (DP1.10017.001) to estimate the leaf area index (LAI) or the plant area index at the plot scale. The data product includes 180-degree images and metadata [47]. In the designated LAI sampling plots, 12 DHP photo points are positioned on a square grid with a spacing of four meters. The LAI of the plot was calculated as the average of the 12 points. DHPs are collected from the leaf-out-to-senescence period with a five-year interval for distributed plots and a two-week interval for towers [48].

A total of 1152 leaf samples from 195 evergreen, 134 deciduous, 24 mixed forest plots, and 12 grass plots were used for RTM parameterization. In addition, calculated LAIs from 28 evergreen and 20 deciduous forest plots were used for validation, combined with EWT.

2.3. MODIS Products

Five MODIS products, namely, MCD43A4, MCD43A2, MOD10A1, MYD10A1, and MCD12Q1, in version 6, available in GEE, were used for CWC retrieval in this study.

The MCD43A4 product contains seven bands (Table 1) and provides a 500 m daily nadir bidirectional reflectance distribution function (BRDF) adjusted reflectance (NBAR). It is produced based on 16-day composites [49]. With the BRDF, the reflectance is simulated as it is obtained from a nadir view in the local afternoon. Therefore, view angle effects are removed. In this study, all bands of the MCD43A4 were utilized as predictors for global CWC retrieval.

Table 1. Spectral specifications of the MCD43A4 bands.

Band	Wavelength
Nadir_Reflectance_Band1	620–670 nm
Nadir_Reflectance_Band2	841–876 nm
Nadir_Reflectance_Band3	459–479 nm
Nadir_Reflectance_Band4	545–565 nm
Nadir_Reflectance_Band5	1230–1250 nm
Nadir_Reflectance_Band6	1628–1652 nm
Nadir_Reflectance_Band7	2105–2155 nm

The MCD43A2 is a quality dataset for BRDF and albedo. At the same time, it contains information on land/water types and solar zenith angles for the MCD43A4 [50]. Here, we used the "BRDF_Albedo_LandWaterType" band to mask the water area and employed "BRDF_Albedo_LocalSolarNoon" to provide a solar zenith angle for pixel-based correction.

The MOD10A1 and MYD10A1 provide daily snow cover and snow albedo and related quality assessments from Terra and Aqua, respectively. The 500 m snow cover data were produced based on a normalized difference snow index mapping algorithm [51]. In this work, the snow cover band was utilized to mask the snow-covered area.

The MCD12Q1 product is a MODIS land cover type product that provides six 500 m annual land cover maps based on different classification schemes. It is produced using MODIS Terra and Aqua reflectance and supervised classifications [52]. In this study, the Annual Plant Functional Types classification was utilized, and the forest was divided into evergreen needle, evergreen broad, deciduous needle, and deciduous broad.

3. Methods

3.1. RTMs Selection and Parameterization

As per the flowchart shown in Figure 2, the spectra of the forest canopy were simulated using three leaf–canopy coupled radiative transfer models (RTMs): PROSPECT-5B + 4SAIL (PRO4SAIL), PROSPECT-5B + 4SAIL2 (PRO4SAIL2), and PROSPECT-5B + GeoSail (PROGeoSail). These coupled RTMs were parameterized based on the integration of observations and previous studies.

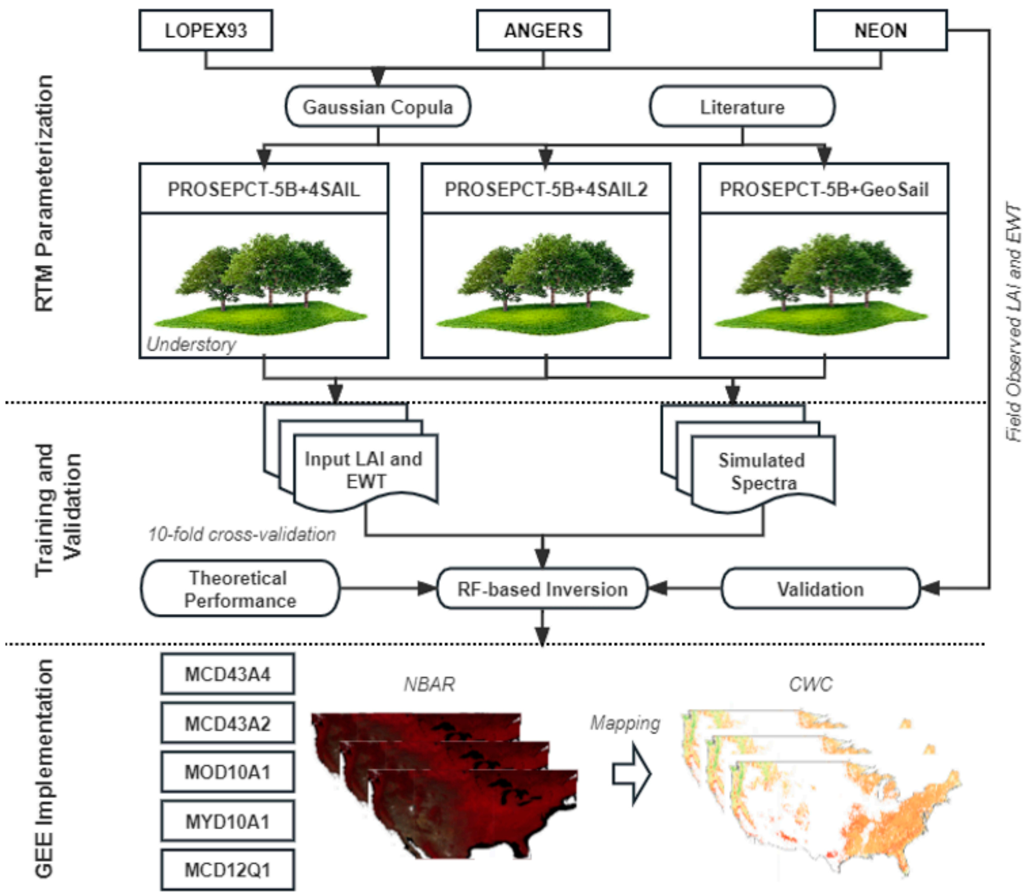

Figure 2. Flowchart of the forest CWC inversion.

The widely used PROSPECT-5 leaf optics model, developed based on the plate model, simulates reflectance and transmittance in the 400–2500 nm spectrum range [19]. In this case, the PROSPECT-5B version provided inputs for canopy RTMs, differentiating chlorophylls and carotenoids. The input variables include the leaf structure parameter and biochemical content (LMA, EWT, Cab, Cbp, and Car). The 4SAIL (an optimized version of SAIL) canopy bidirectional reflectance model, assuming horizontal uniformity, simulates the bidirectional reflectance factor based on the scattering and absorption of four upward/downward radiative fluxes [53].

Building upon the SAIL model, GeoSail incorporates the Jasinski geometric model to calculate a scene's illuminated and shadowed components separately [28]. GeoSail utilizes the optical properties of canopy components to simulate scene reflectance for heterogeneous and discontinuous forests. Within GeoSail, the reflectance and fraction of the illuminated canopy, as well as the illuminated and shadowed background, are used to calculate the forest canopy reflectance [20]

$$\rho_t = C\rho_c + S\rho_s + B\rho_b \tag{1}$$

where ρ_t is the total forest canopy reflectance; C is the canopy fraction; S is the illuminated background fraction; B is the shadowed background fraction; and ρ_c, ρ_s, and ρ_b are the reflectance of the canopy, illuminated, and shadowed background, respectively.

Specifically, the shadowed canopy can be calculated when the solar zenith angle exceeds the aspect angle and the crown shape is conical [20]:

$$\rho_t = (1 - C_s)C\rho_c + C_s\rho_c\tau_s + S\rho_s + B\rho_b \tag{2}$$

where C_s is the fraction of the shadowed canopy, and τ_s is the canopy transmittance.

