Spatially Distributed Overstory and Understory Leaf Area Index Estimated from Forest Inventory Data

Sara A. Goeking [1],* and David G. Tarboton [2]

1 Forest Inventory and Analysis Program, Rocky Mountain Research Station, USDA Forest Service, Ogden, UT 84321, USA
2 Utah Water Research Laboratory, Utah State University, Logan, UT 84322, USA
* Correspondence: sara.goeking@usda.gov

Abstract: Forest change affects the relative magnitudes of hydrologic fluxes such as evapotranspiration (ET) and streamflow. However, much is unknown about the sensitivity of streamflow response to forest disturbance and recovery. Several physically based models recognize the different influences that overstory versus understory canopies exert on hydrologic processes, yet most input datasets consist of total leaf area index (LAI) rather than individual canopy strata. Here, we developed stratum-specific LAI datasets with the intent of improving the representation of vegetation for ecohydrologic modeling. We applied three pre-existing methods for estimating overstory LAI, and one new method for estimating both overstory and understory LAI, to measurements collected from a probability-based plot network established by the US Forest Service's Forest Inventory and Analysis (FIA) program, for a modeling domain in Montana, MT, USA. We then combined plot-level LAI estimates with spatial datasets (i.e., biophysical and remote sensing predictors) in a machine learning algorithm (random forests) to produce annual gridded LAI datasets. Methods that estimate only overstory LAI tended to underestimate LAI relative to Landsat-based LAI (mean bias error ≥ 0.83), while the method that estimated both overstory and understory layers was most strongly correlated with Landsat-based LAI ($r^2 = 0.80$ for total LAI, with mean bias error of -0.99). During 1984–2019, interannual variability of understory LAI exceeded that for overstory LAI; this variability may affect partitioning of precipitation to ET vs. runoff at annual timescales. We anticipate that distinguishing overstory and understory components of LAI will improve the ability of LAI-based models to simulate how forest change influences hydrologic processes.

Keywords: forest inventory; leaf area index; hydrologic modeling; overstory; understory

1. Introduction

Forest cover type, density, and dynamics (i.e., change over time) affect the relative magnitudes of ecohydrologic fluxes such as evapotranspiration (ET) and streamflow [1–3]. Thus, water supply is influenced, at least in part, by how vegetation partitions precipitation into ET vs. streamflow. Forest canopies exert particularly strong effects on this partitioning because they intercept precipitation, which is often then lost to evaporation rather than accumulating as seasonal snowpack or reaching the ground surface to contribute to either runoff or recharge [1,3–5]. After more than a century of research into vegetation–streamflow linkages [6], questions remain about how future forest disturbance, recovery, or conversion to nonforest vegetation, as well as differences between overstory and understory influences on hydrologic processes, may affect future water supply [1,4,7].

While observational studies of individual or paired catchments provide insights to the mechanisms of hydrologic response to vegetation change [8], for logistical and practical reasons, the ability to study individual or paired catchments in detail is largely confined to small watersheds [6]. Broad-scale questions about hydrologic response to forest change can be answered by using process-based models that are capable of representing vegetation in multiple canopy strata, such as the Regional Hydro-Ecologic Simulation System

(RHESSys) [9] and Distributed Hydrology-Soil-Vegetation Model (DHSVM) [10]. Such models, and indeed most quantitative relationships of hydrologic processes to vegetation states, express vegetation density in terms of leaf area index (LAI) [11]. Often these models are applied using coarse-resolution total LAI derived from remote sensing platforms such as MODIS (e.g., [12]) if higher-resolution observations are not available. An advantage of LAI datasets based on spectral remote sensing is their wall-to-wall spatial coverage and fine temporal resolution, i.e., interannual or seasonal variability, but a disadvantage is that they represent total LAI and do not distinguish overstory from understory LAI. To capitalize on the ability of hydrologic models to provide insights into the linkage between water resources and forest disturbance or vegetation type changes, better representations of actual forest vegetation strata (i.e., overstory vs. understory) are required.

In contrast to remote sensing-based LAI, ground-based observations have the potential to distinguish overstory from understory LAI and also attribute causes of change such as type and severity of disturbance. It is possible that interpolation models, e.g., random forests or other machine learning algorithms, that are developed from sparse plot-based LAI and spatially continuous predictors (e.g., reflectance, elevation, aspect) may produce improvements in gridded LAI estimates representing overstory and understory strata, compared to estimates of total LAI produced using remote sensing data alone. Ground-based observations ideally would be unbiased and representative of the full range of variability of forest characteristics that occurs within a domain of interest, but sites are often selected for a specific purpose that may lead to biased inference when taken as generally representative [13]. Several countries, including the USA, are monitored continuously by strategic national forest inventories that conduct probabilistic and repeated sampling of permanent plots that could provide inputs for plot-level LAI estimates. In the USA, the US Forest Service's Forest Inventory and Analysis (FIA) program monitors a plot network of over 300,000 plots nationwide, across all forest types and ownership categories with a mean plot spacing of about 5 km [14]. FIA collects detailed information on trees and the overstory canopy, understory vegetation, and causes and timing of disturbances [15]. However, FIA and other national-scale forest inventories do not measure LAI [16,17]. Thus, a method of using existing inventory data to estimate plot-scale LAI is needed.

In the absence of ground-based LAI measurements such as those obtained via techniques such as light-sensing instruments or hemispherical photography, previous research has derived LAI estimates based on other available measurements such as tree canopy cover or canopy closure (e.g., [18,19]). Methods and mathematical forms typically fall into the following categories, where the choice depends on the intended purpose: (1) a linear scaling factor relating canopy cover to LAI, where a local maximum LAI is imposed on the scaling factor (e.g., [18]); (2) an exponential function relating tree canopy cover to LAI for the purposes of modeling snow cover (presented as Equation (4) in [19], derived from [20]); or (3) allometric equations based on destructive measurements and detailed dissections of a small sample of trees (e.g., [21]). The main limitation of these ground-based methods for measuring or estimating LAI is that they are spatially discontinuous and thus require interpolation to produce gridded LAI datasets for use in spatially distributed hydrologic models.

This study bridges the gap between remote sensing and ground-based estimates of LAI using data from the USA national forest inventory. The objectives of this study were, first, to use multiple LAI estimation methods to estimate plot-scale LAI for overstory and understory canopy strata from standard forestry measurements at FIA plots and, second, to combine plot-scale LAI data with spectral reflectance and other gridded variables in a machine-learning algorithm to produce spatially and temporally explicit maps of overstory and understory LAI on an annual basis. We compared alternative methods for estimating plot-level LAI, developed a machine learning method for interpolating yearly gridded overstory and understory LAI across large watersheds, and compared gridded estimates of annual LAI to Landsat-based LAI. Finally, our third objective was to examine annual

time series of overstory and understory LAI for two test watersheds that have experienced natural disturbances but no land-use change.

2. Materials and Methods

2.1. Study Area

The study area encompassed the South Fork Flathead River and Middle Fork Flathead River watersheds of northwestern Montana, USA (Figure 1). Most of the South Fork watershed and part of the Middle Fork watershed are within the Bob Marshall Wilderness Area, which was designated in 1964 and precluded any substantial vegetation management since that time. The South Fork and Middle Fork are 83% and 75% forested, respectively, with mean elevations of 1870 and 1722 m; areas of 3002 and 2914 km^2; mean annual temperatures of 2.63 and 2.46 °C; and annual precipitation of 1248 and 1268 mm [22].
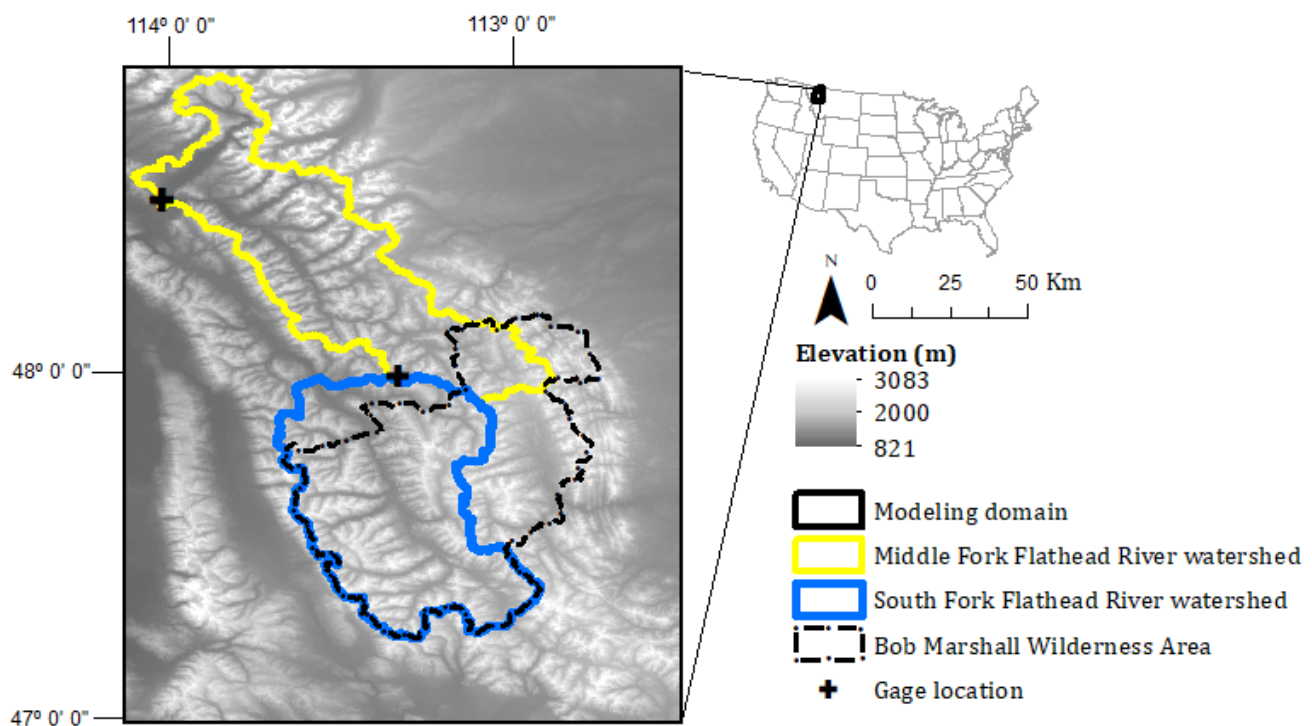

Figure 1. Digital elevation map of the study area. Domain for modeling leaf area index (LAI) is shown by the outer rectangle; domains for time series analysis of LAI and evapotranspiration are the South Fork Flathead River and Middle Fork Flathead River watersheds (outlined in blue and yellow, respectively).

