

University of Nebraska - Lincoln

DigitalCommons@University of Nebraska - Lincoln

---

USDA Forest Service / UNL Faculty Publications

U.S. Department of Agriculture: Forest Service --  
National Agroforestry Center

---

2016

# Refined forest land use classification with implications for United States national carbon accounting

John W. Coulston

*USDA Forest Service*, [jcoulston@fs.fed.us](mailto:jcoulston@fs.fed.us)

Christopher W. Woodall

*USDA Forest Service*

Grant M. Domke

*USDA Forest Service*

Brian F. Walters

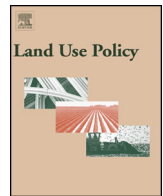
*USDA Forest Service*

Follow this and additional works at: <http://digitalcommons.unl.edu/usdafsfacpub>

---

Coulston, John W.; Woodall, Christopher W.; Domke, Grant M.; and Walters, Brian F., "Refined forest land use classification with implications for United States national carbon accounting" (2016). *USDA Forest Service / UNL Faculty Publications*. 302.  
<http://digitalcommons.unl.edu/usdafsfacpub/302>

This Article is brought to you for free and open access by the U.S. Department of Agriculture: Forest Service -- National Agroforestry Center at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in USDA Forest Service / UNL Faculty Publications by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.



# Refined forest land use classification with implications for United States national carbon accounting



John W. Coulston<sup>a,\*</sup>, Christopher W. Woodall<sup>b</sup>, Grant M. Domke<sup>c</sup>, Brian F. Walters<sup>c</sup>

<sup>a</sup> USDA Forest Service, Southern Research Station, 1710 Research Center Drive, Blacksburg, VA, 24060, USA

<sup>b</sup> USDA Forest Service, Northern Research Station, Durham, NH, USA

<sup>c</sup> USDA Forest Service, Northern Research Station, St. Paul, MN, USA

## ARTICLE INFO

### Article history:

Received 1 July 2015

Received in revised form

23 September 2016

Accepted 1 October 2016

Available online 7 October 2016

### Keywords:

Carbon storage

Woodland forests

Random forests models

Grasslands

## ABSTRACT

The United States provides annual estimates of carbon sources and sinks as part of its National Greenhouse Gas Inventory (NGHGI). Within this effort, carbon stocks and fluxes are reported for six land use categories that are relevant to economic sectors and land use policy. The goal of this study is to develop methodologies that will allow the US to align with an internationally agreed upon forest land use definition which requires forest to be able to reach 5 m in height at maturity. Models to assess height potential are available for a majority of US forests except for woodland ecosystems. We develop a set of models to assess height potential in these systems. Our results suggest that ~13.5 million ha of forests are unlikely to meet the international definition of forests due to environmental limitations to maximum attainable height. The incorporation of this height criteria in the NGHGI results in a carbon stock transfer of ~848 Tg from the forest land use to woodland land use (a sub-category of grasslands) with minimal effect on sequestration rates. The development of a forest land use definition sensitive to climatic factors in this study enables a land use classification system that can be responsive to climate change effects on land uses themselves while being more consistent across a host of international and domestic carbon reporting efforts.

© 2016 Published by Elsevier Ltd.

## 1. Introduction

As signatories to the United Nations Framework Convention on Climate Change (UNFCCC), the United States (US) provides annual estimates of carbon (C) sources and sinks from 1990 to the present following prescribed Intergovernmental Panel on Climate Change (IPCC) good practice guidance (IPCC, 2006; USEPA, 2014) that forms a compendium referred to as the National Greenhouse Gas Inventory (NGHGI) (Woodall et al., 2012). Within the terrestrial components of the NGHGI (as opposed to fossil fuel sources), there resides an important requirement to delineate C stocks and flux by categories of land use, land use change, and forestry. This particular analysis requires the assessment of C by six general land use categories (settlements, grasslands, croplands, wetlands, forests, and other).

In the US, the forest land use category is of critical importance as it accounts for the vast majority (>80%; USEPA, 2014) of the net sequestration of C among all land uses and represents an offset

of annual CO<sub>2</sub> emissions from fossil fuel burning in the US (Joyce et al., 2014). The IPCC good practice guidance (IPCC, 2006) does not dictate the definition of forest land use; rather, it instructs signatories to rely upon their domestic definition. However, the IPCC (2006) guidance suggests that the land use classification should not be influenced by 'rotational or cyclical patterns of land use (e.g., the harvest-regrowth cycle in forestry, or managed cycles of tillage intensity in cropland)'. Further, 'forest land includes systems with vegetation that currently fall below, but are expected to exceed, the threshold of the forest land category'. In accordance with IPCC guidelines the US has adopted the forest land use definition used by the US Department of Agriculture's (USDA) Forest Service, Forest Inventory and Analysis (FIA) program (Smith et al., 2013).

As the recognition of the suite of ecosystem services provided by vegetation has increased (e.g., clean air and water in addition to C sequestration) the need to more objectively delineate between land uses has concomitantly increased beyond that of the NGHGI. In the US, a variety of reporting and domestic policy initiatives have provided impetus to more objectively delineate ecosystem services provided by the variety of land uses in order to facilitate their conservation and monitoring. For example, the Montreal Process Criteria and Indicators evaluate a suite of environmental and

\* Corresponding author.

E-mail address: [jcoulston@fs.fed.us](mailto:jcoulston@fs.fed.us) (J.W. Coulston).

social aspects of US forests (USDA, 2011). The Forest and Rangeland Renewable Resources Planning Act (RPA) of the US requires a comprehensive summary and projection of US forest resources every 10 years (Smith et al., 2009) with updates every 5 years. The US also delineates forest land uses in the Food and Agriculture Organization (FAO) of the United Nations Forest Resource Assessment (FRA, FAO, 2015). Each of these efforts uses a slightly different definition of forest land use which creates inconsistency; however, the FAO forest land use definition is applicable to most of these reporting initiatives.