4SAIL2 is a 3D, hybrid, two-layer canopy radiative transfer model that adheres to the four-stream concept. It incorporates the vertical leaf color gradient description, crown clumping, and numerical robustness from GeoSAIL, FLIM, and SAIL++ [54], respectively. Furthermore, extending a non-Lambertian soil BRDF model eliminated the limitation of 4SAIL2 to a Lambertian soil background. Thus, 4SAIL2 allows for modeling horizontal and vertical heterogeneities, resulting in more realistic simulations of forest canopies [55].

In GeoSail, the representation of the forest understory utilized the simulated canopy reflectance of grass instead of bare soil, as proposed by Quan et al. [28]. This replacement enhanced the realism of the understory reflectance simulation. The bi-hemispherical reflectance was employed, replacing the original Lambertian soil background, assuming that the canopy predominantly intercepts direct radiation.

The parameterization of RTMs relied on prior knowledge derived from field observations and the literature, and more details are shown in Table S1.

3.2. Gaussian Copula

A Gaussian copula was utilized to model the dependencies between leaf traits and constrain the parameterization of PROSPECT-5B. According to Sklar's theorem [56], there is a unique copula that holds for a multivariate absolutely continuous distribution [57]:

$$F\left(x_1, \ldots, x_p\right) = C\left(F_1(x_1), \ldots, F_p\left(x_p\right)\right) \tag{3}$$

where F is a multivariate distribution function, C is a copula, and $F_1(x_1), \ldots, F_p\left(x_p\right)$ are univariate distribution functions (marginal distributions).

The Gaussian copula takes the form [58]

$$C_{Gaussian}\,(u) = \Phi_R\left(\Phi^{-1}(u_1), \ldots, \Phi^{-1}\left(u_p\right)\right) \tag{4}$$

where $C_{Gaussian}$ is a Gaussian copula, Φ_R is the joint standard normal cumulative distribution function with positive definite covariance matrix R and zero mean, and Φ^{-1} is the standard normal quantile function. All marginal distributions and positive definite covariance matrixes allow for Gaussian copulas, and only pairwise dependencies were considered [57].

Gaussian copula modeling involves three steps: (1) calculating the univariate distributions, (2) obtaining the covariance matrix, and (3) modeling the relationships between variables using the copula function. The Gaussian copula was implemented using *openturns* in Python 3.9.5 in this study. Additionally, the univariate distributions were estimated using the Gaussian Kernel Density Estimation based on the data, and the covariance matrix was calculated using the Pearson method. Figures 3 and S1 show the distribution of Gaussian copula-based samples for trees and grass.

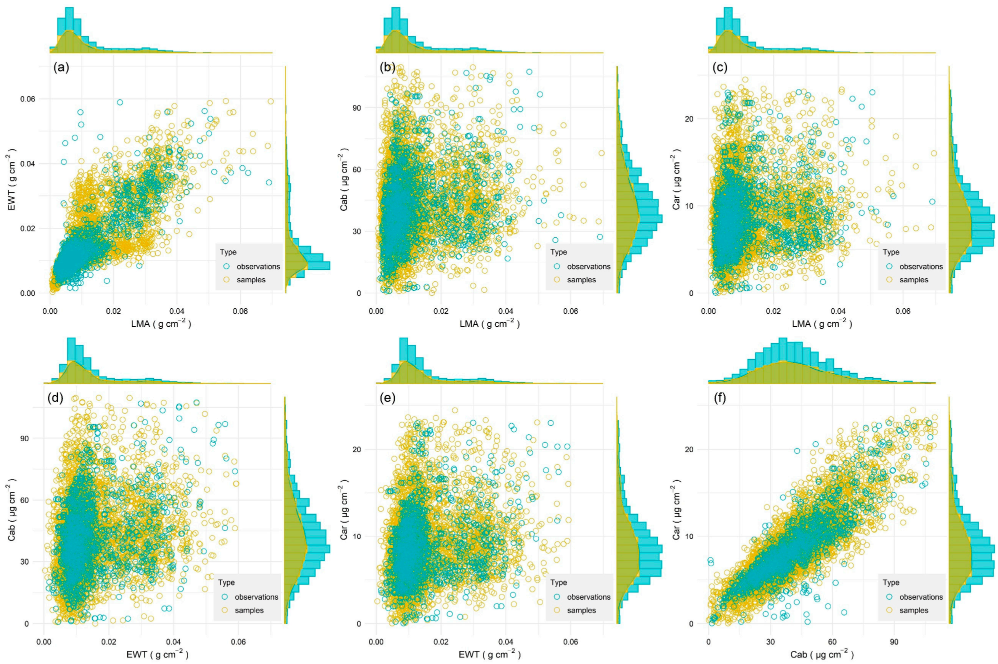

Figure 3. Constrained random samples of tree LMA, EWT, Cab, and Car based on field observations using Gaussian copula: (**a**) LMA vs. EWT; (**b**) LMA vs. Cab; (**c**) LMA vs. Car; (**d**) EWT vs. Cab; (**e**) EWT vs. Car; (**f**) Cab vs. Car.

3.3. RF-Based RTM Inversion

Random forest (RF) regression, a machine learning method, was employed for the inversion of RTMs. RF can overcome the autocorrelation of variables and overfitting, and it exhibits better prediction ability than a decision tree when dealing with many variables and observations because of its hierarchical structures and insensitivity to outliers [59,60]. RF is a combination of bagging algorithms and decision trees. It employs Bootstrap to extract n samples and constructs a decision tree for each sample [59].

Seven MODIS bands, simulated by the RTMs, were chosen as features for training an RF model. A 10-fold cross-validation was employed to evaluate the performance of the RF-based inversion model, using model efficiency (R^2), root-mean-square error ($RMSE$), and mean absolute error (MAE) as criteria for assessing accuracy [61]. Lastly, the forest CWC of the CONUS from 2017 to 2021 was generated in GEE using the RF model and MODIS products.

3.4. Solar Zenith Angle Correction

The geometric position of illumination and observation directly influence the reflectance of the ground target [62]. Geometric and radiometric corrections are essential for accurately retrieving land surface and vegetation parameters, particularly in time-series studies [63]. Regarding the MCD43A4, the bands were adjusted to the nadir view and the solar zenith angle at local noon. Thus, we employed the general method to correct the solar zenith angle [64]:

$$NBAR_\beta = \frac{NBAR_\alpha \times \cos\beta}{\cos\alpha} \quad (5)$$

where $NBAR_\alpha$ and $NBAR_\beta$ are the reflectance before and after correction, while α and β are the original and target solar zenith angles. In this work, the solar zenith angle is unified to 30°.

3.5. LAI Calculation from NEON DHPs

The GBOV (Ground-Based Observations for Validation) algorithm was employed to calculate field LAI based on the DHPs provided by the NEON [65]. To calculate LAIs from DHPs, the following steps were included: (1) image classification, (2) gap fraction calculation, and (3) LAI estimation.

For upward-facing DHPs, the Otsu automatic thresholding method was applied to the blue channel to classify the image cells as sky or vegetation canopy. Meanwhile, the image cells' background and vegetation canopy classification algorithm in downward-facing DHPs followed Meyer and Neto [66]. The method extracts vegetation canopy cells using two spectral indices: the excess green index (ExG, Equation (6)) and the excess red index (ExR, Equation (7) [67]). An image cell is classified as a vegetation canopy when the ExG is greater than the ExR. Otherwise, it is considered background. The ExG is calculated as follows:

$$ExG = 2DN_{\text{green}} - DN_{\text{red}} - DN_{\text{blue}} \tag{6}$$

and the ExR takes the form of

$$ExR = 1.4DN_{\text{red}} - DN_{\text{green}} \tag{7}$$

where DN_{red} is the red channel, DN_{green} is the green channel, and DN_{blue} is the blue channel.