Figure 2 illustrates how this study's workflows involve plot data vs. gridded data as well as comparison of plot-scale vs. gridded LAI estimates.
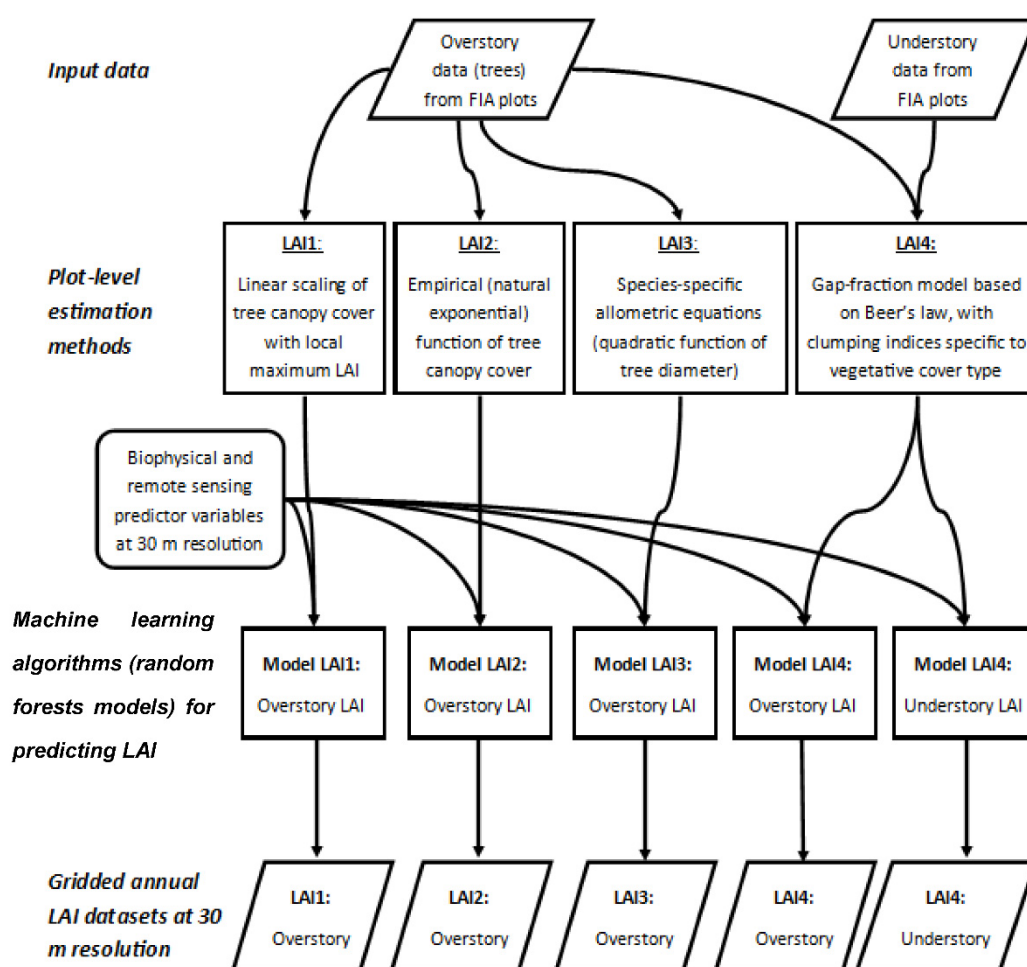
Figure 2. Data sources and methods used to produce plot-level LAI and gridded LAI datasets.

2.2. Field Sampling Protocols

In Montana, FIA began measuring permanent plots in 2003. Plots were established in a semisystematic grid where each plot represents approximately 2400 hectares, with a remeasurement period of 10 years, and a representative sample of 10% of all permanent plots measured each year [15]. Each plot consists of four subplots, each with a radius of 7.3 m, where one subplot is centrally located and the other three subplots are established 36.6 m from the first subplot's center at azimuths of 0, 120, and 240 degrees (Figure 3). On each subplot, field crews measure and record information about the site (e.g., slope and aspect), including tree canopy cover as determined using four line transects per subplot, or 16 total transects per plot; ocular estimates of percent cover of understory vegetation by growth habit (tree, shrub, graminoid, or nongraminoid herbaceous vegetation), where "understory" includes any cover by any of the aforementioned growth habits, regardless of whether those plants occur under trees or not; and individual live and dead trees that are at least 12.7 cm at a height of 1.35 m [17], which is commonly known as breast height and the measurement is thus known as diameter at breast height (DBH). Characteristics of trees with DBH < 12.7 cm are measured on a 2.1 m radius subsample of each subplot [17]. Data collection protocols and definitions are described in full detail in the US Department of Agriculture's published field manual for the national forest inventory [17]. Data collection occurs within the growing season, which for this region is typically June–September.
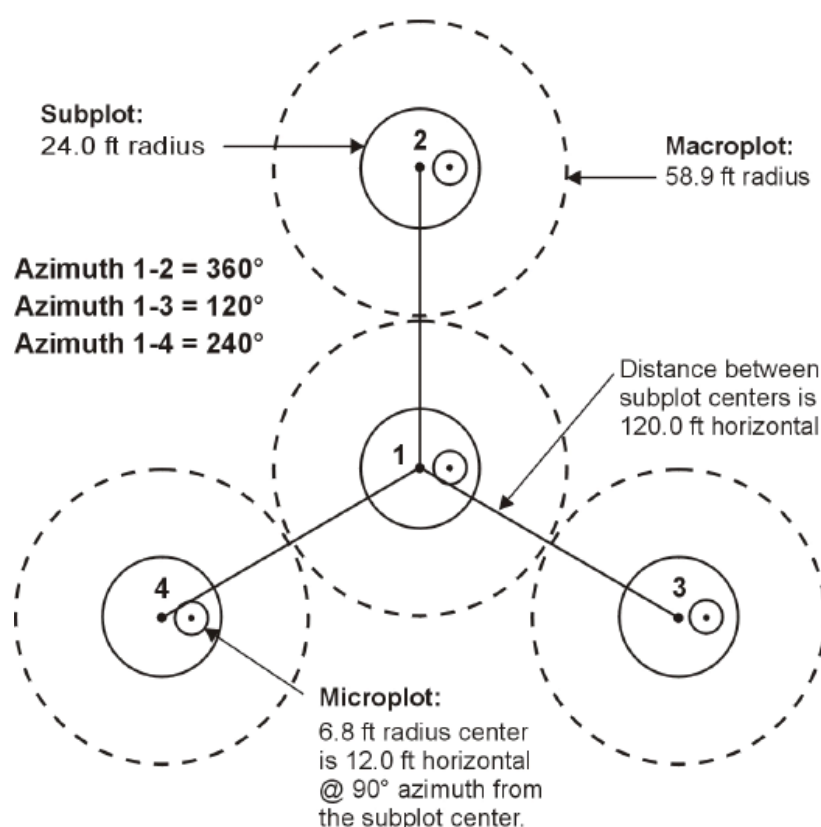
Figure 3. FIA plot configuration [15]. Metric units: subplots have a radius of 7.3 m with subplot centers 36.6 m apart; microplots have a radius of 2.1 m and are located 3.6 m from subplot centers; and macroplots were not used in this study.

FIA defines forest as land that either currently contains, or previously supported, at least 10% tree cover, with a minimum forest patch size of 0.4 ha. Thus, plots that burned, or otherwise lost all tree cover due to natural disturbance, are still defined as "forest" but may have zero percent tree canopy cover until trees regenerate. For plots that did not meet FIA's definition of "forest" (i.e., there is no evidence that the site previously supported at least 10% tree cover), we assumed overstory LAI to be zero. If understory vegetation data were collected on a nonforest plot, we used the field-collected vegetation measurements to estimate understory vegetation cover for that plot. If understory vegetation data were not collected (i.e., for efficiency reasons of focusing on forest plots or if plots were covered with rare summertime snows), we treated the plot's understory vegetation data as nonresponse (i.e., no data rather than zero) and thus did not include them in the training dataset for the understory LAI model.

2.3. Estimation of Plot-Level Leaf Area Index (LAI)

We estimated LAI at each FIA plot using four plot-level estimation methods (Table 1). Among these four methods, the first method is most commonly used in the previous literature and consists of linear scaling of tree canopy cover fraction, C, which ranges from 0 to 1, relative to LAI ranging from 0 to a local maximum that must be obtained from either ground observations or previous studies [18]. We used a local maximum LAI of 5.3, which was determined for coniferous forests of western Montana [23]. Coniferous forests in western Montana typically contain small patches of *Populus tremuloides* as a nonconiferous species that is frequently present as a minor and often temporary (i.e., seral) component of coniferous forest stands [24]. The LAI measurements in this earlier LAI study [23] assumed that leaf area angles can be assumed to be random based on validation of this assumption for broadleaf species (i.e., *Quercus*) [25]. This earlier work is thus representative of the

coniferous forests included in the present study that encompass mostly coniferous species as well as nonconiferous individual species such as *Populus.*

The second method was developed to relate tree canopy cover to snow cover [20], using Equation (1), which is the inverted form of the relationship presented as Equation (4) in [19]:

$$\text{LAI} = e^{(C - 0.55)/0.29} \tag{1}$$

These earlier studies modeled change in snow cover as a function of change in LAI [19,20]. Because direct LAI measurements were not available in those studies but tree canopy cover measurements were available, the studies' authors used the equation above (or an inverted form) to calculate LAI as a function of canopy cover and thus obtain the required input for their snow cover model [19,20].