In regards to domestic US environmental policies, recent executive and legislative guidance has elevated the need to more clearly delineate the ecosystem services provided by woody vegetation among land uses. President Obama's Climate Action plan calls for refining the monitoring of C sequestered among land uses of the US (EOP, 2013). The Agricultural Act of 2014 specifically requires the USDA to identify the capacity and resources needed for refining estimation of forest C and biomass across the US in addition to trees in non-forest land uses such as settlements (US Public Law 113-79). Given the requirement to report ecosystem services such as C sequestration among different land uses for a variety of domestic and international efforts, the likelihood has increased that differing definitions of land uses will result in conflicting estimates of ecosystem services which in turn makes effective rural land policy approaches more difficult to identify. The presence of a variety of estimates has the potential to confound the management, monitoring, and policy development of natural resources. Therefore, the consistent delineation of land uses is needed, especially those that provide the critical function of C sequestration. As an initial step to meet this need, a consistent definition of forest land use should be developed in a fashion that can be implemented across a variety of assessment mechanisms.

Modifying the criteria used to delineate land uses must be done in a consistent fashion for each reporting year as inconsistency may result in misrepresented baselines and unreliable trend information (Grainger, 2008). With the increase in broad-scale forest information the opportunity to draw different inferences regarding status and change in those resources also increases (Coulston et al., 2014; Mather, 1992). Therefore, refined land use criteria must be applicable to the time-series of data that may arise from different sample designs and protocols over time.

The goal of this study is to refine the delineation between forest and non-forest land uses using the FAO forest land use definition for the purpose of improving the consistency of the US' NGHGI estimates with domestic and international reporting instruments with specific objectives being to: (1) develop empirical tree height models as a means to employ an in situ forest land use definition that can be consistently implemented across a range of monitoring mechanisms, across time, and sensitive to climatic attributes (e.g., NGHGI and FRA), and (2) to quantify the implications of this study's refined forest land use definition on forest land use estimates of C stock and C stock change in the US.

## 2. Methods

As our goal is to employ a forest land use definition that is relevant to a range of national and international reporting efforts, we selected the definition used by FAO (2015). The current US definition (developed by FIA) requires land area to have a minimum of 10% tree cover with an areal extent of at least 0.4047 ha with a minimum width of 36.6 m. Further, if the land has less than 10% tree cover it must have the ability to reach 10% cover in situ and not be subject to any non-forest land use such as agriculture or settlements. The FAO definition is similar but further requires trees to have the capacity to reach 5 m at maturity in situ. To employ the

FAO definition models are needed to determine whether the 5 m tree height threshold can be achieved at the maturity of the forest stand.

The FIA program delineates 151 forest community types in the coterminous US and most of these types have associated tree height models (e.g. Carmean et al., 1989) which can be used to apply the FAO definition. However, there is a lack of height models for community types in arid and semi-arid of the coterminous western US (Fig. 1). We focus on these community types and examine their capacity to obtain a 5 m height at maturity. As a means to incorporate an in situ assessment of tree height at forest stand maturity, we develop height models for each of these woodland forest types (Fig. 1).

### 2.1. Data

For this analysis we used FIA data (USDA 2014a,b), 30 year climate norms 1981–2010 (PRISM Climate Group, 2014), and digital elevation products (USGS, 2011). The FIA program employs a repeated measure rotating panel survey design and the nominal sampling intensity is approximately one 674.5 m<sup>2</sup> ground plot per 2403 ha of land area (Bechtold and Patterson, 2005). Each sample location is classified as either forest land use or non-forest land use (in whole or in part based on FIA's definitions) and those locations meeting the forest land use definition (in whole or in part) have additional measurements taken to quantify percent forest and other salient components of biomass, C, stand structure, community type, and health. Data from the FIA program were the basis for stand height, stand age, community type information, stand physiography, and C stock information. The term stand refers to a contiguous unit of trees of similar species composition (e.g., forest type), age structure, stem density, and other conditions so that it forms a distinguishable unit (Smith, 1986). Carbon stocks included C stored in live trees (above and below ground), C in understory vegetation (above and below ground), C in dead trees (standing and downed), C stored in the litter layer, and C in soil organic matter (see Smith et al., 2013 for background on individual C pool predictions). Total C was the sum of all C pools. Climate norms included average annual maximum temperature, average annual precipitation, degree days above 5 °C, degree days above 5 °C during the growing season. From the climate data a growing season moisture index was also calculated as the ratio of precipitation to potential evapotranspiration (Akin, 1991; Coulston and Riitters, 2005). The Digital elevation data were used to model slope and aspect.

### 2.2. Height models

We used an empirical height modeling approach (Avery and Burkhart, 1994) to predict which woodland forest stands were likely to have the capacity to meet the 5 m threshold in situ. The modeling was a probabilistic approach where the probability of the stand being at least 5 m tall was a function of stand age, site characteristics, and regional characteristics. The parameterized models could then be used to estimate the probability of each stand to reach 5 m at any age. We parameterized both random forest models (Breiman, 2001) and logistic regression models for each of the woodland forest type in Fig. 1 using stand that had not been recently disturbed stands (i.e., stands without significant cutting, fire, insects and/or diseases, etc.). If disturbed stands were included our model would include the effects of disturbance on height, age, and site relationships which was not desirable. The general form of the random forest models was

$$P(ht \geq 5m) = f(\text{age, elev, Tmax, gmi, lat, physio, eco, dd, gdd, precip, slope, trAspect})$$

where  $ht$  = maximum tree height,  $age$  = stand age,  $elev$  = elevation,  $Tmax$  = average maximum temperature,  $gmi$  = the