Before gap fraction calculation, each image cell's zenith and azimuth angles need to be provided. The NEON DHPs were collected using a full-frame fisheye with a 180° field of view (FOV). Thus, assuming the zenith angle (α) of the image center is 0°, other cells can be calculated as follows [67]:

$$\alpha = \sqrt{(x - x_c)^2 + (y - y_c)^2} \frac{180}{\sqrt{w^2 + d^2}} \tag{8}$$

where x and y are the coordinates of a cell with an unknown zenith angle, x_c and y_c are the coordinates of the center cell, w is the image's width, and d is the height of the image. All these are expressed in pixels. The azimuth angle of a cell is calculated clockwise from the image center.

The classified images were divided into 36 patches with a 10° interval of azimuth angles to calculate the gap fraction. Only cells with zenith angles ranging from 52.5° to 62.5° were used to calculate the gap fraction of each patch. Therefore, the gap fraction, g_f, was calculated based on the patch as follows [67]:

$$g_f = \frac{n_{background}}{n_{total}} \tag{9}$$

where $n_{background}$ and n_{total} are the sky (or soil for downward images) cell amount and the total cell amount within a patch, respectively.

Then, the plant area index (PAI) was calculated as

$$PAI = \frac{-\overline{\ln g_f}}{0.93} \tag{10}$$

following Lang and Yueqin [68], where $\overline{\ln g_f}$ is the mean of the natural logarithm of the 36 patches' g_f values. Finally, PAI was equal to LAI for the non-woody area, while the LAI of the woody area was estimated according to

$$LAI = PAI_{up}(1 - \beta) + PAI_{\text{down}} \tag{11}$$

where PAI_{up} was calculated from the upward DHPs, PAI_{down} was calculated from the downward DHPs, and β is the area woody-to-total ratio. The means of β are 0.24, 0.16, and 0.18 for the deciduous forest, evergreen forest, and all [65], respectively.

3.6. Validation

The accuracy of CWC inversion was evaluated using NEON field observations. The NLCD land cover data were used to filter the representative NEON plots because of the difference in scale between the field plots and MODIS images. The criterion for screening the plots is that, within a 500 m buffer, one land cover type should occupy more than 90% of the area.

4. Results

4.1. Theoretical Performance of Coupled RTMs Inversion

The scatter plots between the predicted and model-set CWC values from the 10-fold cross-validation showed the theoretical performance of the three RF-based CWC inversion models. High model efficiency (R^2), low error ($RMSE$, MAE, and $rRMSE$), and weak bias ($Bias$) indicated that PRO4SAIL (R^2 = 0.9, $RMSE$ = 0.17 kg m^{-2}, MAE = 0.1 kg m^{-2}, $rRMSE$ = 3.78%, $Bias$ = 0.002 kg m^{-2}), PRO4SAIL2 (R^2 = 0.88, $RMSE$ = 0.17 kg m^{-2}, MAE = 0.1 kg m^{-2}, $rRMSE$ = 3.73%, $Bias$ = 0.001 kg m^{-2}), and PROGeoSail (R^2 = 0.87, $RMSE$ = 0.18 kg m^{-2}, MAE = 0.1 kg m^{-2}, $rRMSE$ = 4.08%, $Bias$ = 0 kg m^{-2}) had the theoretical ability to capture variations in CWC between different forest canopies well (Figure 4).

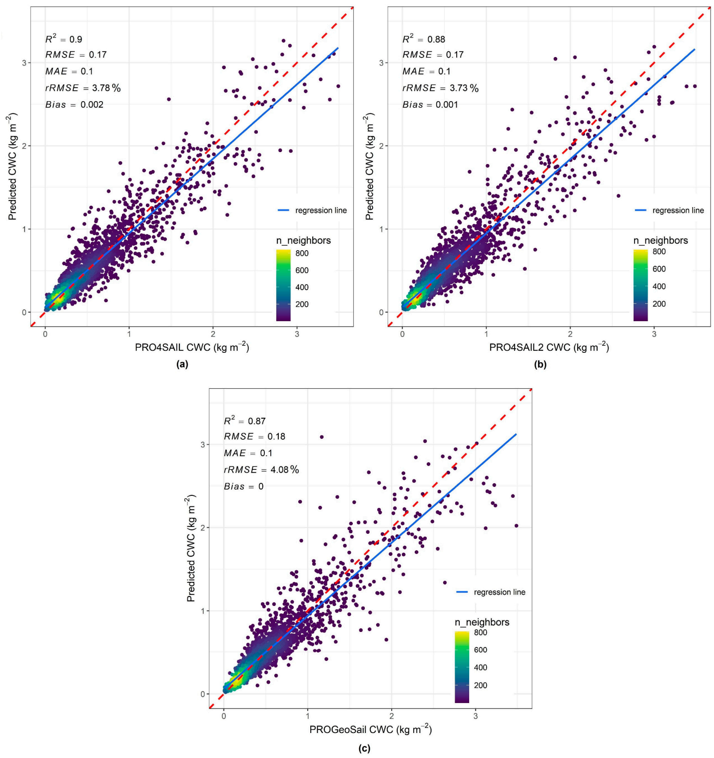

Figure 4. Theoretical performance of the RF-based inversion model over (**a**) PRO4SAIL, (**b**) PRO4SAIL2, and (**c**) PROGeoSail simulations. The 1:1 line is shown in dashed red, and the regression line is in solid blue.

4.2. Forest CWC Validation

The results of CWC validation, based on 48 NEON field observations collected at different sites during the leaf-out-to-senescence period from 2017 to 2021, are displayed in Figure 5. PRO4SAIL2 (R^2 = 0.72, $RMSE$ = 0.15 kg m^{-2}, MAE = 0.12 kg m^{-2}, $rRMSE$ = 12.04%)

and PROGeoSail (R^2 = 0.68, *RMSE* = 0.15 kg m^{-2}, *MAE* = 0.12 kg m^{-2}, *rRMSE* = 12.78%) had higher coefficients of determination and lower errors than PRO4SAIL (R^2 = 0.51, *RMSE* = 0.19 kg m^{-2}, *MAE* = 0.14 kg m^{-2}, *rRMSE* = 15.86%). PROGeoSail exhibited a negative bias of −0.036 kg m^{-2}, and PRO4SAIL2 was −0.081 kg m^{-2}, while the bias of PRO4SAIL was positive: 0.071 kg m^{-2}. Weaker bias, a lower error, and a higher coefficient of determination indicated that PROGeoSail had better accuracy.