For the third method, we estimated LAI as a function of tree diameter and species-specific coefficients, based on destructive sampling of a small number of trees of four tree species [21]: Engelmann spruce (*Picea engelmannii* Parry), subalpine fir (*Abies lasiocarpa* (Hook.) Nutt.), lodgepole pine (*Pinus contorta* var. *latifolia* Engelm.), and quaking aspen (*Populus tremuloides* Michx.). Because this suite of four species does not encompass all species present in our study area, and allometric models of LAI have not been developed for the other species in our area, we adopted the approach of applying an allometric model developed for one species to a species of similar architecture or growth habit [26]. To determine which species should utilize each of the four species-specific allometric equations for LAI [21], we relied on FIA's existing lookup tables for applying a single allometric equation for biomass (including foliage) to groups of species [15].

Table 1. Empirical methods used for estimating plot-level LAI at Forest Inventory and Analysis (FIA) plots.

Method	Description	FIA Inputs	Outputs	Source
LAI1	Linear scaling of tree canopy cover with local max LAI	Tree canopy cover	Overstory LAI	[18] for method; [23] for max LAI
LAI2	Empirical (natural exponential) function of tree canopy cover	Tree canopy cover	Overstory LAI	[19]
LAI3	Species-specific allometric equations (quadratic function of tree diameter)	Tree species and diameter	Overstory LAI	[21]
LAI4	Gap-fraction model based on Beer's law, with clumping indices specific to vegetative cover type	Tree and understory vegetation cover; forest type; understory vegetation type	Overstory and understory LAI	[27] for equation and clumping indices

The fourth method for estimating LAI has not, to the best of our knowledge, been previously used to estimate LAI directly. This method (LAI4) used percent cover of understory and overstory (tree) vegetation as inputs to an inverted gap-fraction model based on Beer's law (Equation (1) in [27]):

$$P(\theta) = e^{-G(\theta) * \text{LAI} * \Omega / \cos(\theta)} \tag{2}$$

In Equation (2), $P(\theta)$ is the gap fraction at zenith angle of view θ representing the fraction of canopy gaps through which light would penetrate to the ground if illuminated from that angle of view; $G(\theta)$ is the extinction coefficient of light, which has a value of 0.5 for random leaf and branch arrangements; and Ω is a dispersion parameter, or clumping index, that is specific to vegetation type and was conceptually developed by [28]. When leaf arrangement is truly random, then Ω is equal to 1.0, but for most vegetation clumping of leaves and branches means that Ω is less than 1. Canopy fraction measured on FIA plots is the fraction of area that is canopy when viewed from above. Thus taking θ to be zero,

canopy fraction becomes 1.0 minus gap fraction, cos(θ) becomes 1, and the equation above can be inverted to solve for LAI, resulting in:

$$\text{LAI} = \frac{\ln(1 - C)}{-0.5 * \Omega} \tag{3}$$

Values for the clumping index, Ω, were based on mean clumping indices developed by [27] (see Table 3 in that paper). We associated each FIA vegetation type with a clumping index category from [27] to determine its clumping index value (Table 2). Understory vegetation was assigned the clumping index that [27] assigned to "sparse herbaceous or sparse shrub cover" areas (i.e., 0.75). Cover was summed for all understory growth habits (graminoids, herbaceous vegetation, and shrubs) prior to applying the clumping index, which yielded LAI for the understory in its entirety.

Table 2. Look-up table used to assign clumping index values to each vegetation type for method LAI4. Clumping index categories that do not occur in the study area are omitted.

FIA Vegetation Type [1]	Clumping Index Category [2]	Clumping Index [2]
Hardwood deciduous forest types (codes 501–988, with canopy cover ≥65%)	2: Tree cover, broadleaf, deciduous, closed	0.69
Hardwood deciduous forest types (codes 501–988, with canopy cover <65%)	3: Tree cover, broadleaf, deciduous, open	0.70
Softwood evergreen forest types (codes 101–319 & 341–391)	4: Tree cover, needleleaf, evergreen	0.62
Softwood deciduous forest type (code 321)	5: Tree cover, needleleaf, deciduous	0.68
Oak/pine forest types (forest type codes 401–409)	6: Tree cover, mixed leaf type	0.69
Nonstocked forest type (code 999) and nonresponse (inaccessible) plots with possible tree cover	9: Mosaic: Tree cover/other natural vegetation	0.72
Understory vegetation and nonforest plots where vegetation data were collected	14: Sparse herbaceous or sparse shrub cover	0.75

[1] FIA forest types and nonforest/nonresponse status are described by the variables FORTYPCD and COND_STATUS_CD in [15]. [2] Clumping index categories and values are from Table 3 of [27].

Evaluation of the four plot-scale LAI estimation methods (Table 1) was accomplished by means of comparison with Landsat-based total LAI at 30 m resolution [29]. For this comparison, we used plot-scale estimates for plots measured in 2019 (*n* = 87) and Landsat-based LAI for the 2019 growing season at those same plot locations. We compared the frequency distributions of the four plot-scale estimation methods with the frequency distribution of Landsat total LAI, produced scatterplots of plot-based vs. Landsat-based LAI for each plot-scale estimation method, and compared the Spearman correlation coefficient of each method's plot-scale LAI with the Landsat LAI. The rationale for using Spearman's r for plot-scale comparisons was that, first, the imposition of LAI maxima affects the correlation with Landsat LAI, and our objective was to evaluate these methods' ability to estimate variation in LAI, rather than evaluating our specific choice of LAI maxima. Second, the relationship between LAI2 and Landsat was not linear but was monotonic.

For all of these comparisons, methods LAI1, LAI2, and LAI3 represent only overstory LAI, while method LAI4 represents overstory LAI, understory LAI, and total LAI as the sum of overstory and understory LAI estimates (Table 1). Landsat-based LAI represents total (i.e., overstory plus understory) LAI but is known to saturate at a value of about 4 [29]. Thus, we expected that our overstory LAI estimates would not precisely equal Landsat-based LAI but that there should be some correlation given that most of the study area is forested with overstory contributing the majority of total LAI. Moreover, ground-based estimates that include both understory and overstory vegetation may exceed Landsat-based estimates because ground-based estimates are not subject to the same saturation constraints as reflectance-based LAI. Therefore, we expected that LAI4 might produce total LAI estimates that exceed Landsat-based total LAI.

2.4. Interpolation of Gridded LAI Datasets

Our second objective was to combine plot-scale LAI with spectral reflectance data and other spatially explicit variables in a machine learning algorithm to produce spatially and temporally explicit maps of overstory and understory LAI on an annual basis. We focused

on annual maximum LAI because variations in forest LAI over time have been mainly attributed to interannual changes in tree cover due to management or disturbance, where such interannual dynamics are greater in magnitude than growing-season variability in LAI [30]. Plot-scale LAI estimates were derived using four different methods, described above. The machine learning algorithm we used was random forests, which is a non-parametric statistical technique that builds an ensemble model from many iterations of individual classification or regression trees and uses bootstrapping, or bagging, to train and improve model performance [31]. We developed five separate random forests regression models: one for each of the four overstory estimation methods (LAI1, 2, 3, and 4) plus one understory estimation method (LAI4).

Like other supervised statistical models or machine learning algorithms, a random forests model requires specification of a response variable and a set of predictor variables. For each model, plot-scale LAI served as the response variable. The training dataset included 1049 plot locations that were measured between 2003 and 2019.

All models included the same set of predictor variables (Table 3).

Table 3. Predictor variables used in all random forests models for predicting LAI.

Description	Source	Citation
Composite maximum annual greenness	Google Earth Engine (GEE) image collections	GEE image collections: LANDSAT_LE07_C01_T1_ANNUAL_ GREENEST_TOA (years 1999–2019) and LANDSAT_LT05_C01_T1_ANNUAL_ GREENEST_TOA (years 1984–1998)
Elevation	Digital elevation model (DEM) from The National Map	[32]
Slope	DEM processed in R package 'raster'	[33]
Aspect	DEM processed in R package 'raster'	[33]
Topographic wetness index	DEM processed in TauDEM	[34,35]
Exposure index	DEM processed in ArcGIS	[36]
Tree canopy cover	National Land Cover Dataset	[37]
Mean annual precipitation	PRISM Dataset	[38]
Min and max annual temperature	PRISM Dataset	[38]
Soils hydrologic group code	STATSGO2 Dataset	[39]

The rationale for including these variables as predictors in our random forests models was because they either provide direct evidence of the density of vegetation at a particular point in time (e.g., NDVI), or are likely to affect the potential vegetation type or density in the long term (e.g., elevation, precipitation, etc.). Predictor variables included composite maximum annual greenness quantified using normalized difference vegetation index (NDVI) from Landsat 5 or Landsat 7 (Table 3), acquired from Google Earth Engine [40]. The maximum annual greenness for each pixel was calculated as the highest recorded NDVI from the images available for that year. The term "composite" simply means that the dataset is composed of pixels representing multiple dates, where the date represented by each pixel represents that pixel location's maximum greenness for that particular year [32]. For any particular plot, only the composite maximum annual greenness for the year that plot was measured was included in the training dataset. We did not attempt to schedule forest inventory plot measurements to coincide with the date of maximum annual greenness, partly for practical reasons and also because the date of maximum annual greenness is not known until after it has occurred. Other predictors included elevation and other topographic variables derived from elevation, including slope and aspect [33], topographic wetness index [34] calculated using TauDEM [35], and a topographic exposure index [36]; tree canopy cover from the National Land Cover Dataset [37]; precipitation and temperature [38]; and soils hydrologic group code [39] as a categorical input variable to random forests. The value of each predictor variable was extracted at the spatial location of each plot to create a plot-based training dataset.