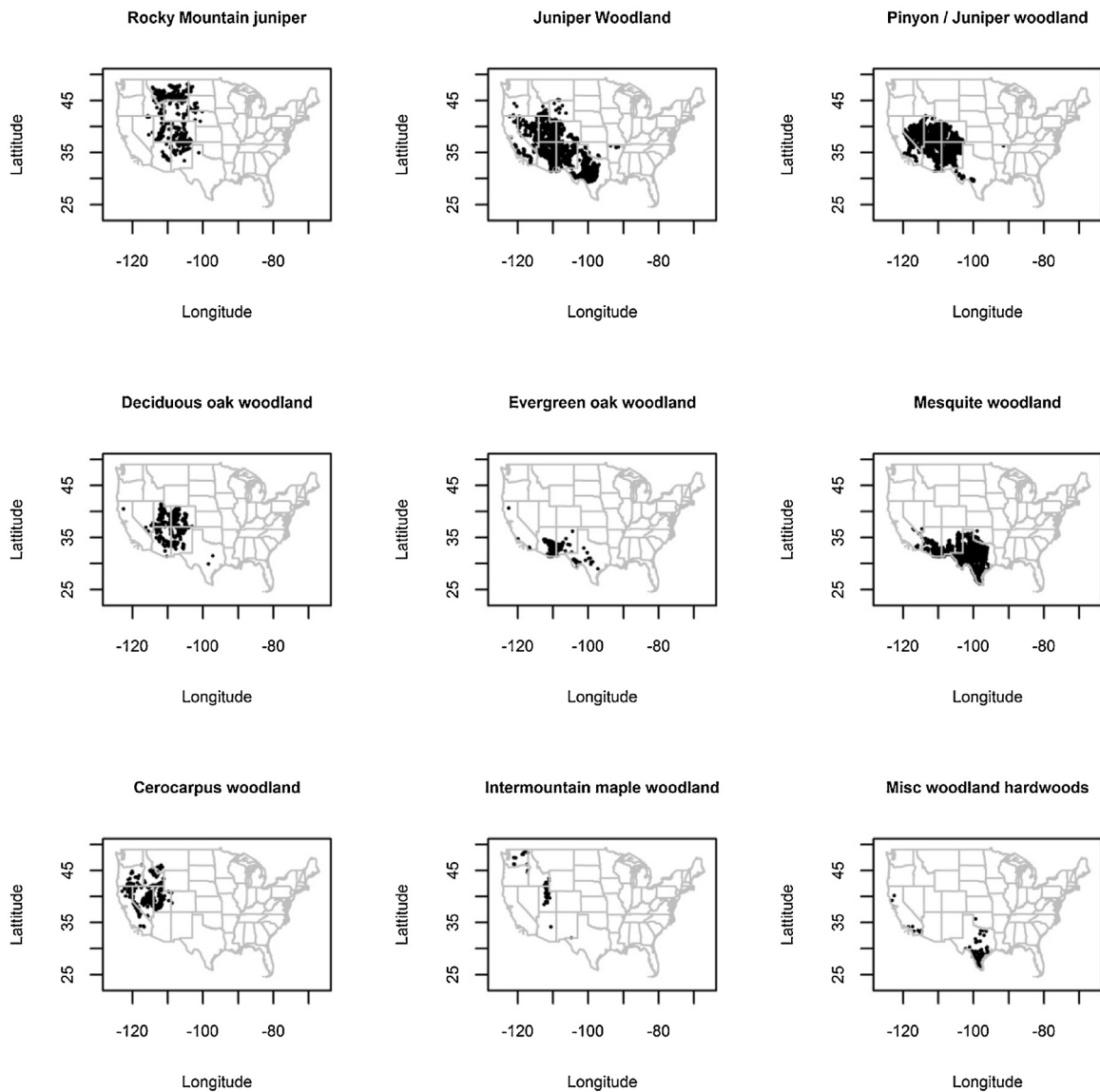


Fig. 1. Spatial distribution of woodland forest types in the coterminous United States.

ratio of precipitation to potential evapotranspiration, *physio* = site physiographic class (xeric, mesic, or hydric), *eco* = ecoregion section, *dd* = degree days, *gdd* = growing degree days, *precip* = precipitation, *slope* = slope, and *trAspect* =  $\cos(\text{Aspect} \cdot 180/\pi)$ .

The general form of the logistic regression models was:

$$P(ht \geq 5m) = \frac{1}{1 + e^{-(a+c \cdot \text{age} + d \cdot \text{elev} + e \cdot \text{precip} + f \cdot \text{gmi} + g \cdot \text{physio})}} + \varepsilon$$

where  $\varepsilon$  = error and all other variables as previously defined. The set of predictor variables used for the logistic modeling approach was selected to reduce correlation among predictor variables.

Random forests is an ensemble method that uses bootstrap aggregating (i.e., bagging) to develop multiple models to improve prediction (Breiman, 2001). Along with bagging random forests also relies on random variable selection to develop a forest of CART-like trees (classification and regression trees). These CART-like trees are uncorrelated. The goal of CART is to understand (learn) the relationship between a dependent variable ( $\mathbf{y}$ ) and a set of predictor variables ( $\mathbf{X}$ ) each of size  $n$ . The learning algorithm employs recursive partitioning which splits the data based on the  $\mathbf{X}$  variables to create homogenous groupings of  $\mathbf{y}$ . The recursive partitioning con-

tinues until either the subset of  $\mathbf{y}$  at each node is the same value or further splitting adds no value. Random forests differs from the CART procedure by (1) employing bootstrap resampling (Efron and Tibshirani, 1993) and (2) random variable selection. Consider a classification tree which is made up of splits and nodes. With random forests a random subset of  $\mathbf{X}$  variables (selected without replacement) is used to determine the split for each node. Call this CART-like model  $\Theta$ . Bootstrap resampling is used to develop  $B$  replicates of  $\Theta$ . Each  $b$  bootstrap sample is selected by sampling  $n$  observation from  $(\mathbf{y}, \mathbf{X})$  with replacement to create  $(\mathbf{y}^b, \mathbf{X}^b)$ . In general, 63% of the original observations will be in the bootstrap sample (in bag) and 37% will be out of bag (OOB) denoted by the superscript  $b$  and  $-b$  respectively.  $\Theta^b$  is then developed for each  $b$  bootstrap sample. The random forest is the ensemble  $\text{RF} = [\Theta^1, \Theta^2, \Theta^3, \dots, \Theta^B]$ .