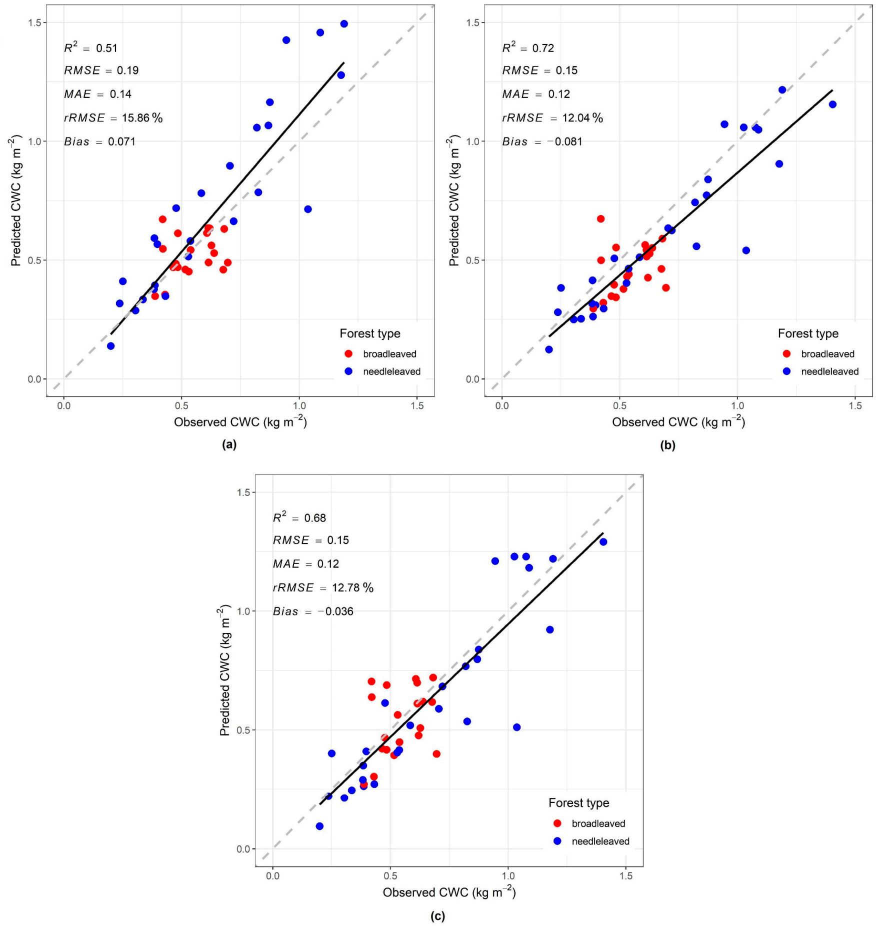

Figure 5. Prediction vs. field observations between (**a**) PRO4SAIL, (**b**) PRO4SAIL2, and (**c**) PROGeoSail-based hybrid inversion. The 1:1 line is shown in dashed gray, and the regression line is in solid black.

4.3. Spatiotemporal Variations in Forest CWC

Retrieval models were implemented in GEE using MODIS products to monitor spatiotemporal variations in forest CWC across the CONUS. The monthly mean forest CWC values from 2017 to 2021 were then generated (Figures 6, S2 and S3).

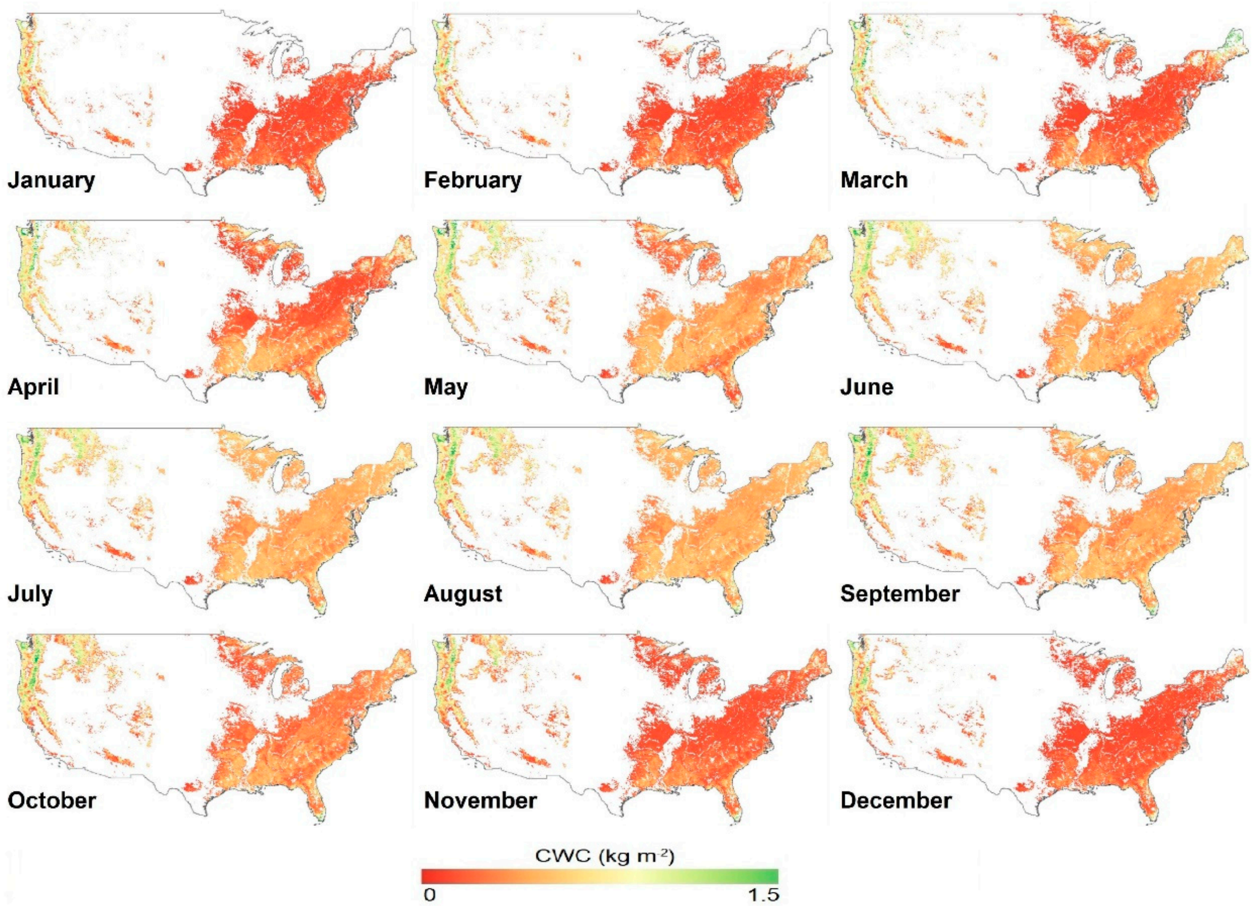

Figure 6. Monthly mean CWC spatial distribution of the CONUS at 500 m resolution in 2017–2021 based on PROGeoSail inversion.

Spatially, CWC exhibited variations between different forest types. Lower CWC values were observed in the DBF located in the eastern part of the CONUS, whereas higher CWC values were generally found in the ENF in the western region. Additionally, CWC dynamics varied with the month. As Figure 7 shows, throughout the year, forest CWC showed an increasing trend from January to August, followed by a decreasing trend until December. The DBF exhibited more significant seasonal variation compared with the ENF and the EBF.

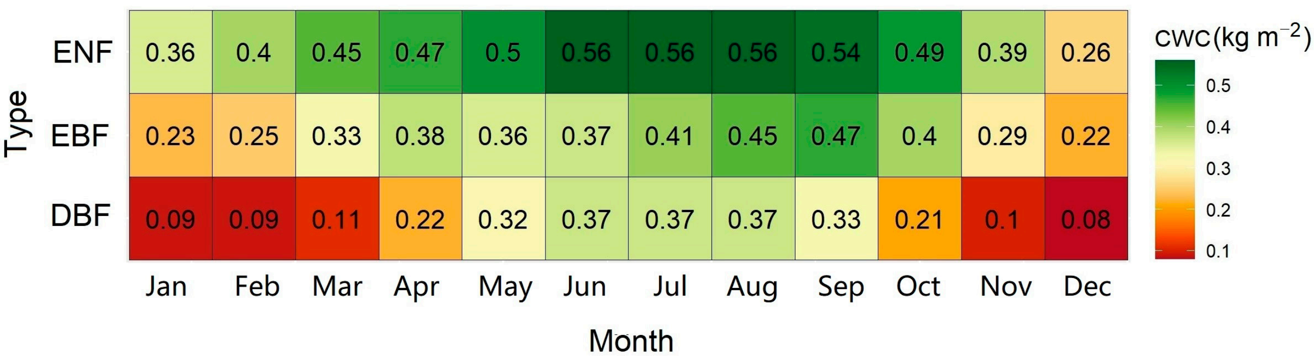

Figure 7. Monthly mean CWC of different forest types in 2017–2021 based on PROGeoSail inversion.