After the models were trained on plot-level observations, each model was then used with spatially gridded predictor variables as inputs (Table 3) to produce annual gridded LAI datasets. Note that the plot-level LAI estimates were used only for model calibration and were not needed for making predictions of LAI based on predictor variables. Thus, we were able to produce annual maximum LAI estimates for any year for which maximum annual greenness exists, including years in which no plots were measured (i.e., as early as 1984). Beyond the plot based out-of-bag random forest validation, we also assessed agreement between our gridded LAI outputs and Landsat-based LAI [29] for a single year. We selected the year 2003 for this assessment because the Landsat-based LAI had minimal missing pixel values for that year. Assessment metrics included Pearson's correlation coefficient, mean absolute error, mean bias error, and root mean squared error, which is similar to mean absolute error in that it ignores the direction of error (i.e., bias) but penalizes for larger individual errors. We used the Pearson correlation coefficient here because it provides a more informative assessment of the ability to produce LAI maps for their intended end-use, which is assessing change over time or using the maps as inputs to hydrologic models.

2.5. Time Series of LAI for Two Large Watersheds

After producing gridded LAI datasets for multiple years (Figure 2, bottom row), we identified the two plot-level LAI estimation methods that showed the best agreement with Landsat-based LAI and used those methods to produce a time series of annual LAI gridded datasets from 1984 to 2019. We examined these time series within two large watersheds, the South Fork Flathead River and Middle Fork Flathead River, in our modeling domain (Figure 1). We tested the time series of overstory LAI for monotonic temporal trend using the Mann–Kendall trend test [41] via R package 'Kendall' [42].

We also examined the time series of annual evaporative (ET) ratios, estimated as the proportion of precipitation that did not result in runoff and calculated as 1 minus the ratio of annual streamflow to annual precipitation for each year. ET ratio for the Middle Fork Flathead River watershed was obtained from the CAMELS dataset for US Geological Survey (USGS) gage 12358500 [43]. For the South Fork Flathead River, daily streamflow data were obtained for USGS gage 12359800 from R package 'dataRetrieval' [44]. Both gage locations are shown in Figure 1. South Fork precipitation data were compiled from Daymet gridded data, which is the same source data used to compile the CAMELS dataset [42], using R package 'daymetr' [45]. We assessed the strength of correlations between each LAI estimation method and ET ratio based on these time series. Although we tested for a lagged correlation between LAI and ET ratio, there was no significant lag detected and we therefore did not implement a lag in the correlation analysis.

3. Results

3.1. Estimation of Plot-Level LAI

Correlation coefficients for estimates of plot-scale overstory LAI relative to Landsat-based total LAI were between 0.703 and 0.710, while for the combination of overstory and understory based on method LAI4, the correlation was 0.578 (Figure 4a). Plot-based methods for estimating overstory LAI were strongly correlated with each other (pairwise r > 0.85 for all pairs), which suggests that the choice of one method over another may not be tremendously impactful for estimating overstory LAI alone. Scatterplots between each plot-scale estimation method vs. Landsat reveal that methods LAI2, LAI3, and LAI4 (overstory) underestimate LAI (Figure 4a) relative to Landsat. Violin plots also demonstrate this pattern (Figure 4b), which is unsurprising given that Landsat-based LAI detects all vegetation without distinguishing between overstory and understory strata, while our ground-based overstory LAI methods did not include the understory vegetation component. This result is somewhat expected because most forests in the region are relatively open and do not have fully closed canopies. In contrast, the sum of overstory and understory LAI estimates produced by method LAI4 overestimate LAI relative to Landsat-based estimates (Figure 4a,b), which may reflect the fact that ground-based estimates of multilayered LAI

are not subject to the saturation that occurs at a value of about 4 for Landsat-based LAI [29]. The highest estimates produced by method LAI4_total were 16.5, which is higher than estimates produced by any other method but also within the range observed in a study in nearby Glacier National Park [46], which found values as high as 20.8. Thus, the ability of method LAI4 to estimate total LAI represents a contribution in overcoming a known limitation of total LAI as estimated from spectral remote sensing. Among all methods, method LAI4 showed the widest dispersion of LAI values (Figure 4b).

(**a**)

(**b**)
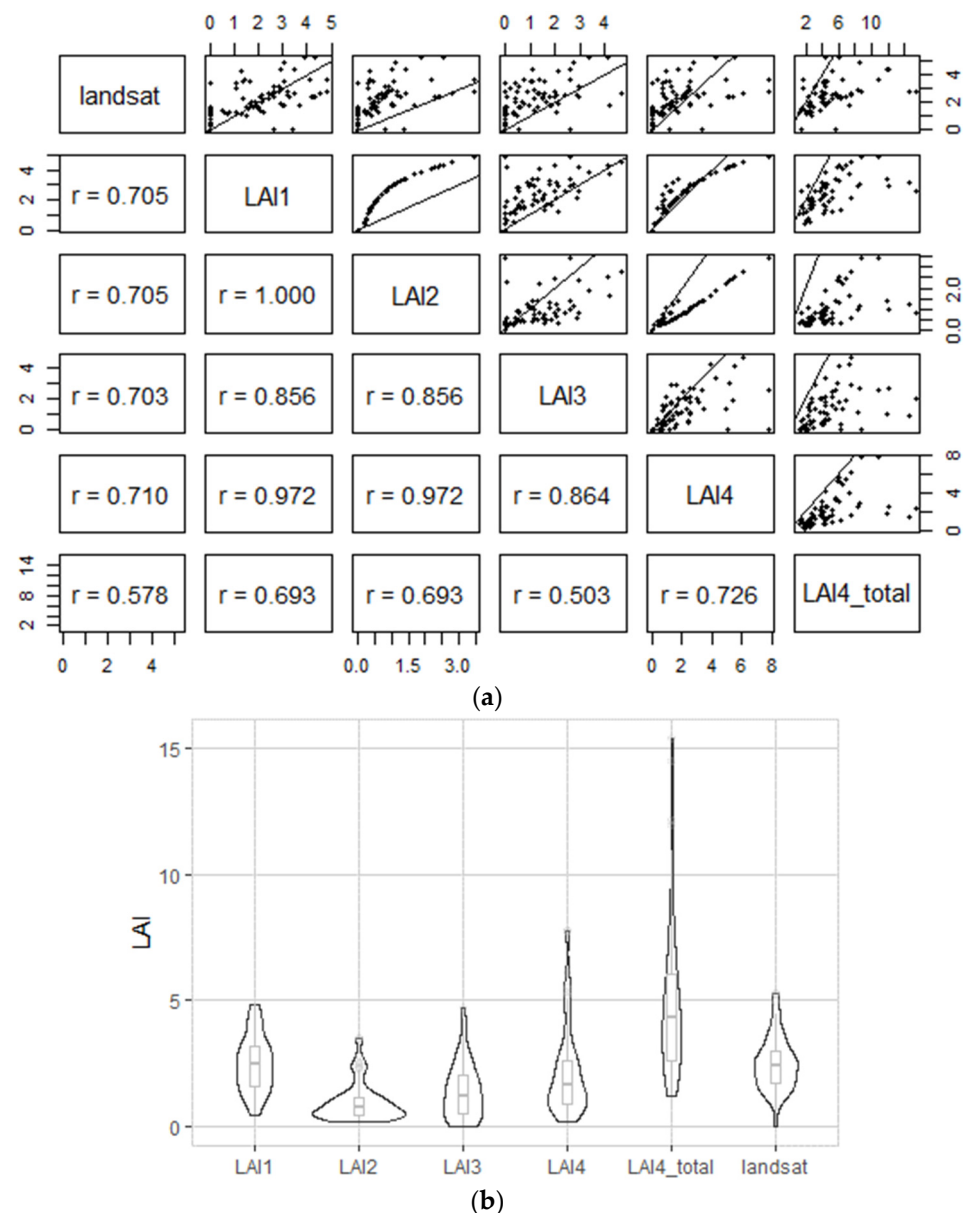
Figure 4. Comparisons of plot-level leaf area index (LAI) at Forest Inventory and Analysis (FIA) plots: (**a**) scatterplots and correlations; R = Spearman's rank correlation coefficient; and (**b**) violin shapes show frequency distribution, and boxplots show median (horizontal bar) and interquartile range (box). Four methods estimate overstory LAI (LAI1, LAI2, LAI3 and LAI4), one method estimates both overstory and understory LAI (LAI4_total), and Landsat-based estimates represent total LAI from the method of [29] at plot locations, all using data collected in 2019.

