



# Uncertainty of biomass stocks in Spanish forests: a comprehensive comparison of allometric equations

Aitor Ameztegui<sup>1,2</sup> · Marcos Rodrigues<sup>1,3</sup> · Víctor Granda<sup>4</sup>

Received: 11 October 2021 / Revised: 12 February 2022 / Accepted: 5 March 2022 / Published online: 9 April 2022  
© The Author(s) 2022

## Abstract

Biomass and carbon content are essential indicators for monitoring forest ecosystems and their role in climate action, but their estimation is not straightforward. A typical approach to solve these limitations has been the estimation of tree or stand biomass based on forest inventory data, using either allometric equations or biomass expansion factors. Many allometric equations exist, but very few studies have assessed how the calculation methods used may impact outcomes and how this impact depends on genera, functional group, climate or forest structural attributes. In this study we evaluate the differences in biomass estimates yielded by the most widely used biomass equations in Spain. We first quantify the discrepancies at tree level and among the main forest tree species. We observed that the divergences in carbon estimations between different equations increased with tree size, especially in the case of hardwoods and for diameters beyond the range used to calibrate the equations. At the plot level, we found considerable differences between the biomass values predicted using different methods (above 25% in one out of three plots), which constitutes a warning against the uncritical choice of equations to determine biomass or carbon values. The spatial representation of the differences revealed geographical patterns related to the dominance of fast-growing species such as *Eucalyptus* or *Pinus pinaster*, with a minor effect of forest structure, and almost no effect of climate. Finally, we observed that differences were mostly due to the data source rather than the modelling approach or equation used. Based on our results, BEF equations seem a valid and unbiased option to provide nation-level estimations of carbon balance, although local equations should preferably be used if they are available for the target area.

**Keywords** Allometries · Biomass estimations · Carbon stocks · Forest inventories · Climate change mitigation

## Introduction

Biomass and carbon content are essential indicators for monitoring forest ecosystems and their role in climate action (Ruiz-Peinado et al. 2017). Accurate estimations of forest biomass are critical for assessing carbon balance and greenhouse gases and to inform about the effectiveness of mitigation policies such as REDD+ (Schepaschenko et al. 2019), the Kyoto Protocol (Breidenich et al. 1998) or the Paris Agreement (Roelfsema et al. 2020). Countries adhering to those international agreements commit to conduct regular estimations of forest-related CO<sub>2</sub> release and removals, to account for the role that forests play in the global carbon cycle (Eggleston et al. 2006).

The carbon sink capacity of forests is closely linked to the amount of biomass they harbor. Measuring or estimating the biomass of a tree or forest is not straightforward, and several approaches have been developed, each with its limitations. In situ, direct measurement of forest biomass in

---

Communicated by Peter Annighoefer.

---

✉ Aitor Ameztegui  
aitor.ameztegui@udl.cat

<sup>1</sup> Department of Agriculture and Forest Engineering (EAGROF), University of Lleida, Av. Alcalde Rovira Roure 191, 25198 Lleida, Spain

<sup>2</sup> Joint Research Unit CTFC-AGROTECNIO-CERCA, Ctra. St. Llorenç de Morunys km 2, 25280 Solsona, Catalonia, Spain

<sup>3</sup> GEOFOREST Group, Department of Geography and Land Management, University of Zaragoza, Zaragoza, Spain

<sup>4</sup> Ecosystem Modelling Facility, Centre de Recerca Ecològica i Aplicacions Forestals (CREAF), Bellaterra, Catalonia, ES, Spain

the field requires destructive sampling and extensive field and lab work, which includes the separation and weighting of the different tree compartments (stem, branches, foliage and roots) (Neumann et al. 2016). A typical approach to solve these limitations has been the estimation of tree or stand biomass based on forest inventory data (Chave et al. 2001). Different methods are however available, leading to lack of consistency in the reporting of biomass stocks between and within countries (Neumann et al. 2016; Petersson et al. 2012). The two main methods to convert field measurements into aboveground biomass consist of either allometric biomass equations or biomass expansion factors (BEFs) (Brown, 2002). Allometric equations require information on individual tree dimensions, which are then converted into biomass as a function of diameter at breast height alone or diameter and tree height (Soares and Tomé, 2004). In the latter case, height must be estimated from diameter if it is not measured, thus requiring additional height–diameter allometric equations. The second method consists of converting volume measurements or estimates through the use of biomass expansion factors (BEFs). In either method, biomass estimations can then be transformed into carbon stocks using species- or genus-specific carbon fraction ratios (Brown, 2002).