### 2.3. Model assessment

The Brier score was used to examine the performance of each model and to select a single modeling approach (random forest or logistic regression). The Brier score is a measure of the accu-

racy of probabilistic predictions when prediction probabilities are assigned to a set of mutually exclusive discrete outcomes. The Brier score is the mean squared error in probability space and was defined as:

$$Brier = 1/n \sum_{i=1}^n (P(ht \geq 5m)_i - O_i)^2$$

where  $O_i$  is the observed outcome [ $ht \geq 5m$  (outcome = 1), or  $ht < 5m$  (outcome = 0)] for observation  $i$ . The range of the Brier score is 0 to 1 with 0 representing no error in the predicted probability. We examined the Brier score for the random forest models for each woodland forest type based on both the in bag sample and the out of bag sample. The Brier score was also calculated for the logistic regression models parameterized with the full dataset. To develop an appropriate out of bag Brier score for the logistic regression models we used a bootstrap approach. To accomplish this for each forest type we used the same approach as the random forest model where each  $b$  bootstrap sample was selected by sampling  $n$  observation from  $(\mathbf{y}, \mathbf{X})$  with replacement to create  $(\mathbf{y}^b, \mathbf{X}^b)$ . A logistic regression model was then parameterized based on  $(\mathbf{y}^b, \mathbf{X}^b)$  and we predicted the probability of  $P(ht \geq 5m)$  for each observation in  $\mathbf{X}^{-b}$  (call this  $\hat{\mathbf{y}}^{-b}$ ) and the Brier score was calculated using  $\mathbf{y}^{-b}$  and  $\hat{\mathbf{y}}^{-b}$ . This was repeated 200 times for each forest type and the out of bag Brier score was the mean of the Brier scores across bootstrap replicates for each woodland forest type. Modeling approaches (random forest vs. logistic regression) were compared based on both the in bag and out of bag Brier scores.

#### 2.4. Model application

In order to use the selected site models we developed an optimization approach to convert  $P(ht \geq 5m)$  to a discrete predicted outcome,  $\hat{O}$ . The optimization function was:

$$Pt = \min_{P(ht \geq 5m) \in P} \left( 0.95 - \frac{\sum P \hat{O}}{\sum O} \right)^2$$

where  $Pt$  was the probability threshold,  $P$  ranged from 0 to 1 and all other variables were as previously defined. The optimization was performed for each model. This approach ensured that 95% of the observations where  $ht \geq 5m$  were predicted to have  $ht \geq 5m$ .

To determine which observations, currently <5 m in height, would likely be able to reach the 5 m height threshold in situ we used the models. Note that the models were parameterized with variables that were relatively stable over time (temperature and precipitation norms, slope, and elevation). The exception was stand age. For all observations with a height <5 m we evaluated  $\hat{O}$  based on  $Pt$  at the 1st quartile of age, the mean, and the 3rd quartile of observed stand ages within each woodland forest type.

#### 2.5. Population estimates

Our intent was to quantify the impacts of removing forest stands that did not have the capacity to reach a height of 5 m in situ from the forest land use classification of the US. We examined the impact in terms of forest area, forest C stocks, and expected forest C stock change. For woodland forest observations predicted to be unable to reach the 5 m height threshold we considered the observations “non-forest” and they did not contribute to either forest area or forest C stocks. We used a post-stratified estimator (Bechtold and Patterson, 2005; Cochran, 1977) to construct estimates of forest area and forest C stocks for (1) the original data without removing woodland forest stands unlikely to reach 5 m in height, (2) the data when considering woodland forest stands unlikely to reach

5 m in height at the 1st quartile of age to be non-forest, (3) the data when considering woodland forest types stands to reach 5 m in height at mean age to be non-forest, and (4) the data when considering woodland forest stands unlikely to reach 5 m in height at 3rd quartile of age to be non-forest.

To approximate the influence of removing these lands from forest land use estimates of C stock change we used a simple age-based population model (Coulston et al., 2015). We summarized total woodland forest area ( $\mathbf{W}_t$ ) by 5 year age class from 0 to 295 years based on the observations that were unlikely to reach 5 m in height at mean age. We also summarized C stock density ( $\mathbf{D}_t$ ) in the same fashion. We assumed that the age transition matrix  $\mathbf{T}$  was:

$$\mathbf{T} = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

with this version of  $\mathbf{T}$ , disturbance (e.g. fire, cutting) did not influence the aging process. However, disturbance did influence  $\mathbf{D}_t$  as observed in the data. Forest stands age incrementally until the terminal age class of 295. We set the time step  $m$  to 5 years. The approximate annual C stock change ( $\Delta C$ ) was:

$$\Delta C = ((\mathbf{T} \cdot \mathbf{W}_t)' \cdot \mathbf{D}_t - C_t) / m$$

where  $C_t$  was the total current C stock.