3.2. Interpolation of Gridded LAI Datasets

The assessment above described how well plot-level LAI estimates correspond to Landsat-based LAI, and, here, we describe machine learning models that used those plot-

level LAI estimates as calibration data for estimating LAI from gridded biophysical and remote sensing predictor variables that we extracted at FIA plot locations (Figure 2). For all LAI estimation methods, the machine learning algorithms had value of model $r^2 > 0.6$ (Table 4). In contrast, understory LAI (method LAI4_under) had a very weak model ($r^2 = 0.03$), which might reflect that understory vegetation may be undetectable by remote sensing-based predictor variables (i.e., composite maximum annual greenness and the NLCD tree canopy cover layer) due to tree canopies obscuring this vegetation on forest plots. Indeed, we found that climatic variables (precipitation and temperature) were the most important predictors of understory LAI (method LAI4_under), while both climatic and remote sensing-based predictor variables were important predictors for all models of overstory LAI. Among all plot-scale LAI estimation methods, the random forests model based on method LAI1 had the highest proportion of LAI variability explained by the model while method LAI2 had the lowest model error (mean squared residual, or MSR, in Table 4). Method LAI4's overstory model had the lowest model r^2 among overstory estimation methods, although differences among models were not large. Note that these metrics of model performance reflect the ability of the predictor variables to explain (via random forests regression) the plot-to-plot variability in each plot-level LAI estimation method, and, thus, they do not reflect the accuracy of any particular method.

Table 4. Performance metrics for random forests models of LAI based on four overstory estimation methods (LAI1, LAI2, LAI3, and LAI4) and one understory method (LAI4_under). Mean of squared residuals (MSR) and model r^2 are based on out-of-bag samples from 500 trees.

Method	MSR	Model r^2
LAI1	0.46	0.77
LAI2	0.17	0.66
LAI3	0.40	0.69
LAI4	0.92	0.64
LAI4_under	5.79	0.03

All four LAI estimation methods produced gridded datasets that are strongly correlated with Landsat-based gridded LAI, which provides evidence of whether our gridded LAI values are realistic (Table 5). Method LAI4's total LAI (i.e., overstory plus understory) was most strongly correlated with Landsat (r = 0.80). Mean absolute error, mean bias error, and root mean squared error were all lowest for method LAI1, followed closely by method LAI4, and were all highest for method LAI2. These pixel-based comparisons show a bias similar to that exhibited by plot-based comparisons: all overstory methods have a positive mean bias error, and are thus lower than Landsat-based total LAI, while total LAI produced by method LAI4 is higher than Landsat-based LAI and has a negative mean bias error (Table 5). As discussed above, it is realistic that overstory LAI alone would be lower than Landsat-based total LAI and that ground-based total LAI (method LAI4) would be higher than Landsat-based LAI.

The wider dispersion of total LAI values produced by method LAI4, particularly when understory vegetation is included, more closely resemble the spatial variability and dispersion of values demonstrated by Landsat-based total LAI (Figure 5). This pattern is evident for two different years, 2003 and 2019, which were used to assess not only single-year LAI estimates (above) but also change in LAI over time. Based on visual comparisons of change in Landsat-based estimates versus change in our gridded LAI datasets between 2003 and 2019, methods for estimating overstory LAI—particularly LAI2 and LAI3, and to a lesser extent LAI1 and LAI4 overstory—are missing a lot of change that is captured by Landsat-based LAI and total LAI as estimated by method LAI4. Notable differences still exist in the amount of change in LAI observed from Landsat compared to total LAI as estimated by method LAI4; these differences may be due to inaccuracies in either Landsat-based LAI or our plot-based method of estimating LAI but resolving the ultimate cause

of the discrepancy would require ground validation data and was beyond the scope of this study.

Table 5. Pixel-to-pixel comparisons of multiple gridded LAI datasets relative to Landsat-derived LAI as estimated for 2003. MAE = mean absolute error; MBE = mean bias error; RMSE = root mean squared error. LAI1, LAI2, LAI3, and LAI4, represent overstory LAI; LAI4_total represents the sum of overstory and understory LAI.

2003	Pearson's *r*	MAE	MBE	RMSE
LAI1	0.75	0.97	0.83	1.28
LAI2	0.72	1.78	1.77	2.12
LAI3	0.71	1.48	1.44	1.80
LAI4	0.72	1.15	1.04	1.46
LAI4_total	0.80	1.11	−0.99	1.42
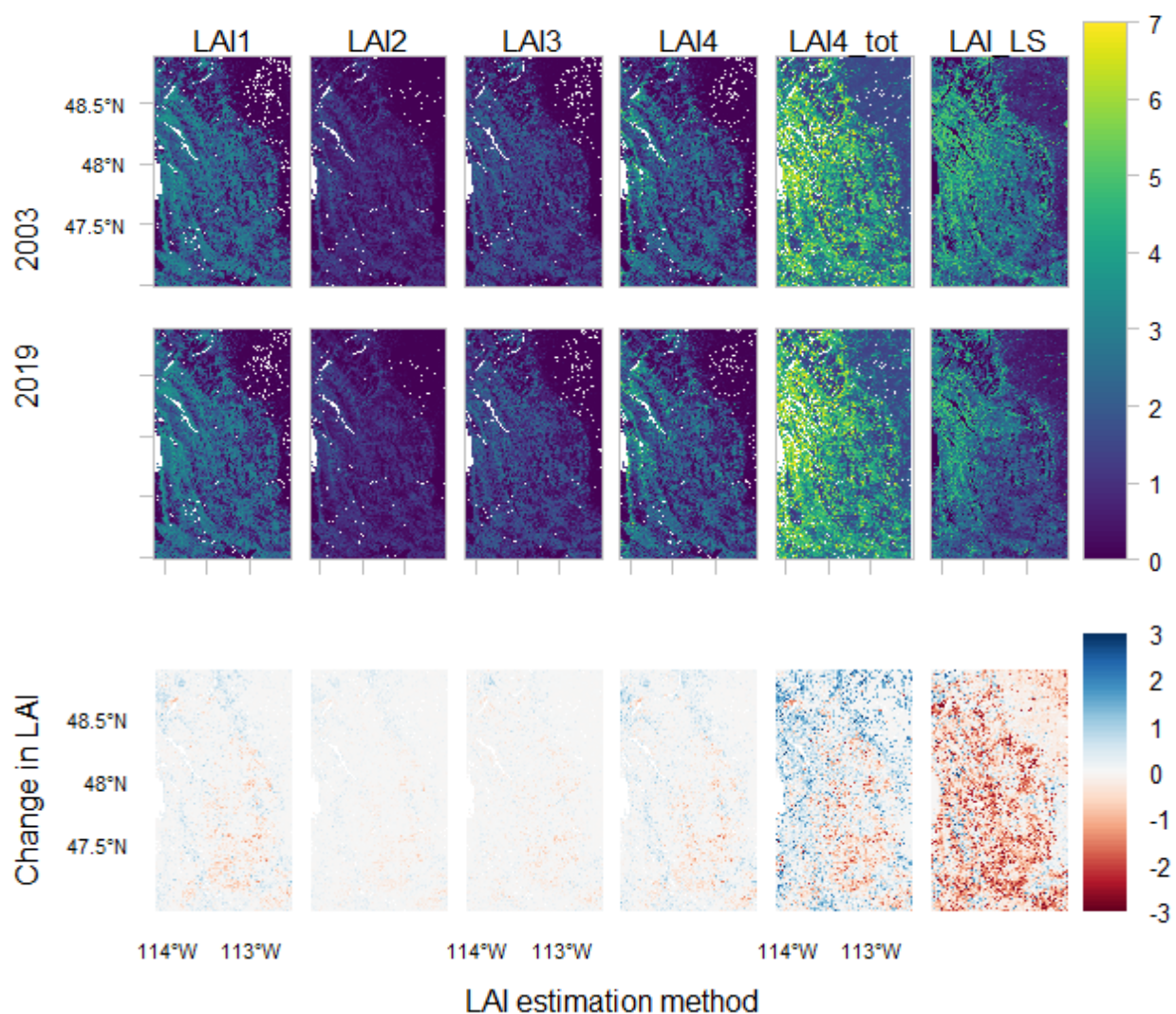
Figure 5. Maps of LAI based on random forests interpolations of empirical plot-based methods (LAI1, LAI2, LAI3, and LAI4 for overstory LAI; LAI4_tot for overstory + understory LAI), and Landsat total LAI (LAI_LS) for 2003 and 2019 (**top** and **middle** rows, respectively) and for the difference between 2003 and 2019 (**bottom** row) for the entire modeling domain shown in Figure 1. Negative change represents decreases in LAI.