BEFs are commonly used for greenhouse gas reporting, as recommended by several international agencies, including the IPCC (Eggleston et al. 2006), the FAO (Neumann et al. 2016), and more recently by the Spanish Ministry for the Ecological Transition and the Demographic Challenge (MITECO, 2019), among others. This method presents the advantage of not requiring individual tree information, since stand-level volume information—common in national forest inventories—can be multiplied by species-specific constants to convert to tree biomass and carbon content (Condés and McRoberts, 2017). However, several studies comparing biomass estimates from allometric equations and BEFs highlight that the relationship between volume and biomass (i.e., BEF) is not constant but depends on both intrinsic and extrinsic factors such as tree age, stand volume, site index or climate (Aguirre et al. 2021; Jalkanen et al. 2005; Petersson et al. 2012; Soares and Tomé 2004). To reduce the risk of bias, the use of age-dependent or volume-dependent BEFs is recommended, and biomass allometric equations are preferred when available (Petersson et al. 2012).

Many available predictive allometric equations exist; most of them developed for specific study sites or small regions (Muukkonen, 2007). Likewise, comprehensive reviews and collections of biomass equations for broader regions are available, mostly for North America (Jenkins et al. 2003; Rojas-García et al. 2015; Ter-Mikaelian and Korzukhin, 1997) and Europe (Annighöfer et al. 2016; Muukkonen, 2007; Zianis et al. 2005). The recent Forest Observation System initiative (Schepaschenko et al. 2019) even

aims at building a global dataset of ground-based biomass observations as reference for the calibration and validation of algorithms implemented from the various space-borne missions existent and to come (e.g., GEDI mission from NASA (Dubayah et al. 2020), ESA's BIOMASS mission, (Scipal et al. 2010)). However, these datasets aggregate data from plots contributed by several existing networks, in which biomass is typically not measured but estimated from site-specific, often inconsistent allometric equations (Schepaschenko et al. 2019).

Unless equations for the species or region of interest do exist, the choice often comes out of habit (Muukkonen, 2007), without conducting any systematic evaluation of the impact of the available sources on biomass estimates. Several studies have reported that the use of constant, species-specific BEFs can lead to biased estimates of forest biomass (Castedo-Dorado et al. 2012; Jalkanen et al. 2005; Neumann et al. 2016; Petersson et al. 2012), but very few have assessed how the calculation methods used in allometric equations may impact outcomes (but see, for example Neumann et al. (2016)); or if/how this impact depends on genera, functional group, or forest attributes such as tree size.

In this study we evaluate the differences in biomass estimates yielded by the most widely used biomass equations in Spain, including both allometric equations and BEFs. We first quantify the discrepancies at tree level and among the main forest tree species, providing further insights into the role played by forest attributes (namely, tree size, species and functional group) in the observed discrepancies. Then, we assess the implications of the choice of equations on plot-level biomass estimates for the whole Spanish territory. We expected tree-level differences in biomass to grow with tree size, which would lead to greater differences in mature or large-stocking plots. Finally, we identify the factors that determine the disagreement in biomass estimates, which allows us to endorse specific recommendations on the criteria for choosing the best alternative among the available sources.

## Methods

### Equation data sources

We used the biomass equations most used in Spain. These equations include different sources and data acquisition methods that can be resumed as follows. The first group of sources are based on allometric equations, which estimate aerial biomass—either total biomass or separated by tree compartments—directly from tree diameter and, in some cases, also from tree height. We gathered several equations using this approach: the first set of equations was developed by the National Institute for Agronomic Research (INIA) as