### 3. Results

#### 3.1. Model assessment and selection

We developed nine logistic regression models and nine random forest models; two for each woodland forest type. When considering the logistic regression models, age was the most significant predictor across models (probability of a greater Z score typically less than 0.01) followed by *elev* (probability of a greater Z score typically less than 0.1). The Brier score based on the in bag assessment ranged between 0.025 and 0.202 for the intermountain maple woodlands and the deciduous oak woodland types respectively (Fig. 2). Based on the out of bag assessment the Brier score ranged from 0.06 to 0.22 for the intermountain maple woodlands and the miscellaneous woodland hardwood types respectively (Fig. 2).

The random forest models were relatively similar in terms of important predictors and out of bag Brier scores. The strongest predictors (determined by Gini importance) across woodland types were *age*, *precip*, *gmi*, and *elev*. Based on the in bag samples the Brier score was less than 0.04 for all models, with the lowest being for intermountain maple woodlands (Fig. 2). The Brier score based on the out of bag samples ranged between 0.03 and 0.216 for the intermountain maple woodland type and the miscellaneous woodland hardwoods respectively.

The random forest models had lower Brier scores than the logistic regression models across woodland types based on the in bag assessment. However, the out of bag assessment was more relevant for selecting a single modeling approach. Based on the out of bag Brier scores the random forest models typically outperformed the logistic regression models across woodland forest types. The exceptions were the Rocky Mountain juniper and evergreen oak woodland forest types. Because of the overall better performance of the random forest models we selected those models for subsequent analysis.



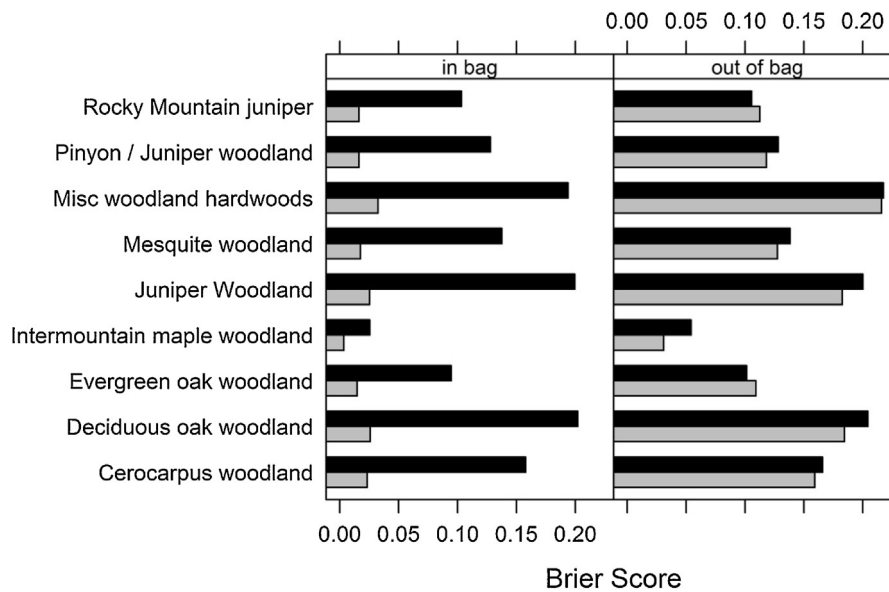


Fig. 2. In bag and out of bag Brier scores for the site model for each woodland forest type based on logistic regression (black bars) and random forest (gray bars).

### 3.2. Model application

The probability threshold ( $P_t$ ) was optimized for each model in order to assure that 95% of the observed stand  $\geq 5$  m were predicted to be  $\geq 5$  m.  $P_t$  ranged from 0.775 mesquite woodland forest to 0.99 for intermountain maple woodlands. For the other woodland forest type  $P_t$  was between 0.78 and 0.98. This threshold was used to evaluate woodland forest type stands currently  $< 5$  m tall at the 1st quartile of age, mean age, and 3rd quartile of age. For example, consider juniper woodland forests which had a mean age of 99 years. To apply the model we used all predictor variables from the original data except age. We then evaluated  $P(ht \geq 5 \text{ m})$  for ages 5–140 years. We examined whether each stand exceeded  $P_t$  (0.78 for juniper woodland forests) at the mean age (See Fig. 3 for example of 9 randomly selected stands). In the juniper woodland forest example (Fig. 3) we note that one of the stands exceed the  $P_t = 0.78$  at 99 years. These two stands would be considered forest land use under the FAO definition with the others considered non-forest.

### 3.3. Population estimates

To understand the implications of aligning the US' forest land use definition with FAO standards we examined estimates of population totals of area, C stocks, and C stock change. Results were similar for area and C stocks regardless of whether we used the 1st quartile, mean, or 3rd quartile of age in our assessment (Table 1). Given the similarity we focused on the results arising from examining the data when considering woodland forest types unlikely to reach 5 m in height at mean age to be non-forest land use.