3.3. Time Series of LAI for Two Large Watersheds

Time series of LAI for 1984–2019 yielded a subtle but statistically significant ($p < 0.05$) decreasing trend in overstory, understory, and total LAI in the South Fork Flathead River watershed (Figure 6a). Methods LAI1 and LAI4 both detected a decrease in overstory LAI, although this decrease was small compared to the decrease in understory and total LAI detected by method LAI4. In contrast to the South Fork, the Middle Fork watershed did not exhibit any significant trend in overstory LAI, although method LAI4 did detect a significant decrease in total LAI from 1984 to 2019 (Figure 6b). There were no significant trends in ET ratio in either watershed.
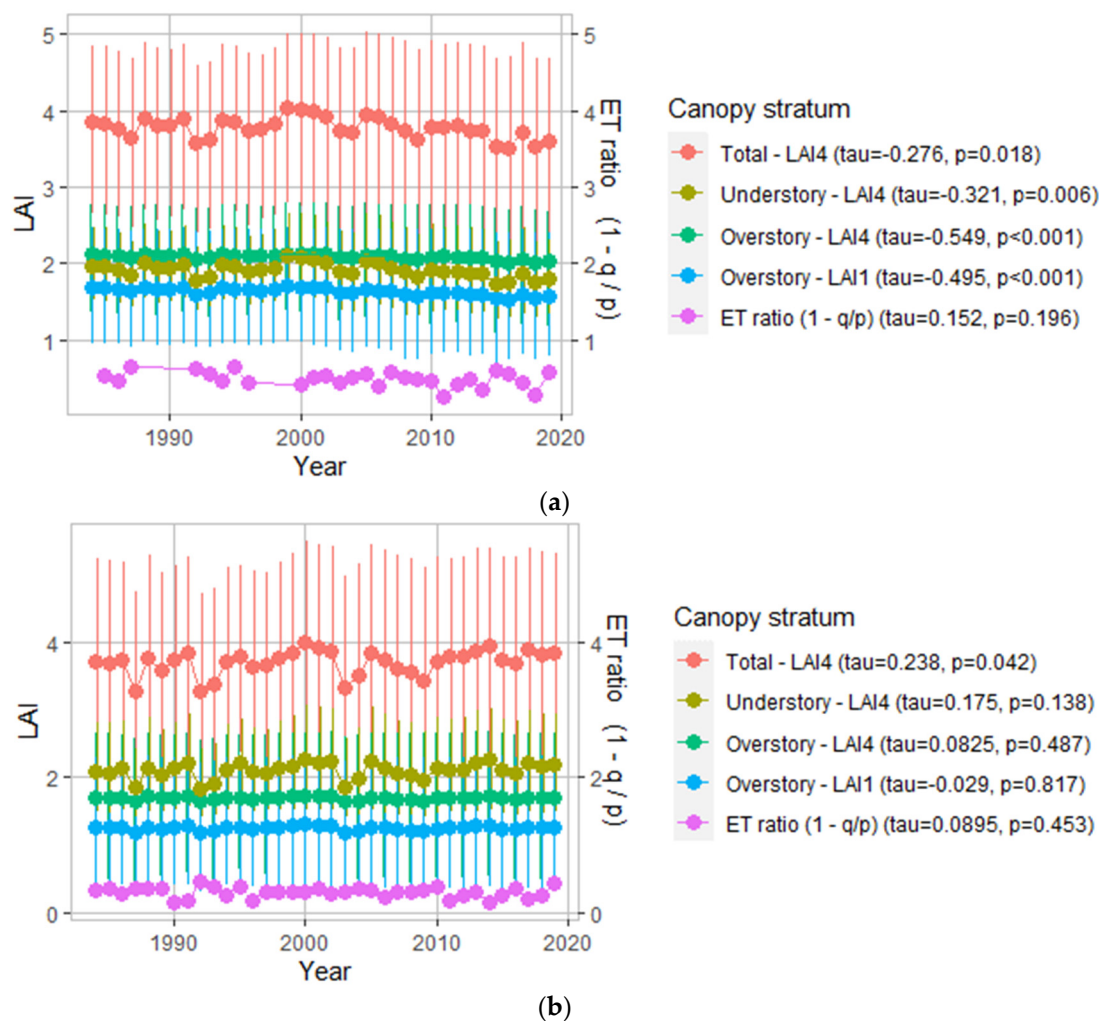
Figure 6. Annual time series of overstory, understory, and total leaf area index (LAI) in: (**a**) South Fork Flathead Watershed and (**b**) Middle Fork Flathead Watershed, for 1984–2019, as produced by random forests models based on methods LAI1 and LAI4, and ET ratio (1 – ratio of mean annual streamflow to mean annual precipitation). LAI points represent watershed-scale medians and bars represent the 1st (lower) and 3rd (upper) quartiles. Values of tau and associated *p*-values in the legend represent results of the Mann–Kendall trend test. Note missing observations for some years in the South Fork Flathead River due to lack of streamflow data.

Annual median LAI for overstory, understory, and total LAI based on methods LAI1 and LAI4 were strongly correlated with one another, with R > 0.8 for all pairwise comparisons (Figure 7). Thus, methods LAI1 and LAI4 produce highly correlated watershed-scale LAI estimates (r = 0.977 for the South Fork, and r = 0.962 for the Middle Fork, for LAI1 vs. LAI4 overstory LAI). Correlations with ET ratio show some minor differences among

LAI estimation methods. In the South Fork watershed, ET ratio is very weakly correlated with overstory LAI as estimated by methods LAI1 and LAI4, and slightly more strongly correlated with understory and total LAI produced by method LAI4 (Figure 7a). In the Middle Fork watershed, correlations between annual median LAI and ET ratio were stronger, and as in the South Fork, ET was more strongly correlated with total LAI as estimates by method LAI4 than with other estimates.
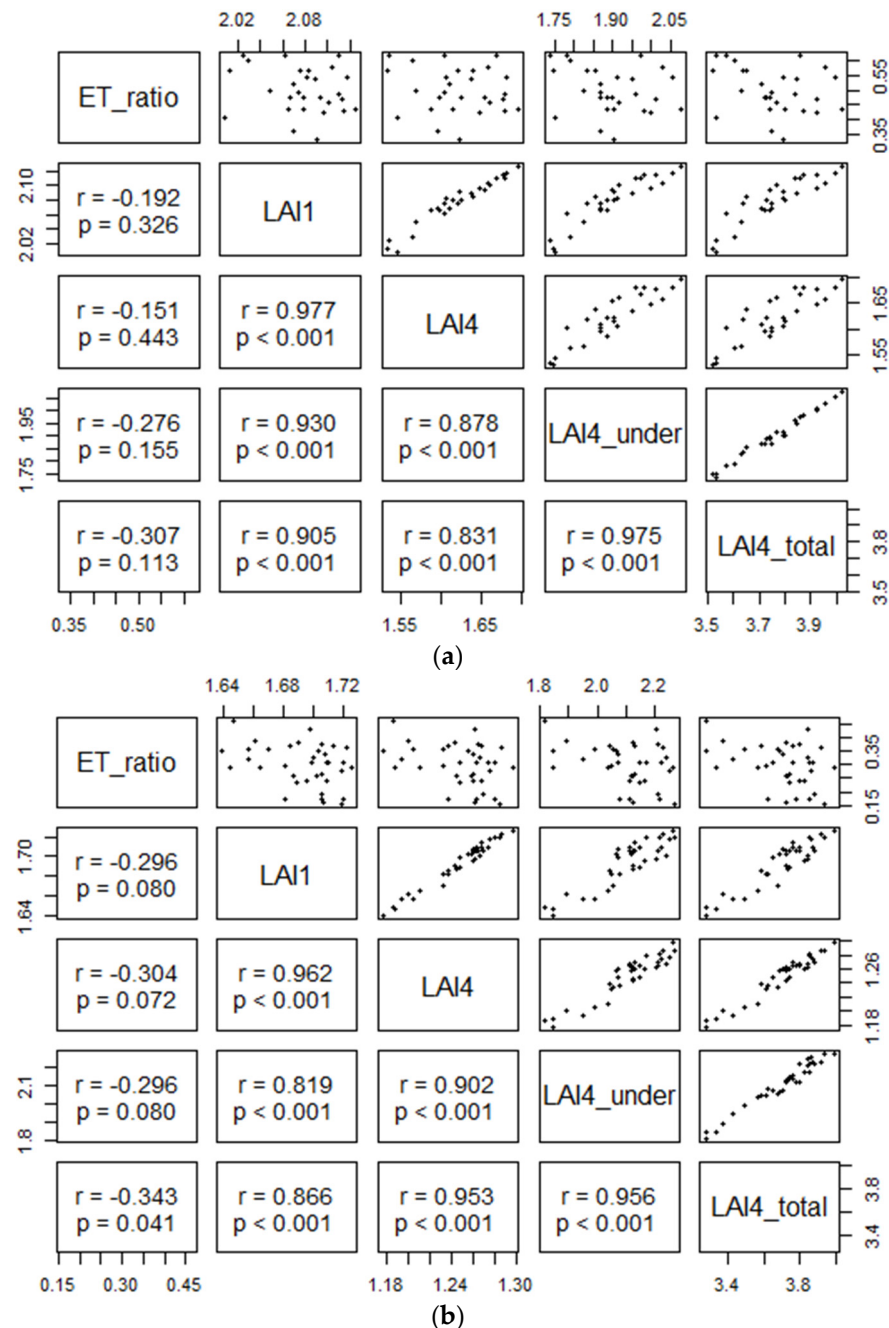

Figure 7. Comparisons of annual ET ratio and watershed-scale median leaf area index for two watersheds: (**a**) South Fork Flathead River and (**b**) Middle Fork Flathead River, as estimated using methods LAI1 and LAI4 (overstory), LAI4under (understory) and LAI4total (total LAI) for water years 1984–2019. ET ratio is defined as 1 minus the ratio of mean annual streamflow to mean annual precipitation. R represents the Spearman's rank correlation coefficient.

4. Discussion

4.1. Summary of Our Comparison of Multiple Methods for Estimating LAI

This study demonstrated development of gridded LAI datasets that are not subject to the same constraints as remote sensing-based datasets, namely the inability to separate LAI into overstory vs. understory strata, as well as the saturation that occurs at specific LAI values and above which remote sensing-based methods cannot distinguish variability in LAI densities. Although new methods for estimating Landsat-based LAI have become computationally efficient and publicly available on Google Earth Engine [29], this algorithm was developed for Landsat 5 and later versions of Landsat and is thus not applicable to the full Landsat record. In contrast, our machine learning models used only a single Landsat-based predictor variable that is available beginning in 1984, and it is thus possible to produce gridded datasets of maximum annual LAI for any year from 1984 to the most recent growing season.

Among the four alternative methods we tested for estimating plot-scale LAI and then interpolating gridded LAI datasets, methods LAI1 and LAI4 performed the best overall in comparison to Landsat-based LAI. Each of these methods has specific strengths and weaknesses. Method LAI1 has the advantage of requiring only a single ground measurement—plot-level tree canopy cover—and is thus more parsimonious. Because method LAI4 uses a clumping index that is specific to forest type, it requires a more complex crosswalk of FIA's forest types to the cover types specified for clumping indices in [27]. However, scripting is provided to accomplish this task, and this method could more reliably be applied to FIA data anywhere in the USA, unlike method LAI1 which may require tuning of local maximum LAI based on direct field measurements of LAI. Method LAI4 has the advantage of producing not only overstory LAI but also understory and thus total LAI. Total and understory LAI, as produced by method LAI4, demonstrated the greatest sensitivity to change in LAI over time, as compared to Landsat-based LAI, and LAI4's total LAI was more strongly correlated than LAI1 to evaporative fraction.