a requirement by the Ministry of Environment. Published as a monograph (Montero et al. 2005), they were the first attempt to provide accurate biomass estimations for the main tree species in Spain and have thus been extensively used since then. Montero et al. (2005) sampled 32 forest species across different provinces in Spain (13 softwood and 19 hardwood species), in order to quantify the carbon stocks of Spanish forests and their potential as carbon sinks. The set of equations obtained from this source are hereafter named "INIA" and used the same log–log equation for all species, where:

$$\ln(\text{biomass}) = a + b \ln(\text{diameter})$$

This approach was later updated by Ruiz-Peinado (2012, 2011), who used the same original dataset to build biomass equations for 10 softwood (Ruiz-Peinado et al. 2011) and 13 hardwood species (Ruiz-Peinado et al. 2012). As opposed to Montero et al. (2005), the new equations by Ruiz-Peinado include total height as a potential independent variable and ensure that the models fulfill the additivity property (i.e., the sum of biomass estimation of the different tree fractions equals the estimation of total aerial biomass). They also considered a wider set of equations, evaluating 15 different biomass models for each component and species. The set of equations obtained from Ruiz-Peinado (2012, 2011) are hereafter named "RUIZ" (Table 1).

A third set of equations was obtained from the Ecological and Forestry Inventory of Catalonia (Inventari Ecologic i Forestal de Catalunya; IEFC). The IEFC was carried out between 1988 and 1998 by CREAM and the Catalan Government and aimed at complementing traditional forest inventories by focusing on environmental issues such as the role of forests in the carbon cycle. It included a total of 10,644 plots randomly distributed within the forested area of Catalonia, at a rate of approximately one plot for every 100 hectares (Vayreda et al. 2005). It incorporated the systematic sampling, in 20% of the plots, of a set of variables including wood, branches, leaves and bark biomass, which were then

used to derive species-specific allometric regressions as a function of tree diameter (we named those "IEFC1") and as a function of diameter and height ("IEFC2") (Ibáñez et al. 2002). It also included volume estimates and biomass expansion factors for the main species (this approach is hereafter named "BEF"), so both approaches were developed using the same dataset. Moreover, it included height–diameter allometric equations that we used to obtain height values for our calculations (see next section). The methodology and results of the IEFC have been published in a series of books and are also available online (<http://www.creaf.uab.es/iefc/>), and the equations and parameters are available in the Catalan Forestry Lab (<https://laboratoriforestal.creaf.cat/>). All the equations used for this work, and the species-specific parameters can be found in the Supplementary Material S1.

## Predicting individual-tree aboveground biomass and differences across sources

We determined total aerial biomass for each source of equations and species in the following manner: we generated tree diameter pseudo-observations as a sequence of diameter values equally spaced at a 1-cm interval ranging from 5 to 80 cm (Muukkonen, 2007), and we computed for each pseudo-observation species-wise values of total aerial biomass according to each of the five equation sources selected and summarized in Table 1. To avoid differences in the results due to the height–diameter allometric equation, the height values for those equations requiring it were always computed in the same way, using the height–diameter equations developed in the IEFC, which determine tree height based on tree diameter (See Supplementary Material S1). We only kept those species that were parameterized for all the sources, leading to a final set of 17 species: 10 hardwoods and 7 softwoods (Table 2). Altogether, these 17 species are widely represented in Spanish forests, being present in 89% of the forest plots and representing 81% of the

**Table 1** Main sources of biomass estimation equation used in this comparative study. For each source, we indicate the main reference, the number of species included, and the equations used, as well as an indicator of model performance

Source Name	Reference	Number of species	Equation used	R2 range
INIA	(Montero et al. 2005)	32	$AGB = a \cdot DBH^b$	0.962 (0.901–0.996)
RUIZ	(Ruiz-Peinado et al. 2011, 2012)	23	Several (see Tables S2 and S3 in Suppl. Mat. 1)	Not reported
IEFC1	(Ibáñez et al. 2002; Vayreda et al. 2005)	34	$AGB = a \cdot DBH^b$	0.945 (0.865–0.988)
IEFC2			$AGB = a \cdot DBH^b \cdot Ht^c$	0.945 (0.865–0.988)
BEF			$AGB = VOB * BEF$	Not reported
Height-diameter		64	$Height = a \cdot DBH^b$	0.496 (0.05–0.89)

*AGB* aboveground biomass, *DBH* diameter at breast height, *VOB* volume over bark, *Ht* total height