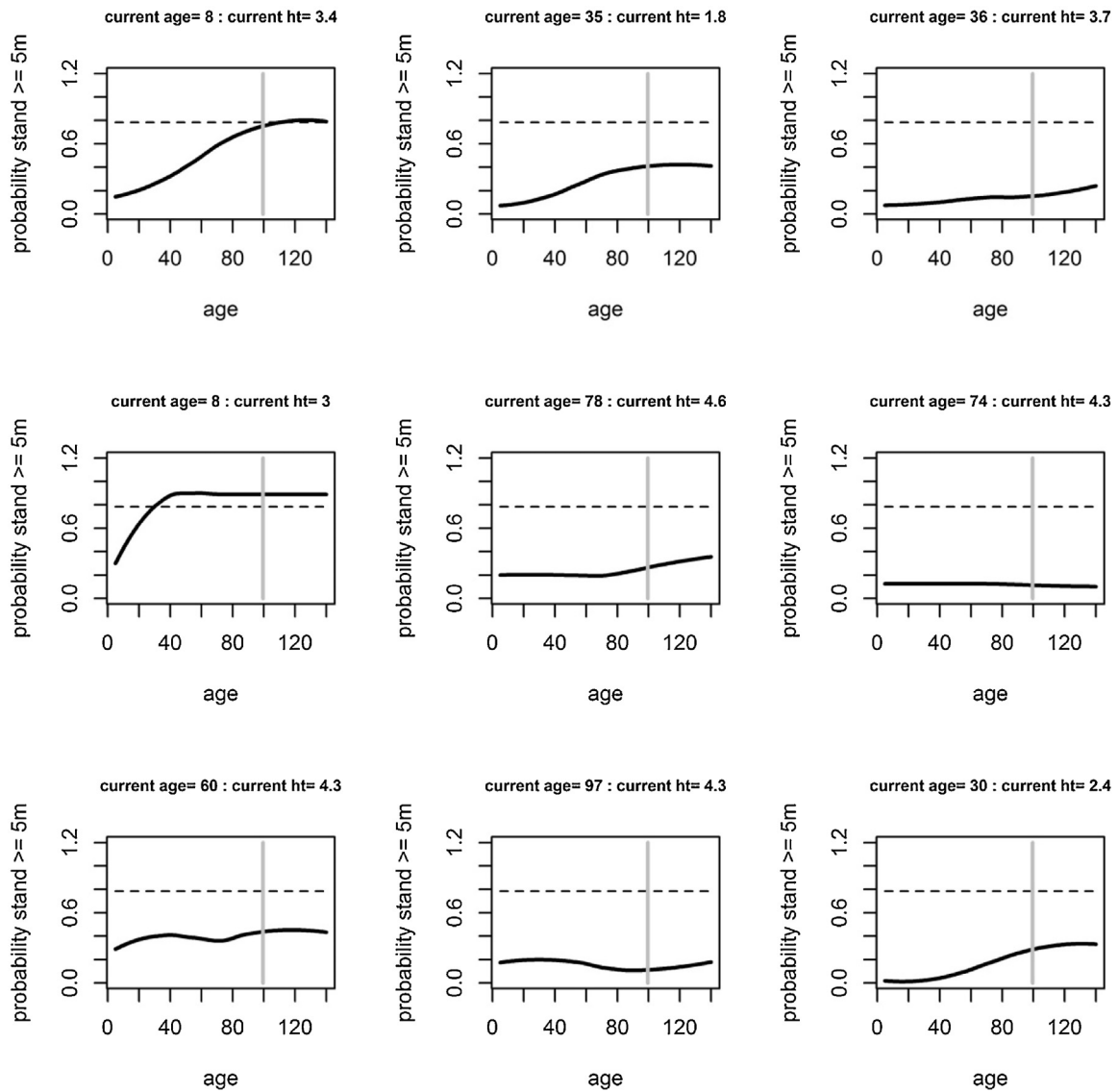
Considering woodland forest types that were not likely to reach 5 m in height as non-forest reduced total estimated forest land use area of the coterminous United States by approximately 4.75% (from 284.4 million ha to 270.9 million ha) (Table 1). Adjusting the forest land base also impacted total estimated C storage in the United States. The largest impact was on above and below ground C stocks in tree species in woodland forest types where in both cases C stocks would be reduced by approximately 12%. Carbon stocks in the litter and above and below ground understory would also be reduced by  $\sim 4.26\%$ – $6.58\%$ . Above and below ground C stocks for all live tree species were reduced by  $\sim 0.35\%$ . Total US forest C stock was reduced by 2.01%.

While removing woodland forests that were unlikely to reach the 5 m height threshold from the forest land use base reduced the total C stock by 2.01%, annual C stock change was relatively unaffected. Based on our age-based population model these woodlands may be a net source of C ( $-0.144 \text{ Tg C yr}^{-1}$ ). This was due to the disturbance in these areas and the relatively constant C stock densities across age classes. The mean stock density across age classes was  $57.2 \text{ Mg C ha}^{-1}$  and an interquartile range of  $3.96 \text{ Mg C ha}^{-1}$ . Based on the slow C accumulation rate of these areas (as compared to temperate forests) we expect C stock change for forest remaining forest in the US to be minimally affected.

## 4. Discussion

Here we present a consistent and biologically relevant technique to separate forest land use from non-forest land using an international definition of forests that incorporates climate information. Further, we quantified the effects of implementing the addition of the tree height requirement as part of the forest land definition on both C stock and C stock change for the US. To our knowledge this is the first broad-scale effort to develop height models that are sensitive to climate as a means to delineate forests from woodlands in the context of a NGHGI. These models can be applied to data collected by the FIA program under the current statistical design. As Grainger (2008) suggests, this is key in ensuring that baseline and trend information remains reliable. Beyond implementation in the US NGHGI, the parsimonious techniques forwarded in this study should be broadly applicable to NGHGIS in other nations where climate data and standard forest inventories are available.

Our results suggest that there are approximately 13.5 million ha of woodland forests that are unlikely to reach the 5 m tree height threshold to be considered part of the forest land use base in the US' NGHGI. Under IPCC land use definitions these areas would be classified as grassland and under FAO definitions these area would be classified as other wooded lands. If only interpreted on the basis of areal extent, this would appear to be a substantial reduction of the US forest to C sink strength, but because of low productivity rates these areas actually contribute very little to annual forest C sequestration. However, the overall effect of removing these lands on official US C stock change estimates is complicated due to the combination of C sequestration from forests remaining forest (i.e.,



**Fig. 3.** Example of  $P(ht > 5\text{ m})$  by age for nine randomly selected juniper woodland stands. The horizontal dashed line denotes the  $P_t$ , the vertical gray line denotes mean age across all juniper woodland stands, and the solid black line represent the probability of exceeding the 5 m threshold between 5 years and 140 years of stand age.