One outcome of this study is the comparative evaluation of several previously used and published methods for estimating LAI when direct measurements or remote sensing-based data are not available. Specifically, methods LAI1, LAI2, and LAI3 have been used by prior studies to estimate overstory on the basis of tree canopy cover [18–20] or the combination of species identity and DBH [21]. We found that the simplest of these methods, LAI1, produced the best agreement with Landsat-based LAI. However, although method LAI2 did not agree as well with Landsat data, it may be preferable for snow modeling applications, which was the original purpose of that LAI estimation method [19,20].

The choice of the most appropriate plot-scale LAI estimation method depends largely on the intended application or question. For consideration of overstory LAI only, linear scaling of tree canopy cover with LAI, i.e., method LAI1, may be sufficient and parsimonious. For studies investigating processes within overstory vs. understory strata, e.g., carbon allocation or partitioning of precipitation into various hydrologic pathways, LAI4 would be most appropriate because it is capable of estimating not only overstory but also understory and thus total LAI, without the saturation constraint imposed by remote sensing-based total LAI datasets. Although we used composite maximum annual greenness as a predictor variable in our machine learning models, and greenness is subject to the same saturation effect as Landsat-based LAI, the fact that we used machine learning models with additional biophysical (nonspectral) predictor variables may help to overcome constraints imposed by spectral remote sensing.

4.2. Benefits and Caveats of Using Forest Inventory Plot Data to Estimate LAI

The need to couple understory and overstory vegetation in models has been previously recognized due to the interactions between forest vegetation layers [47,48]. Overstory and understory vegetation have distinct but interacting responses to both disturbance and postdisturbance recovery [49,50], which may have unknown ramifications for fluxes of water, energy, and carbon. The ability to estimate LAI for multiple canopy strata could be

leveraged in models that have distinct representations of separate strata [9,10], and such applications could enhance our understanding of the process-level responses to disturbances that alter forest structure or result in type changes from forest to nonforest vegetation.

A unique benefit of using national forest inventory (NFI) data, such as that collected by FIA which we used to produce plot-based estimates of LAI, is that ground data are acquired from an ongoing data collection program with a probabilistic sample and methods that are consistent over time. Although lidar acquisitions such as ICESat-1, ICESat-2, and GEDI all present promising capabilities to estimate vertically integrated LAI [51], thus far, such space-based missions are temporally constrained compared to ongoing ground data collection by NFIs [6,13,14]. Thus, this approach to estimating annual LAI could be applied in other countries that have ongoing NFIs that include measurements of overstory and understory canopy cover.

Although this study presents novel methods of estimating LAI both as plot-scale estimates and as gridded datasets, it does come with caveats. First, the methods tested here obviously require ground observations and do not provide LAI estimates within a year or season. In contrast, Landsat-based LAI from radiative transfer models can provide greater temporal resolution and thus seasonal (intra-annual) LAI, but again with the caveat that it estimates only total LAI. Second, validating any LAI dataset is challenging, and our study was limited by having only Landsat-based total LAI to use as a point of reference. However, the Landsat-based dataset we used was well-calibrated and validated using widely distributed, intensively studied sites throughout the U.S. [29]. Although we did not present the results within this paper, we did investigate the use of LAI datasets derived from MODIS and ICESat-1 as comparison datasets. We found that pairwise correlations among these datasets were very weak (r < 0.2), possibly due to spatial offsets and scale discrepancies that are difficult to resolve, and thus they are not included here.

4.3. Recommendations for Future Research and Development

We recommend that future studies conduct validation of both plot-based and gridded LAI produced using methods LAI1 and LAI4. Plot-based validation could be accomplished by measuring LAI using light-sensing devices (e.g., LI-COR sensors or analysis of hemispherical photography) on a subsample of FIA plots and then using those LAI values for calibrating and validating various methods that estimate LAI based on other FIA measurements. Calibration might include improved parameterization of clumping indices for specific vegetation or forest types in support of method LAI4. Plot-based LAI measurements could also be used to test the assumption inherent in method LAI1, i.e., that a single maximum LAI value could be implemented across broad regions as the constraining upper limit for scaling canopy cover against overstory LAI. Gridded LAI datasets could be further evaluated by comparing the performance of a physically based hydrologic model using alternative LAI datasets as inputs, specifically, the impact of using strictly remote sensing-based total LAI, compared to overstory and understory gridded LAI, on the ability to estimate hydrologic fluxes (e.g., canopy evapotranspiration, soil evaporation, maximum snow water equivalent, etc.).

Finally, we recommend that broad applicability of our findings will likely require development of tools that allow others to define an area of interest and produce multistrata gridded LAI datasets. To facilitate widespread and innovative use of these methods, such applications would ideally occur within cloud-based computing platforms that consume FIA plot data as well as the predictor variables used in our random forests models. However, in the case of the USA, such development will require cooperative agreements between FIA and cloud-based computing platforms that protect the confidentiality of plot locations as required by federal law [52]. Such advances could allow scientists in forestry and hydrology to use overstory and understory LAI datasets, rather than simply total LAI, for multiple purposes. Incorporation of overstory and understory LAI into physically based ecological and hydrologic models could enhance future understanding of how forest disturbance, recovery, and land-cover change affect both forest and water resources.

5. Conclusions

This study compared four methods for estimating plot-scale leaf area index, and then used those plot-scale estimates in a machine learning algorithm to produce gridded LAI datasets on an annual basis. We evaluated these four alternative methods by comparing both plot-based and gridded LAI estimates against Landsat-based total LAI. We found that the simplest LAI estimation method performed well at estimating overstory LAI but did not address our objective of estimating both overstory and understory LAI. The method based on an inverted gap-fraction model, combined with previously published clumping indices that are specific to vegetation type, performed best at capturing trends over time and also produced separate estimates of understory, overstory, and total LAI. Future research could improve validation and parameterization of plot-based LAI estimates and test the assumption that gridded LAI datasets that partition LAI into multiple canopy strata will lead to enhanced performance in hydrologic models.

The separate overstory and understory LAI datasets produced using method LAI4 for years 1984-2019 are published and available for download from [53].