Furthermore, the three RTMs exhibited consistent spatial and temporal patterns. However, the retrieved CWC from PRO4SAIL was higher than that of PRO4SAIL2 and PROGeoSail, particularly in the ENF.

5. Discussion

Leaf traits are correlated, and forest canopy structures are heterogeneous and discontinuous. The accurate simulation of leaf and canopy characteristics can improve the accuracy of CWC inversion using coupled RTMs. Here, the 3D RTMs produced more accurate CWC retrieval than 1D RTMs. Our results demonstrated better theoretical performance than Campos-Taberner et al. [17] and superior validation accuracy compared with Trombetti et al. [29] and García-Haro et al. [1].

The successful inversion of forest CWC can be attributed to two key improvements. Firstly, in contrast to previous studies [21,29,69], we employed a Gaussian copula to constrain the leaf traits in the leaf-level RTM. This approach was adopted to alleviate the well-known ill-posed inversion problem in remote sensing. Prior studies [23–25,70] have highlighted the importance of constraining the fully independent variables using prior information to mitigate this issue. Quan et al. [25] and Reyes-Muñoz et al. [27] employed different methods to constrain the inputs of the RTMs for vegetation parameter inversion, resulting in favorable outcomes. As demonstrated in Figures 3 and S1, the leaf traits exhibit natural correlations with each other.

Secondly, a more accurate simulation of the forest canopy structure is needed. When retrieving canopy parameters, the widely used SAIL (or 4SAIL) model is generally applied to describe the canopy structure of grass, crops, shrubs, and forests [1,17,21,26,29,71] without considering their differences. However, the forest canopy exhibits heterogeneity and discontinuity, and modeling it solely with the turbid-based PROSAIL model may introduce bias. 4SAIL2 incorporates horizontal and vertical heterogeneities to achieve more realistic modeling by combining crown clumping and leaf color gradients [54]. GeoSail models crown clumping using Jasinski's parameterization [20]. Although including additional parameters may exacerbate the ill-posed problem, 3D canopy RTMs constrained by prior information, such as 4SAIL2 and GeoSail, are more suitable for forest modeling [28,55].

The forest CWC undergoes dynamic changes, particularly in the DBF. This variation was also observed by Trombetti et al. [29]. According to this definition, the CWC is determined by its EWT and LAI, both of which are influenced by plant phenology, climate, soil moisture, and other factors [2,72,73].

Our results also revealed a spatial pattern of higher CWC in the ENF of the CONUS and lower CWC in the DBF, which is consistent with the findings of Campos-Taberner et al. [17] and Reyes-Muñoz et al. [27]. This pattern can be attributed to the significant difference in leaf EWT, as indicated by field observations (Figures 5 and S4). However, the difference in CWC between ENF and DBF is more pronounced than in previous studies [1,17,26,29]. This difference could be explained by the variations in the selection and parameterization of RTMs [17], the inversion method [29,74], and remote sensing images [1,26]. Given the limitations of the data, more research on this difference is needed in the future.

In our study, a few limitations and uncertainties remain in the following aspects: Firstly, the field observations are limited and unevenly distributed in time and space. There are 170, 226, and 989 leaf samples from trees from LOPEX93, ANGERS, and the NEON, respectively, and the majority of these samples were obtained from deciduous trees. Limited and uneven leaf samples may lead to bias in EWT distribution and eventually affect the inversion of CWC. In addition, the NEON field observations of leaf traits and DHPs were primarily collected from leaf-out to senescence [48]. In this study, only 48 field CWC measurements (28 evergreen and 20 deciduous) were available for accuracy assessment, and there is a lack of observations for the winter period. Therefore, future research should prioritize additional efforts in forest field validation.

Secondly, this study focused on simulating the canopy while neglecting the leaf. PROSPECT-5, based on the plate model, was developed to simulate the reflectance and transmittance of broad leaves [19]. Although numerous studies have demonstrated the suitability of PROSPECT-5 for retrieving physiological parameters in coniferous species [13,28,75], the use of LIBERTY to simulate the reflectance and transmittance of coniferous species would provide greater robustness [76].

Thirdly, the simulated reflectance of the grass canopy was used instead of bare soil. This improvement can make the simulation of the forest more realistic [28]. However, the ground cover of CONUS forests is usually dominated by herbaceous and shrub layers [77]. Further study on the influence of forest underlayer on the inversion of CWC is needed.

Finally, compared with 4SAIL, 4SAIL2 and GeoSail require more inputs, which exacerbates the ill-posed problem. In this study, we employed prior information to constrain the inputs of the RTMs, including leaf-level parameters such as LAM, EWT, Cab, and Car, as well as the canopy cover fraction at the canopy level. However, there remain several independent parameters for the RTMs' inputs. Thus, future research should aim to impose further constraints on these free parameters based on additional prior information.

In this study, we estimated the forest CWC at a coarse spatial resolution. To study the species-specific characterization of CWC, hyperspectral, LiDAR, terahertz technology, and high-resolution images from unmanned aerial vehicles (cm to m level), Sentinel-2 (10 m), and Landsat 8–9 (30 m) could be used [61,78].

6. Conclusions

In this study, we aimed to compare the performance of different RTMs and provide an efficient strategy for forest CWC monitoring. A Gaussian copula and prior-knowledge-constrained parameters were input into coupled RTMs, and RF regression was used to retrieve forest CWC values across the CONUS. We found that accurately simulating the characteristics of leaf traits and canopy structures can improve the accuracy of CWC inversion. Additionally, 3D RTMs demonstrated better performance in retrieving forest CWC. We found that forest CWC showed a dynamically changed spatiotemporal pattern, and ENF had a strong CWC capacity. However, uncertainties still exist because of data limitations and the ill-posed problem. Future research should focus more on simulating forests with heterogeneous and discontinuous canopy structures and conducting field validation.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/f14071418/s1: Figure S1. Constrained random samples of herbaceous LMA, EWT, Cab, and Car based on field observations using a Gaussian copula for (a) LMA vs. EWT; (b) LMA vs. Cab; (c) LMA vs. Car; (d) EWT vs. Cab; (e) EWT vs. Car; (f) Cab vs. Car. Figure S2. Monthly mean CWC spatial distribution of CONUS at 500 m resolution in 2017–2021 based on PRO4SAIL inversion; Figure S3. Monthly mean CWC spatial distribution of CONUS at 500 m resolution in 2017–2021 based on PRO4SAIL2 inversion; Figure S4. Box plot for the EWT of NEON field observations comparing deciduous and evergreen forests. The statistical significance is shown by lowercase letters (a and b) at the level of $p < 0.01$ using one-way analysis of variance (ANOVA); Table S1. Parameterization of RTM inputs for forest CWC inversion.

Author Contributions: Conceptualization, S.L. and L.L.; methodology, L.L.; software, L.L.; validation, L.L., X.W., and X.L.; resources, S.L. and W.Y.; data curation, L.L. and P.R; writing—original draft preparation, L.L.; writing—review and editing, L.L., S.L., W.Y., X.W., X.L., P.R., and H.Z.; supervision, S.L. and W.Y.; funding acquisition, W.Y. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China, grant number 41671432.

Data Availability Statement: The MODIS products (MCD43A4, MCD43A2, MOD10A1, MYD10A1, and MCD12Q1) used here are available on GEE. The LOPEX93 and ANGERS datasets can be downloaded from http://opticleaf.ipgp.fr/index.php?page=database (accessed on 10 May 2023). The Plant Foliar Traits (DP1.10026.001) and Digital Hemispheric Photos (DHP) of the Plot Vegetation data products (DP1.10017.001) are accessible from https://data.neonscience.org (accessed on 10 May 2023). All the scripts are available upon request.

Acknowledgments: The National Ecological Observatory Network is a program sponsored by the National Science Foundation and operated under a cooperative agreement by Battelle. The Plant Foliar Traits (DP1.10026.001) and Digital Hemispheric Photos of Plot Vegetation (DP1.10017.001) data products are based in part upon work supported by the National Science Foundation through the NEON Program.