**Table 1**  
Percent reduction in forest area and C stocks when removing woodland forest stands unlikely to reach 5 m in tree height.

Parameter	Original estimate	% Reduction in estimate Age break point		
		1st quartile	mean	3rd quartile
Forest area (million ha)	284	4.92	4.75	4.71
C total (Tg)	42,171	2.09	2.01	1.99
C all live trees above ground (Tg)	13,162	0.36	0.35	0.35
C all live trees belowground (Tg)	2775	0.38	0.37	0.37
C live trees above ground in woodland types (Tg)	384	12.14	11.95	11.88
C live trees below ground in woodland types (Tg)	86	12.25	12.07	11.99
C standing dead trees (Tg)	1150	0.20	0.20	0.20
C down dead (Tg)	1734	0.98	0.94	0.93
C litter (Tg)	5001	4.32	4.26	4.25
soil organic C (Tg)	17,562	3.05	2.89	2.84
C understory above ground (Tg)	709	6.82	6.58	6.51
C understory below ground (Tg)	79	6.82	6.58	6.51

IPCC land use matrix terminology) with C stock transfers resulting from other land uses transitioning to forest use and out of forest use to construct stock change estimates (USEPA, 2014). Care must be taken in applying these models to the full time series of C inventory data so that the 848 Tg of C in these areas is not incorrectly treated

as an emission from the forest sector but rather a stock transfer to the IPCC grassland land use category (the appropriate IPCC land use category).

While our efforts focused on the height requirement in woodland forest types, there are other vegetation types that may require

similar analyses. For example, mangrove forests typically occupy a variety of conditions from shallow fresh and salt water wetlands and marshes to dry land in coastal Florida and along the Gulf Coast of the Southern US (Giri et al., 2011). In some cases these community types may not be able to reach the requisite 5 m height threshold (Simard et al., 2006) and may be more appropriately classified under a wetland land use. While our general modeling approach is relevant to this type of system it is likely that other predictor variables, such as salinity and flood frequency (Feller et al., 2002), may be more relevant in model development. We recommend further research on mangrove and other systems in the US (e.g., boreal forests of interior Alaska and alpine forests in the Western US) that may not have the capacity to reach 5 m in height.

The height models developed during this research are sensitive to climate shifts. We have however made our initial assessments of C implications based on current climate but given the expected future climate changes our results should be re-assessed as new climate data become available. This is particularly relevant to the arid and semi-arid systems of the western US where most woodland types exist. For example, Melillo et al., 2014 suggests increased drought, fire, and other disturbances across these areas. Given that moisture is a limiting factor in most of these systems (Coulston et al., 2010; Floyd et al., 2009; Marlon et al., 2012) future evaluations will need to be made as dominant tree species may change and disturbance impacts become apparent. Further, a re-evaluation of results under current and potential future climate will elucidate potential rural land use implications of plausible climate change impacts.

As signatories to the UNFCCC, the US is required to annually monitor C stocks and fluxes across a matrix of land uses. The woodland forest stands identified in our analysis may be removed from the forest land use and used to inform estimates of C stocks and stock change in the grassland use. Currently, only estimates of soil C stocks and stock changes are reported in the NGHGI for the grassland use (USEPA, 2014) but this work highlights potential C stocks of treed lands within the grassland land use. As this study merely explored the exclusion of some treed areas within woodland forest types from a national inventory of forests, future work should involve exploring opportunities to conduct a consistent national inventory of the tree resource within grasslands to include not only a consistent national inventory but also in situ measurement of carbon pools.

## Acknowledgements

We thank Dr. Chris Edgar (Texas A&M Forest Service) and Dr. Greg Reams (USDA Forest Service) for discussions that led to this research.