**Table 2** List of species included in the final dataset, and information on minimum and maximum diameter used in the calibration sample of each equation

Source	INIA; RUIZ Montero et al. (2005)			IEFC1, IEFC2, BEFIbáñez et al. (2002)		
	N	Dmin	Dmax	N	Dmin	Dmax
<i>Alnus glutinosa</i>	16	8.3	47.3	13	7	25.05
<i>Castanea sativa</i>	24	10.6	50.6	14	7.2	41.8
<i>Eucalyptus spp.</i>	24	9.8	54.0	11	6.3	52.1
<i>Fagus sylvatica</i>	72	9.5	74.8	72	5.7	49
<i>Fraxinus spp.</i>	27	7.2	52.2	20	5	37.1
<i>Populus x canadensis</i>	32	11.3	50.7	13	8.9	50.1
<i>Quercus canariensis</i>	23	10.0	60.0	8	6.8	19
<i>Quercus faginea</i>	24	9.5	46.5	34	5.05	26.9
<i>Quercus ilex</i>	43	7.8	85.9	722	5	36.4
<i>Quercus suber</i>	33	10.5	69.0	50	5.2	24.3
<i>Abies alba</i>	29	9.0	57.5	101	6.15	69.7
<i>Pinus halepensis</i>	55	8.0	44.0	1387	5	46.65
<i>Pinus nigra</i>	50	9.9	77.3	919	5	54.8
<i>Pinus pinaster</i>	198	7.0	64.0	47	6.85	36.1
<i>Pinus pinea</i>	47	7.5	63.0	140	5.6	39.2
<i>Pinus sylvestris</i>	316	6.2	76.0	1422	5	57.4
<i>Pinus uncinata</i>	21	8.0	41.0	417	5.9	57.7

N represents the number of trees used to calibrate the equations for each source and species, whereas Dmin and Dmax are the minimum and maximum diameter (in cm) of the calibration sample

total number of trees according to the Third National Forest Inventory (Dirección General para la Biodiversidad, 2007).

We estimated the pairwise difference between biomass estimations of different sources (i.e., different equations) through the normalized difference, calculated as:

$$\text{Diff}_{\text{eq1-2}} = \frac{\text{AGB}_{\text{eq2}} - \text{AGB}_{\text{eq1}}}{\text{AGB}_{\text{eq2}} + \text{AGB}_{\text{eq1}}}$$

where  $\text{AGB}_{\text{eq1}}$  and  $\text{AGB}_{\text{eq2}}$  are aboveground biomass estimations for equations 1 and 2, respectively. This type of formulation has a series of advantages as compared to the absolute difference, ratios or percentage change: it defines values from  $-1.0$  to  $1.0$ , where values close to zero indicate small differences in biomass estimation between the two sources. Moreover, as opposed to ratios or percentage change, the mean normalized difference is symmetrical, i.e.,  $\text{Diff}_{1-2} = -\text{Diff}_{2-1}$ .

We calculated the mean normalized difference between each pair of sources for each species and diameter in our sequence, and then we assessed the distribution of normalized differences across diameter classes, species and functional groups (hardwoods vs. softwoods). The width of the distribution indicates the range of variability in the normalized differences across all the sources, so that narrower distributions indicate more agreement in biomass estimates between all sources. In turn, the number of peaks of the distribution of normalized differences indicates how the

predictions from different sources are clustered into similar values.

## Plot-level biomass predictions

To assess the source-dependent effects on biomass estimations at the plot level we calculated total aerial biomass for all the plots measured in the Third National Forest Inventory (NFI3) in Peninsular Spain. NFI3 provides a comprehensive data set on forest composition and structure through permanent plots located at the intersections of a  $1 \times 1$  UTM grid and sampled about every 10 years. Each plot is composed of four nested subplots of 5 to 25 m radii in which trees are inventoried depending on diameter at breast height (Alberdi et al. 2017a).