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

References

1. García-Haro, F.J.; Campos-Taberner, M.; Moreno, A.; Torbern Tagesson, H.; Camacho, F.; Martinez, B.; Sanchez, S.; Piles, M.; Camps-Valls, G.; Yebra, M.; et al. A global canopy water content product from AVHRR/Metop. *ISPRS J. Photogramm. Remote Sens.* **2020**, *162*, 77–93. [CrossRef]
2. Lyons, D.S.; Dobrowski, S.Z.; Holden, Z.A.; Maneta, M.P.; Sala, A. Soil moisture variation drives canopy water content dynamics across the western US. *Remote Sens. Environ.* **2021**, *253*, 112233. [CrossRef]
3. Zhou, S.; Zhang, Y.; Williams, A.P.; Gentine, P. Projected increases in intensity, frequency, and terrestrial carbon costs of compound drought and aridity events. *Sci. Adv.* **2019**, *5*, eaau5740. [CrossRef] [PubMed]
4. Bevacqua, E.; Zappa, G.; Lehner, F.; Zscheischler, J. Precipitation trends determine future occurrences of compound hot-dry events. *Nat. Clim. Change* **2022**, *12*, 350. [CrossRef]
5. Senf, C.; Pflugmacher, D.; Yang, Z.Q.; Sebald, J.; Knorn, J.; Neumann, M.; Hostert, P.; Seidl, R. Canopy mortality has doubled in Europe's temperate forests over the last three decades. *Nat. Commun.* **2018**, *9*, 4978. [CrossRef]
6. Jiao, W.Z.; Wang, L.X.; Smith, W.K.; Chang, Q.; Wang, H.L.; D'Odorico, P. Observed increasing water constraint on vegetation growth over the last three decades. *Nat. Commun.* **2021**, *12*, 3777. [CrossRef]
7. Hartmann, H.; Bastos, A.; Das, A.J.; Esquivel-Muelbert, A.; Hammond, W.M.; Martinez-Vilalta, J.; McDowell, N.G.; Powers, J.S.; Pugh, T.A.M.; Ruthrof, K.X.; et al. Climate Change Risks to Global Forest Health: Emergence of Unexpected Events of Elevated Tree Mortality Worldwide. *Annu. Rev. Plant Biol.* **2022**, *73*, 673–702. [CrossRef]
8. Brodrick, P.G.; Anderegg, L.D.L.; Asner, G.P. Forest Drought Resistance at Large Geographic Scales. *Geophys. Res. Lett.* **2019**, *46*, 2752–2760. [CrossRef]
9. Konings, A.G.; Yu, Y.; Xu, L.; Yang, Y.; Schimel, D.S.; Saatchi, S.S. Active microwave observations of diurnal and seasonal variations of canopy water content across the humid African tropical forests. *Geophys. Res. Lett.* **2017**, *44*, 2290–2299. [CrossRef]
10. Quemada, C.; Perez-Escudero, J.M.; Gonzalo, R.; Ederra, I.; Santesteban, L.G.; Torres, N.; Iriarte, J.C. Remote Sensing for Plant Water Content Monitoring: A Review. *Remote Sens.* **2021**, *13*, 2088. [CrossRef]
11. Holtzman, N.M.; Anderegg, L.D.L.; Kraatz, S.; Mavrovic, A.; Sonnentag, O.; Pappas, C.; Cosh, M.H.; Langlois, A.; Lakhankar, T.; Tesser, D.; et al. L-band vegetation optical depth as an indicator of plant water potential in a temperate deciduous forest stand. *Biogeosciences* **2021**, *18*, 739–753. [CrossRef]
12. Frappart, F.; Wigneron, J.P.; Li, X.J.; Liu, X.Z.; Al-Yaari, A.; Fan, L.; Wang, M.J.; Moisy, C.; Le Masson, E.; Lafkih, Z.A.; et al. Global Monitoring of the Vegetation Dynamics from the Vegetation Optical Depth (VOD): A Review. *Remote Sens.* **2020**, *12*, 2915. [CrossRef]
13. Jacquemoud, S.; Verhoef, W.; Baret, F.; Bacour, C.; Zarco-Tejada, P.J.; Asner, G.P.; François, C.; Ustin, S.L. PROSPECT+SAIL models: A review of use for vegetation characterization. *Remote Sens. Environ.* **2009**, *113*, S56–S66. [CrossRef]
14. Ceccato, P.; Flasse, S.; Tarantola, S.; Jacquemoud, S.; Gregoire, J.M. Detecting vegetation leaf water content using reflectance in the optical domain. *Remote Sens. Environ.* **2001**, *77*, 22–33. [CrossRef]
15. Casas, A.; Riano, D.; Ustin, S.L.; Dennison, P.; Salas, J. Estimation of water-related biochemical and biophysical vegetation properties using multitemporal airborne hyperspectral data and its comparison to MODIS spectral response. *Remote Sens. Environ.* **2014**, *148*, 28–41. [CrossRef]
16. Ceccato, P.; Flasse, S.; Gregoire, J.M. Designing a spectral index to estimate vegetation water content from remote sensing data—Part 2. Validation and applications. *Remote Sens. Environ.* **2002**, *82*, 198–207. [CrossRef]
17. Campos-Taberner, M.; Moreno-Martinez, A.; Javier Garcia-Haro, F.; Camps-Valls, G.; Robinson, N.P.; Kattge, J.; Running, S.W. Global Estimation of Biophysical Variables from Google Earth Engine Platform. *Remote Sens.* **2018**, *10*, 1167. [CrossRef]
18. Verrelst, J.; Malenovsky, Z.; Van der Tol, C.; Camps-Valls, G.; Gastellu-Etchegorry, J.P.; Lewis, P.; North, P.; Moreno, J. Quantifying Vegetation Biophysical Variables from Imaging Spectroscopy Data: A Review on Retrieval Methods. *Surv. Geophys.* **2019**, *40*, 589–629. [CrossRef]
19. Feret, J.-B.; François, C.; Asner, G.P.; Gitelson, A.A.; Martin, R.E.; Bidel, L.P.R.; Ustin, S.L.; le Maire, G.; Jacquemoud, S. PROSPECT-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments. *Remote Sens. Environ.* **2008**, *112*, 3030–3043. [CrossRef]
20. Huemmrich, K.F. The GeoSail model: A simple addition to the SAIL model to describe discontinuous canopy reflectance. *Remote Sens. Environ.* **2001**, *75*, 423–431. [CrossRef]
21. García-Haro, F.J.; Campos-Taberner, M.; Muñoz-Marí, J.; Laparra, V.; Camacho, F.; Sánchez-Zapero, J.; Camps-Valls, G. Derivation of global vegetation biophysical parameters from EUMETSAT Polar System. *ISPRS J. Photogramm. Remote Sens.* **2018**, *139*, 57–74. [CrossRef]
22. Wang, Y. Quantitative Remote Sensing Inversion in Earth Science: Theory and Numerical Treatment. In *Handbook of Geomathematics*; Freeden, W., Nashed, M.Z., Sonar, T., Eds.; Springer: Berlin/Heidelberg, Germany, 2010; pp. 785–812.