## References

- Akin, W.E., 1991. *Global Patterns: Climate, Vegetation, and Soils*. University of Oklahoma Press, Norman, OK, pp. 370p.
- Avery, T.E., Burkhart, H.E., 1994. *Forest Measurements*, 4th ed. McGraw-Hill Inc., New York, NY, pp. 408p.
- Bechtold, W.A., Patterson, P.L. (Eds.), 2005. The Enhanced Forest Inventory and Analysis Program—National Sampling Design and Estimation Procedures. Gen. Tech. Rep. SRS-80. US Department of Agriculture, Forest Service, Southern Research Station, Asheville, NC, 85p.
- Breiman, L., 2001. *Random forests*. *Mach. Learn.* 45 (1), 5–532.
- Carnean, W.H., Hahn, J.T., Jacobs, R.D. 1989. Site index curves for forest species in the eastern United States. USDA Forest Service, North Central Forest Experiment Station Gen. Tech. Rep. NC-128. St. Paul, MN. 142p.
- Cochran, W.G., 1977. *Sampling Techniques*, 3rd ed. Wiley, New York, NY, pp. 428p.
- Coulston, J.W., Riitters, K.H., 2005. Preserving biodiversity under current and future climates: a case study. *Global Ecol. Biogeogr.* 14, 31–338.
- Coulston, J.W., Oswalt, S.N., Carraway, A.B., Smith, W.B., 2010. Assessing forest land area based on canopy cover in a semi-arid region: a case study. *Forestry* 83, 143–151.
- Coulston, J.W., Reams, G.A., Wear, D.N., Brewer, C.K., 2014. An analysis of forest land use, forest land cover and change at policy-relevant scales. *Forestry* 87, 267–276.
- Coulston, J.W., Wear, D.N., Vose, J.M., 2015. Complex forest dynamics indicate potential for slowing carbon accumulation. *Sci. Rep.* 5, 8002, <http://dx.doi.org/10.1038/srep08002>.
- Executive Office of the President, White House, 2013. President's Climate Action Plan, Available at: <http://www.whitehouse.gov/sites/default/files/image/president27sclimateactionplan.pdf>.
- Efron, B., Tibshirani, R.J., 1993. *An Introduction to the Bootstrap*. Chapman & Hall, New York, NY, pp. 436p.
- Food and Agriculture Organization of the United Nations, 2015. *Global Forest Resources Assessment 2015: Desk Reference*. Food and Agriculture Organization of the United Nations, Rome, Italy, pp. 244p, ISBN 978-92-5-108826-5.
- Feller, I.C., McKee, K.L., Whigham, D.F., O'Neill, J.P., 2002. Nitrogen vs. phosphorus limitation across an ecotonal gradient in a mangrove forest. *Biogeochemistry* 62, 145–1175.
- Floyd, M.L., Clifford, M., Cobb, M.S., Hanna, D., Delph, R., Ford, P., Turner, D., 2009. Relationship of stand characteristics to drought-induced mortality in three Southwestern pinon-juniper woodlands. *Ecol. Appl.* 19, 1223–1230.
- Giri, C., Ochieng, E., Tieszen, L.L., Zhu, Z., Singh, A., Loveland, T., Masek, J., Duke, N., 2011. Status and distribution of mangrove forests of the world using each observation satellite data. *Global Ecol. Biogeogr.* 20, 154–159.
- Grainger, A., 2008. Difficulties in tracking the long-term global trend in tropical forest area. *Proc. Natl Acad. Sci. U. S. A.* 105, 818–823.
- Intergovernmental Panel on Climate Change (IPCC), 2006. Good Practice Guidance for Land Use, Land-Use Change and Forestry. <http://www.ipcc-nggip.iges.or.jp/public/2006gl/vol4.html>.
- Joyce, L.A., Running, S.W., Breshears, D.D., Dale, V.H., Malmshiemer, R.W., Sampson, R.N., Sohngen, B., Woodall, C.W., 2014. Ch. 7: Forests. *Climate Change Impacts the United States: The Third National Climate Assessment*. In: Melillo, J.M., Richmond, Terese (T.C.), Yohe, G.W. (Eds.). U.S. Global Change Research Program, pp. 175–194.
- Marlon, J.R., Bartlein, P.J., Gavin, D.G., Long, C.J., Anderson, R.S., Briles, C.E., Brown, K.J., Colombaroli, D., Hallett, D.J., Power, M.J., Schrif, E.A., Walsh, M.K., 2012. Long-term perspectives on wildfires in the western USA. *Proc. Natl Acad. Sci. U. S. A.*, E535–E543.
- Mather, A.S., 1992. The forest transition. *Area* 24, 367–379.
- Melillo, J.M., Richmond, Terese (T.C.), Yohe, G.W. (Eds.), 2014. U.S. Global Change Research Program, p. 841pp, <http://dx.doi.org/10.7930/J0Z31WJ2>.
- PRISM Climate Group, 2014. 2.5-Arcmin (4 [Online] <http://www.prism.oregonstate.edu>).
- Simard, M., Zhang Keqi, Rivera-Monroy, V.H., Ross, M.S., Ruiz, P.L., Castaneda-Moya, E., Twilley, R.R., Rodriguez, E., 2006. Mapping height and biomass of mangrove forests in Everglades National Park with SRTM elevation data. *Photogramm. Eng. Remote Sens.* 72, 299–311.
- Smith, W.B., Miles, P.D., Perry, C.H., Pugh, S.A. 2009. Forest resources of the United States, 2007. USDA, Forest Service, Gen. Tech. Rep. WO-78, Washington, D.C. 336p.
- Smith, J.E., Heath, L.S., Hoover, C.M., 2013. Carbon factors and models for forest carbon estimates for the 2005–2011 National Greenhouse Gas Inventories of the United States. *For. Ecol. Manage.* 307, 7–719.
- Smith, D.M., 1986. *The Practice of Silviculture*, 8th ed. John Wiley & Sons, New York, NY, pp. 527p.
- U.S. Public Law 113-79, 2014. Agricultural Act of 2014. Act of February 7, 2014, Online: <http://www.gpo.gov/fdsys/pkg/BILLS-113hr2642enr/pdf/BILLS-113hr2642enr.pdf>.
- USDA Forest Service, 2011. *National Report on Sustainable Forests—2010*. USDA Forest Service, FS-979 Washington, DC214p.
- USDA, Forest Service, 2014a. Forest Inventory and Analysis National Program – Data and Tools – FIA Data Mart, FIADB Version 5.1. U.S. Department of Agriculture, Forest Service, Washington, DC <http://apps.fs.fed.us/fiadb-downloads/datamart.html>.
- USDA, Forest Service, 2014b. The Forest Inventory and Analysis Database: Database Description and User Guide Version 6.0 for Phase 2. U.S. Department of Agriculture, Forest Service, Washington, DC <http://www.fia.fs.fed.us/library/database-documentation/current/ver6.0/FIADB.user%20guide.6-0.p2.5-6-2014.pdf>.
- US Environmental Protection Agency, 2014. *Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990–2012*. Environmental Protection Agency EPA 430-R-14-003, Washington DC, pp. 529p.
- US Geological Survey, 2011. Global multi-resolution terrain elevation data 2010 [Online] <https://lta.cr.usgs.gov/GMTED2010>.
- Woodall, C.W., Domke, G.M., MacFarlane, D.W., Oswalt, C.M., 2012. Comparing field- and model-based standing dead tree carbon stock estimates across forests of the United States. *Forestry* 85, 125–133.