References

1.　Adams, H.D.; Luce, C.H.; Breshears, D.D.; Allen, C.D.; Weiler, M.; Hale, V.C.; Smith, A.M.S.; Huxman, T.E. Ecohydrological consequences of drought- and infestation- triggered tree die-off: Insights and hypotheses. *Ecohydrology* **2011**, *5*, 145–159. [CrossRef]
2.　Bosch, J.M.; Hewlett, J.D. A Review of Catchment Experiments to Determine the Effect of Vegetation Changes on Water Yield and Evapotranspiration. *J. Hydrol.* **1982**, *55*, 3–23. [CrossRef]
3.　Hibbert, A.R. Forest Treatment Effects on Water Yield. *Int. Symp. For. Hydrol.* **1967**, 527–543.
4.　Molotch, N.P.; Blanken, P.D.; Williams, M.W.; Turnipseed, A.A.; Monson, R.K.; Margulis, S.A. Estimating sublimation of intercepted and sub-canopy snow using eddy covariance systems. *Hydrol. Process.* **2007**, *21*, 1567–1575. [CrossRef]
5.　Stottlemyer, R.; Troendle, C. Effect of canopy removal on snowpack quantity and quality, Fraser experimental forest, Colorado. *J. Hydrol.* **2001**, *245*, 165–176. [CrossRef]
6.　Andréassian, V. Waters and forests: From historical controversy to scientific debate. *J. Hydrol.* **2004**, *291*, 1–27. [CrossRef]
7.　Tague, C.L.; Moritz, M.; Hanan, E. The changing water cycle: The eco-hydrologic impacts of forest density reduction in Mediterranean (seasonally dry) regions. *WIREs Water* **2019**, *6*, e1350. [CrossRef]
8.　Brown, A.E.; Zhang, L.; McMahon, T.A.; Western, A.W.; Vertessy, R.A. A review of paired catchment studies for determining changes in water yield resulting from alterations in vegetation. *J. Hydrol.* **2005**, *310*, 28–61. [CrossRef]
9.　Tague, C.L.; Band, L.E. RHESSys: Regional Hydro-Ecologic Simulation System—An Object-Oriented Approach to Spatially Distributed Modeling of Carbon, Water, and Nutrient Cycling. *Earth Interact.* **2004**, *8*, 1–42. [CrossRef]
10.　Wigmosta, M.S.; Vail, L.W.; Lettenmaier, D.P. A distributed hydrology-vegetation model for complex terrain. *Water Resour. Res.* **1994**, *30*, 1665–1679. [CrossRef]
11.　Goeking, S.A.; Tarboton, D.G. Forests and Water Yield: A Synthesis of Disturbance Effects on Streamflow and Snowpack in Western Coniferous Forests. *J. For.* **2020**, *118*, 172–192. [CrossRef]
12.　Rouhani, S.; Schaaf, C.L.; Huntington, T.G.; Choate, J. Simulation of Dissolved Organic Carbon Flux in the Penobscot Watershed, Maine. *Ecohydrol. Hydrobiol.* **2021**, *21*, 256–270. [CrossRef]
13.　Klesse, S.; Derose, R.J.; Guiterman, C.; Lynch, A.M.; O'Connor, C.D.; Shaw, J.D.; Evans, M.E.K. Sampling bias overestimates climate change impacts on forest growth in the southwestern United States. *Nat. Commun.* **2018**, *9*, 5336. [CrossRef]
14.　McRoberts, R.E.; Bechtold, W.A.; Patterson, P.L.; Scott, C.T.; Reams, G.A. The Enhanced Forest Inventory and Analysis Program of the USDA Forest Service: Historical Perspective and Announcement of Statistical Documentation. *J. For.* **2005**, *103*, 304–308.
15.　Burrill, E.; Wilson, A.; Turner, J.; Pugh, S.; Menlove, J.; Christiansen, G.; Conkling, B.; David, W. *The Forest Inventory and Analysis Database: Database Description and User Guide Version 8.0 for Phase 2*; US Department of Agriculture Forest Service: Washington, DC, USA, 2018.
16.　Härkönen, S.; Lehtonen, A.; Manninen, T.; Tuominen, S.; Peltoniemi, M. Estimating Forest Leaf Area Index Using Satellite Images: Comparison of k -NN Based Landsat-NFI LAI with MODIS- RSR Based LAI Product for Finland. *Boreal Environ. Res.* **2015**, *20*, 181–195.
17.　USDA Interior West Forest Inventory & Analysis P2 Field Procedures, v. 8.0. Available online: https://www.fs.fed.us/rm/ogden/data-collection/pdf/V80_IW_FIA_P2manualMarch31_2019.pdf (accessed on 6 March 2021).
18.　Broxton, P.D.; Harpold, A.; Biederman, J.; A Troch, P.; Molotch, N.P.; Brooks, P. Quantifying the effects of vegetation structure on snow accumulation and ablation in mixed-conifer forests. *Ecohydrology* **2014**, *8*, 1073–1094. [CrossRef]
19.　Varhola, A.; Coops, N.C.; Weiler, M.; Moore, R.D. Forest canopy effects on snow accumulation and ablation: An integrative review of empirical results. *J. Hydrol.* **2010**, *392*, 219–233. [CrossRef]
20.　Pomeroy, J.W.; Gray, D.M.; Hedstrom, N.R.; Janowicz, J.R. Physically based estimation of seasonal snow accumulation in the boreal forest. In Proceedings of the 59th Eastern Snow Conference, Stowe, VT, USA, 5–7 June 2002; pp. 93–108. [CrossRef]
21.　Kaufmann, M.R.; Edminster, C.B.; Troendle, C.A. *Leaf Area Determinations for Subalpine Tree Species in the Central Rocky Mountains*; US Department of Agriculture, Forest Service, Research Paper; US Department of Agriculture: Washington, DC, USA, 1982.
22.　U.S. Geological Survey The StreamStats Program. Available online: http://streamstats.usgs.gov (accessed on 10 October 2021).
23.　Pierce, L.L.; Running, S.W. Rapid Estimation of Coniferous Forest Leaf Area Index Using a Portable Integrating Radiometer. *Ecology* **1988**, *69*, 1762–1767. [CrossRef]
24.　Pfister, R.D.; Kovalchik, B.L.; Arno, S.F.; Presby, R.C. *Forest Habitat Types of Montana*; Gen. Tech. Rep. 34; U.S. Department of Agriculture, Intermountain Forest & Range Experiment Station: Ogden, UT, USA, 1977.
25.　Caldwell, M.; Meister, H.-P.; Tenhunen, J.; Lange, O. Canopy structure, light microclimate and leaf gas exchange of Quercus coccifera L. in a Portuguese macchia: Measurements in different canopy layers and simulations with a canopy model. *Trees* **1986**, *1*, 25–41. [CrossRef]
26.　Van Tuyl, S.; Law, B.; Turner, D.; Gitelman, A. Variability in net primary production and carbon storage in biomass across Oregon forests—an assessment integrating data from forest inventories, intensive sites, and remote sensing. *For. Ecol. Manag.* **2005**, *209*, 273–291. [CrossRef]
27.　Chen, J.M.; Menges, C.H.; Leblanc, S.G. Global mapping of foliage clumping index using multi-angular satellite data. *Remote Sens. Environ.* **2005**, *97*, 447–457. [CrossRef]
28.　Nilson, T. A theoretical analysis of the frequency of gaps in plant stands. *Agric. Meteorol.* **1971**, *8*, 25–38. [CrossRef]
29.　Kang, Y.; Ozdogan, M.; Gao, F.; Anderson, M.C.; White, W.A.; Yang, Y.; Yang, Y.; Erickson, T.A. A data-driven approach to estimate leaf area index for Landsat images over the contiguous US. *Remote Sens. Environ.* **2021**, *258*, 112383. [CrossRef]

30. Le Dantec, V.; Dufrêne, E.; Saugier, B. Interannual and spatial variation in maximum leaf area index of temperate deciduous stands. *For. Ecol. Manag.* **2000**, *134*, 71–81. [CrossRef]

31. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]

32. U.S. Geological Survey, National Elevation Dataset (NED). 2002. Available online: https://apps.nationalmap.gov/services (accessed on 10 June 2022).

33. Hijmans, R.J. Raster: Geographic Data Analysis and Modeling; R Package Version 3.4-13. 2021. Available online: https://CRAN.R-project.org/package=raster (accessed on 10 June 2022).

34. Mikita, T.; Klimánek, M. Topographic Exposure and its Practical Applications. *J. Landsc. Ecol.* **2010**, *3*, 42–51. [CrossRef]

35. Beven, K.J.; Kirkby, M.J. A Physically Based Variable Contributing Area Model of Basin Hydrology. *Hydrol. Sci. Bull.* **1979**, *24*, 43–69. [CrossRef]

36. Tarboton, D.G. *Terrain Analysis Using Digital Elevation Models (TauDEM)*; Utah Water Research Laboratory, Utah State University: Logan, UT, USA, 2016. Available online: https://hydrology.usu.edu/taudem/taudem5 (accessed on 10 June 2022).

37. Yang, L.; Jin, S.; Danielson, P.; Homer, C.; Gass, L.; Bender, S.M.; Case, A.; Costello, C.; Dewitz, J.; Fry, J.; et al. A new generation of the United States National Land Cover Database: Requirements, research priorities, design, and implementation strategies. *ISPRS J. Photogramm. Remote Sens.* **2018**, *146*, 108–123. [CrossRef]

38. Daly, C.; Neilson, R.P.; Phillips, D.L. A Statistical-Topographic Model for Mapping Climatological Precipitation over Mountainous Terrain. *J. Appl. Meteorol.* **1994**, *33*, 140–150. [CrossRef]

39. Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. U.S. General Soil Map (STATSGO2). 2020. Available online: https://sdmdataaccess.sc.egov.usda.gov (accessed on 10 June 2022).

40. 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]

41. Helsel, D.R.; Hirsch, R.M.; Ryberg, K.R.; Archfield, S.A.; Gilroy, E.J. *Statistical Methods in Water Resources*; Techniques and Methods; U.S. Geological Survey: Reston, VA, USA, 2020; p. 484.

42. McLeod, A.I. Kendall: Kendall Rank Correlation and Mann-Kendall Trend Test. 2011. Available online: https://CRAN.R-project.org/package=Kendall (accessed on 10 June 2022).

43. Addor, N.; Newman, A.J.; Mizukami, N.; Clark, M.P. The CAMELS data set: Catchment attributes and meteorology for large-sample studies. *Hydrol. Earth Syst. Sci.* **2017**, *21*, 5293–5313. [CrossRef]

44. A De Cicco, L.; Hirsch, R.M.; Lorenz, D.; Watkins, D. dataRetrieval. Available online: https://code.usgs.gov/water/dataRetrieval/-/tree/v2.7.11 (accessed on 10 June 2022).

45. Hufkens, K.; Basler, D.; Milliman, T.; Melaas, E.K.; Richardson, A.D. An integrated phenology modelling framework in r. *Methods Ecol. Evol.* **2018**, *9*, 1276–1285. [CrossRef]

46. White, J.D.; Running, S.W.; Nemani, R.; E Keane, R.; Ryan, K.C. Measurement and remote sensing of LAI in Rocky Mountain montane ecosystems. *Can. J. For. Res.* **1997**, *27*, 1714–1727. [CrossRef]

47. Landuyt, D.; Perring, M.; Seidl, R.; Taubert, F.; Verbeeck, H.; Verheyen, K. Modelling understorey dynamics in temperate forests under global change–Challenges and perspectives. *Perspect. Plant Ecol. Evol. Syst.* **2018**, *31*, 44–54. [CrossRef] [PubMed]

48. Thrippleton, T.; Bugmann, H.; Kramer-Priewasser, K.; Snell, R.S. Herbaceous Understorey: An Overlooked Player in Forest Landscape Dynamics? *Ecosystems* **2016**, *19*, 1240–1254. [CrossRef]

49. Carter, T.A.; Fornwalt, P.J.; Dwire, K.A.; Laughlin, D.C. Understory plant community responses to widespread spruce mortality in a subalpine forest. *J. Veg. Sci.* **2022**, *33*, e13109. [CrossRef]

50. Laughlin, D.C.; Fulé, P.Z. Wildland fire effects on understory plant communities in two fire-prone forests. *Can. J. For. Res.* **2008**, *38*, 133–142. [CrossRef]

51. Tang, H.; Dubayah, R.; Brolly, M.; Ganguly, S.; Zhang, G. Large-scale retrieval of leaf area index and vertical foliage profile from the spaceborne waveform lidar (GLAS/ICESat). *Remote Sens. Environ.* **2014**, *154*, 8–18. [CrossRef]

52. Sabor, A.A.; Radeloff, V.C.; McRoberts, R.E.; Clayton, M.; Stewart, S.I. Adding uncertainty to forest inventory plot locations: Effects on analyses using geospatial data. *Can. J. For. Res.* **2007**, *37*, 2313–2325. [CrossRef]

53. Goeking, S.A.; Tarboton, D. Data for Spatially distributed overstory and understory leaf area index estimated from forest inventory data. HydroShare. 2022. Available online: https://www.hydroshare.org/resource/ff7ced18a3234f63b9c3cfae03702c30/ (accessed on 10 June 2022).