23. Combal, B.; Baret, F.; Weiss, M.; Trubuil, A.; Mace, D.; Pragnere, A.; Myneni, R.; Knyazikhin, Y.; Wang, L. Retrieval of canopy biophysical variables from bidirectional reflectance—Using prior information to solve the ill-posed inverse problem. *Remote Sens. Environ.* **2003**, *84*, 1–15. [CrossRef]
24. Yebra, M.; Chuvieco, E. Linking ecological information and radiative transfer models to estimate fuel moisture content in the Mediterranean region of Spain: Solving the ill-posed inverse problem. *Remote Sens. Environ.* **2009**, *113*, 2403–2411. [CrossRef]
25. Quan, X.; He, B.; Li, X. A Bayesian Network-Based Method to Alleviate the Ill-Posed Inverse Problem: A Case Study on Leaf Area Index and Canopy Water Content Retrieval. *Ieee Trans. Geosci. Remote Sens.* **2015**, *53*, 6507–6517. [CrossRef]
26. Martínez-Ferrer, L.; Moreno-Martínez, Á.; Campos-Taberner, M.; García-Haro, F.J.; Muñoz-Marí, J.; Running, S.W.; Kimball, J.; Clinton, N.; Camps-Valls, G. Quantifying uncertainty in high resolution biophysical variable retrieval with machine learning. *Remote Sens. Environ.* **2022**, *280*, 113199. [CrossRef]
27. Reyes-Muñoz, P.; Pipia, L.; Salinero-Delgado, M.; Belda, S.; Berger, K.; Estévez, J.; Morata, M.; Rivera-Caicedo, J.P.; Verrelst, J. Quantifying Fundamental Vegetation Traits over Europe Using the Sentinel-3 OLCI Catalogue in Google Earth Engine. *Remote Sens.* **2022**, *14*, 1347. [CrossRef]
28. Quan, X.; Yebra, M.; Riano, D.; He, B.; Lai, G.; Liu, X. Global fuel moisture content mapping from MODIS. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *101*, 102354. [CrossRef]
29. Trombetti, M.; Riaño, D.; Rubio, M.A.; Cheng, Y.B.; Ustin, S.L. Multi-temporal vegetation canopy water content retrieval and interpretation using artificial neural networks for the continental USA. *Remote Sens. Environ.* **2008**, *112*, 203–215. [CrossRef]
30. Djamai, N.; Zhong, D.T.; Fernandes, R.; Zhou, F.Q. Evaluation of Vegetation Biophysical Variables Time Series Derived from Synthetic Sentinel-2 Images. *Remote Sens.* **2019**, *11*, 1547. [CrossRef]
31. Hosgood, B.; Jacquemoud, S.; Andreoli, G.; Verdebout, J.; Pedrini, G.; Schmuck, G. *Leaf Optical Properties Experiment 93 (LOPEX93)*; Report EUR 16095 EN.; Joint Research Centre, European Commission: Ispra, Italy, 1994.
32. Jacquemound, S.; Bidel, L.; Francois, C.; Pavan, G. ANGERS Leaf Optical Properties Database. Available online: http://opticleaf.ipgp.fr/index.php?page=database (accessed on 10 May 2023).
33. Kattge, J.; Bonisch, G.; Diaz, S.; Lavorel, S.; Prentice, I.C.; Leadley, P.; Tautenhahn, S.; Werner, G.D.A.; Aakala, T.; Abedi, M.; et al. TRY plant trait database—enhanced coverage and open access. *Glob. Change Biol.* **2020**, *26*, 119–188. [CrossRef]
34. Kao, R.H.; Gibson, C.M.; Gallery, R.E.; Meier, C.L.; Barnett, D.T.; Docherty, K.M.; Blevins, K.K.; Travers, P.D.; Azuaje, E.; Springer, Y.P.; et al. NEON terrestrial field observations: Designing continental-scale, standardized sampling. *Ecosphere* **2012**, *3*, 1–17. [CrossRef]
35. Amani, M.; Ghorbanian, A.; Ahmadi, S.A.; Kakooei, M.; Moghimi, A.; Mirmazloumi, S.M.; Moghaddam, S.H.A.; Mahdavi, S.; Ghahremanloo, M.; Parsian, S.; et al. Google Earth Engine Cloud Computing Platform for Remote Sensing Big Data Applications: A Comprehensive Review. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2020**, *13*, 5326–5350. [CrossRef]
36. Hu, Q.; Yang, J.Y.; Xu, B.D.; Huang, J.X.; Memon, M.S.; Yin, G.F.; Zeng, Y.L.; Zhao, J.; Liu, K. Evaluation of Global Decametric-Resolution LAI, FAPAR and FVC Estimates Derived from Sentinel-2 Imagery. *Remote Sens.* **2020**, *12*, 912. [CrossRef]
37. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27. [CrossRef]
38. Sanchez-Ruiz, S.; Moreno-Martinez, A.; Izquierdo-Verdiguier, E.; Chiesi, M.; Maselli, F.; Gilabert, M.A. Growing stock volume from multi-temporal landsat imagery through google earth engine. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *83*, 101913. [CrossRef]
39. Weintraub, S. NEON User Guide to Plant Foliar Traits (DP1.10026.001). Available online: https://data.neonscience.org/documents/10179/2237401/NEON_CFC_userGuide_vD/3f1404e9-1a6b-9c13-7807-742893f88ef8 (accessed on 10 May 2023).
40. Hosgood, B.; Jacquemound, S.; Andreeoli, G.; Verdebout, J.; Pedrini, G.; Schmuck, G. Leaf Optical Properties Experiment Database (LOPEX93). Available online: http://opticleaf.ipgp.fr/index.php?page=database (accessed on 10 May 2023).
41. Majasalmi, T.; Bright, R.M. Evaluation of leaf-level optical properties employed in land surface models. *Geosci. Model Dev.* **2019**, *12*, 3923–3938. [CrossRef]
42. Yang, B.; He, Y.; Chen, W. A simple method for estimation of leaf dry matter content in fresh leaves using leaf scattering albedo. *Glob. Ecol. Conserv.* **2020**, *23*, e01201. [CrossRef]
43. Liu, L.; Song, B.; Zhang, S.; Liu, X. A Novel Principal Component Analysis Method for the Reconstruction of Leaf Reflectance Spectra and Retrieval of Leaf Biochemical Contents. *Remote Sens.* **2017**, *9*, 1113. [CrossRef]
44. Schimel, D. NEON Observatory Design. Available online: https://data.neonscience.org/documents/10179/11206/NEON.DOC.000001vD/af0626a5-7e95-4b0e-aaf1-c330d1b6a554 (accessed on 10 May 2023).
45. Barnett, D. TOS Science Design for Spatial Sampling. Available online: https://data.neonscience.org/documents/10179/1723439/NEON.DOC.000913vC/92b5da9f-8c35-f441-4413-c743ea4d01f7 (accessed on 10 May 2023).
46. Meier, C. NEON Terrestrial Observation System: Spatial and Temporal Sampling Strategy. Available online: https://data.neonscience.org/documents/10179/1723439/NEON.DOC.005108vB/1eda32ae-84f0-d263-0665-ab08c7d4eaca (accessed on 10 May 2023).
47. Meier, C. NEON User Guide to Digital Hemispheric Photos of Plot Vegetation (NEON.DP1.10017). Available online: https://data.neonscience.org/documents/10179/2237401/NEON_dhp_userGuide_vE/a7d3ae35-5595-cd1d-f1ad-aeed03bdb62e (accessed on 10 May 2023).
48. Meier, C. TOS Protocol and Procedure: DHP—Measurement of Leaf Area Index. Available online: https://data.neonscience.org/documents/10179/1883155/NEON.DOC.014039vM/94cd2218-598c-935f-829e-5003d7b25950 (accessed on 10 May 2023).