We only kept those plots in which the studied species represented at least 80% of the total basal area, resulting in a final dataset of 61,983 plots (82% out of the 75,629 plots in the original database) containing 1,006,826 measures of individual trees. For each plot, we calculated total plot biomass as the sum of biomass estimations of all the individual trees in the plot, using each of the equation sources described above. We then estimated differences in plot biomass estimations across different sources and represented the results spatially to depict spatial patterns of discrepancies or similarities across equations. In this case we calculate the differences between plots as relative differences, in order

to obtain values that are more easily interpretable at the plot scale, as follows:

$$\text{Eq1 vs Eq2} = 100 * \frac{\text{AGB}_{\text{eq1}} - \text{AGB}_{\text{eq2}}}{\text{AGB}_{\text{eq2}}}$$

Finally, we assessed the importance of different variables as predictors of the pairwise differences in biomass estimation across all the tested sources. For each pairwise difference in plot-level estimates, we fit a random model in which we included as potential predictors variables related to species composition (dominant species, proportion in basal area of the main species), forest structure (mean quadratic diameter, mean tree height, stem density and basal area), climate (mean annual temperature and annual rainfall) and plot topography (elevation, slope) (Table 3). We avoided the potential misspecification of the model due to spatial autocorrelation through bootstrapping: we fitted 100 models each with a subsample of ~ 10% randomly chosen data points, from which 60% were used for calibration and 40% for validation. We retrieved variable importance for each predictor as the mean variable importance across the 100 models, and we calculated the R-squared—using the validation sample—as indicative of model performance. All the analyses were carried out in R version 4.2, and we used the packages *caret* (Kuhn, 2008) and *randomForest* (Liaw and Wiener, 2002) to fit the models.

## Results

### Tree-level estimates of aboveground biomass

The different equations produced similar estimates of aboveground biomass for trees of small diameter, and the discrepancies between them grew with tree size and only emerged

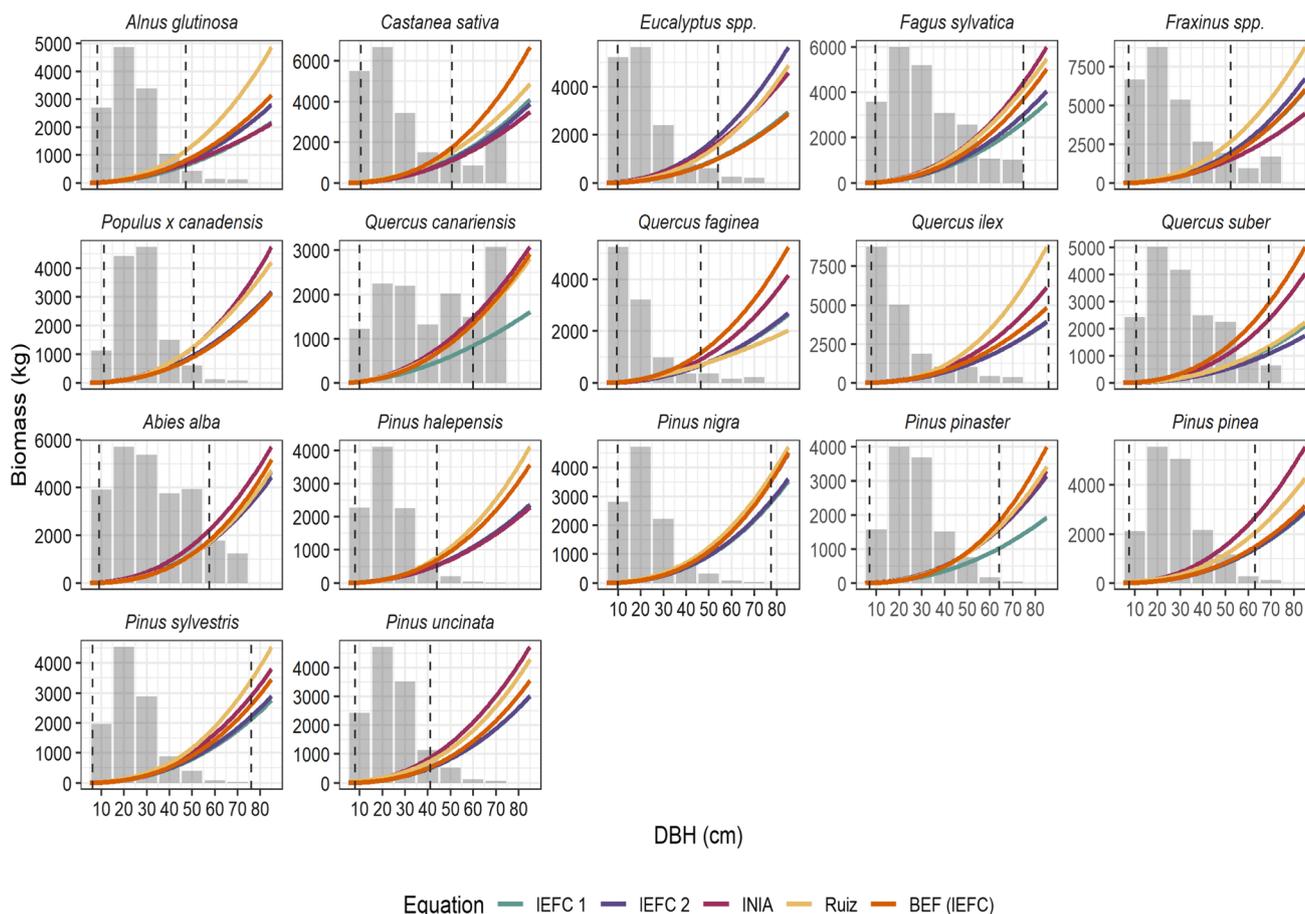
for trees with DBH > 30–40 cm (Fig. 1). None of the equations produced consistently higher biomass values for all species, and the ranking among the different data sources varied markedly among species. For most species, the range of diameters used to calibrate the equations was sufficiently wide as to include more than 80% of the trees measured in the National Forest Inventory. This proportion was lower for *Castanea sativa* (64%), *Quercus faginea* (66%), *Eucalyptus spp.* (72%) and *Quercus canariensis* (78%), mostly due to the presence of many trees with lower DBH than the minimum sampled to calibrate the equations (Table 2). In the case of *Quercus canariensis*, however, up to 9% of the trees in the NFI3 were larger than the largest tree sampled (Fig. 1).