49. Schaaf, C.; Wang, Z. MCD43A4 MODIS/Terra+Aqua BRDF/Albedo Nadir BRDF Adjusted Ref Daily L3 Global—500m V006. Available online: https://lpdaac.usgs.gov/products/mcd43a4v006 (accessed on 20 June 2022).
50. Schaaf, C.; Wang, Z. MCD43A2 MODIS/Terra+Aqua BRDF/Albedo Quality Daily L3 Global—500m V006. Available online: https://lpdaac.usgs.gov/products/mcd43a2v006 (accessed on 20 June 2022).
51. Hall, D.K.; Riggs, G.A. MODIS/Terra Snow Cover Daily L3 Global 500m SIN Grid, Version 6. Available online: https://doi.org/10.5067/MODIS/MOD10A1.006 (accessed on 10 May 2023).
52. Friedl, M.; Sulla-Menashe, D. MCD12Q1 MODIS/Terra+Aqua Land Cover Type Yearly L3 Global 500m SIN Grid V006. Available online: https://lpdaac.usgs.gov/products/mcd12q1v006 (accessed on 10 May 2023).
53. Verhoef, W.; Jia, L.; Xiao, Q.; Su, Z. Unified optical-thermal four-stream radiative transfer theory for homogeneous vegetation canopies. *Ieee Trans. Geosci. Remote Sens.* **2007**, *45*, 1808–1822. [CrossRef]
54. Verhoef, W.; Bach, H. Coupled soil-leaf-canopy and atmosphere radiative transfier modeling to simulate hyperspectral multi-angular surface reflectance and TOA radiance data. *Remote Sens. Environ.* **2007**, *109*, 166–182. [CrossRef]
55. le Maire, G.; Marsden, C.; Verhoef, W.; Ponzoni, F.J.; Lo Seen, D.; Begue, A.; Stape, J.-L.; Nouvellon, Y. Leaf area index estimation with MODIS reflectance time series and model inversion during full rotations of Eucalyptus plantations. *Remote Sens. Environ.* **2011**, *115*, 586–599. [CrossRef]
56. Sklar, A. *Fonctions de Répartition à n Dimensions et Leurs Marges*; Publications de L'institut de Statistique de L'université de Paris: Paris, France, 1959; Volume 8.
57. Žežula, I. On multivariate Gaussian copulas. *J. Stat. Plan. Inference* **2009**, *139*, 3942–3946. [CrossRef]
58. Van de Vyver, H.; Van den Bergh, J. The Gaussian copula model for the joint deficit index for droughts. *J. Hydrol.* **2018**, *561*, 987–999. [CrossRef]
59. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
60. Tang, X.; Fan, S.; Zhang, W.; Gao, S.; Chen, G.; Shi, L. Global variability in belowground autotrophic respiration in terrestrial ecosystems. *Earth Syst. Sci. Data* **2019**, *11*, 1839–1852. [CrossRef]
61. de Almeida, C.T.; Galvao, L.S.; de Oliveira Cruz e Aragao, L.E.; Henry Balbaud Ometto, J.P.; Jacon, A.D.; de Souza Pereira, F.R.; Sato, L.Y.; Lopes, A.P.; Lima de Alencastro Graca, P.M.; Silva, C.V.D.J.; et al. Combining LiDAR and hyperspectral data for aboveground biomass modeling in the Brazilian Amazon using different regression algorithms. *Remote Sens. Environ.* **2019**, *232*, 111323. [CrossRef]
62. Widlowski, J.L.; Taberner, M.; Pinty, B.; Bruniquel-Pinel, V.; Disney, M.; Fernandes, R.; Gastellu-Etchegorry, J.P.; Gobron, N.; Kuusk, A.; Lavergne, T.; et al. Third Radiation Transfer Model Intercomparison (RAMI) exercise: Documenting progress in canopy reflectance models. *J. Geophys. Res. Atmos.* **2007**, *112*, D09111. [CrossRef]
63. Breunig, F.M.; Galvao, L.S.; dos Santos, J.R.; Gitelson, A.A.; de Moura, Y.M.; Teles, T.S.; Gaida, W. Spectral anisotropy of subtropical deciduous forest using MISR and MODIS data acquired under large seasonal variation in solar zenith angle. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *35*, 294–304. [CrossRef]
64. Shahi, A.P.; Rai, P.K.; Rabi ul, I.; Mishra, V.N. Chapter 5—Remote sensing data extraction and inversion techniques: A review. In *Atmospheric Remote Sensing*; Kumar Singh, A., Tiwari, S., Eds.; Elsevier: Amsterdam, The Netherlands, 2023; pp. 85–104.
65. Brown, L.A.; Fernandes, R.; Djamai, N.; Meier, C.; Gobron, N.; Morris, H.; Canisius, F.; Bai, G.; Lerebourg, C.; Lanconelli, C.; et al. Validation of baseline and modified Sentinel-2 Level 2 Prototype Processor leaf area index retrievals over the United States. *ISPRS J. Photogramm. Remote Sens.* **2021**, *175*, 71–87. [CrossRef]
66. Meyer, G.E.; Neto, J.C. Verification of color vegetation indices for automated crop imaging applications. *Comput. Electron. Agric.* **2008**, *63*, 282–293. [CrossRef]
67. Dash, L.B.H.M.J. *Algorithm Theoretical Basis Document—Vegetation Products RM4 (Tcanopy), RM6 (FIPAR) and RM7 (LAI) and FCOVER*; University of Southampton: Southampton, UK, 2021.
68. Lang, A.R.G.; Yueqin, X. Estimation of leaf area index from transmission of direct sunlight in discontinuous canopies. *Agric. For. Meteorol.* **1986**, *37*, 229–243. [CrossRef]
69. Boren, E.J.; Boschetti, L. Landsat-8 and Sentinel-2 Canopy Water Content Estimation in Croplands through Radiative Transfer Model Inversion. *Remote Sens.* **2020**, *12*, 2803. [CrossRef]
70. de Sá, N.C.; Baratchi, M.; Hauser, L.T.; van Bodegom, P. Exploring the Impact of Noise on Hybrid Inversion of PROSAIL RTM on Sentinel-2 Data. *Remote Sens.* **2021**, *13*, 648. [CrossRef]
71. Estevez, J.; Salinero-Delgado, M.; Berger, K.; Pipia, L.; Rivera-Caicedo, J.P.; Wocher, M.; Reyes-Munoz, P.; Tagliabue, G.; Boschetti, M.; Verrelst, J. Gaussian processes retrieval of crop traits in Google Earth Engine based on Sentinel-2 top-of-atmosphere data. *Remote Sens. Environ.* **2022**, *273*, 112958. [CrossRef] [PubMed]
72. Joswig, J.S.; Wirth, C.; Schuman, M.C.; Kattge, J.; Reu, B.; Wright, I.J.; Sippel, S.D.; Ruger, N.; Richter, R.; Schaepman, M.E.; et al. Climatic and soil factors explain the two-dimensional spectrum of global plant trait variation. *Nat. Ecol. Evol.* **2022**, *6*, 36–50. [CrossRef]
73. Savoy, P.; Mackay, D.S. Modeling the seasonal dynamics of leaf area index based on environmental constraints to canopy development. *Agric. For. Meteorol.* **2015**, *200*, 46–56. [CrossRef]
74. Asner, G.P.; Brodrick, P.G.; Anderson, C.B.; Vaughn, N.; Knapp, D.E.; Martin, R.E. Progressive forest canopy water loss during the 2012–2015 California drought. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E249–E255. [CrossRef]

75. Cheng, Y.-B.; Zarco-Tejada, P.J.; Riano, D.; Rueda, C.A.; Ustin, S.L. Estimating vegetation water content with hyperspectral data for different canopy scenarios: Relationships between AVIRIS and MODIS indexes. *Remote Sens. Environ.* **2006**, *105*, 354–366. [CrossRef]
76. Dawson, T.P.; Curran, P.J.; Plummer, S.E. LIBERTY—Modeling the effects of leaf biochemical concentration on reflectance spectra. *Remote Sens. Environ.* **1998**, *65*, 50–60. [CrossRef]
77. Krebs, M.A.; Reeves, M.C.; Baggett, L.S. Predicting understory vegetation structure in selected western forests of the United States using FIA inventory data. *For. Ecol. Manag.* **2019**, *448*, 509–527. [CrossRef]
78. Gente, R.; Koch, M. Monitoring leaf water content with THz and sub-THz waves. *Plant Methods* **2015**, *11*, 15. [CrossRef] [PubMed]