The pattern of normalized differences in biomass estimation along size classes presented a marked species-specific character (Fig. 2). For some species, such as *Abies alba* or *Pinus uncinata*, the normalized differences were not size-dependent, i.e., the differences in biomass estimations produced by the five equation sources were similar regardless tree size. For others (e.g., *Pinus halepensis*, *Pinus pinaster*, *Quercus faginea*) there was a marked size-dependent effect, and the width of the normalized differences widened as tree size increased, creating two to three modes produced by the five equation sources tested.

These patterns became more evident when we analyzed the two main functional groups (Fig. 3). For softwoods we could detect two main modes in the differences, indicating that biomass estimates produced by the five equations were concentrated in two groups with similar values. One of them had mean zero, indicating consistent and similar biomass estimates for at least two of the sources. Moreover, the difference between these two groups was not size-dependent. For hardwoods, on the other hand, there was more discrepancy among biomass estimates, with a wide range of values clustered around the zero mean. These differences, however, increased with tree size, dispersing without clustering around any single value.

**Table 3** Proportion of plots, according to the 3NFI within each class of differences in plot-level biomass estimation

Source comparison	< -50%	(-50, -25]	(-25, -10]	(-10,10]	(10,25]	(25,50]	+50%
IEFC2 versus IEFC1	0.4	7.6	33.9	35.5	10.0	5.9	6.7
INIA versus IEFC1	0.0	0.0	1.7	29.2	24.7	28.7	15.6
RUIZ versus IEFC1	0.1	3.9	7.1	23.5	17.5	30.2	17.6
INIA versus IEFC2	0.0	1.8	5.7	22.6	23.2	23.3	23.5
RUIZ versus IEFC2	0.0	2.0	9.0	22.2	14.8	29.6	22.3
RUIZ versus INIA	0.6	8.9	23.2	38.5	18.4	7.1	3.3
BEFs versus IEFC1	0.0	0.8	3.1	57.7	25.2	7.8	5.4
BEFs versus IEFC2	3.3	4.4	8.0	29.1	23.2	21.1	11.0



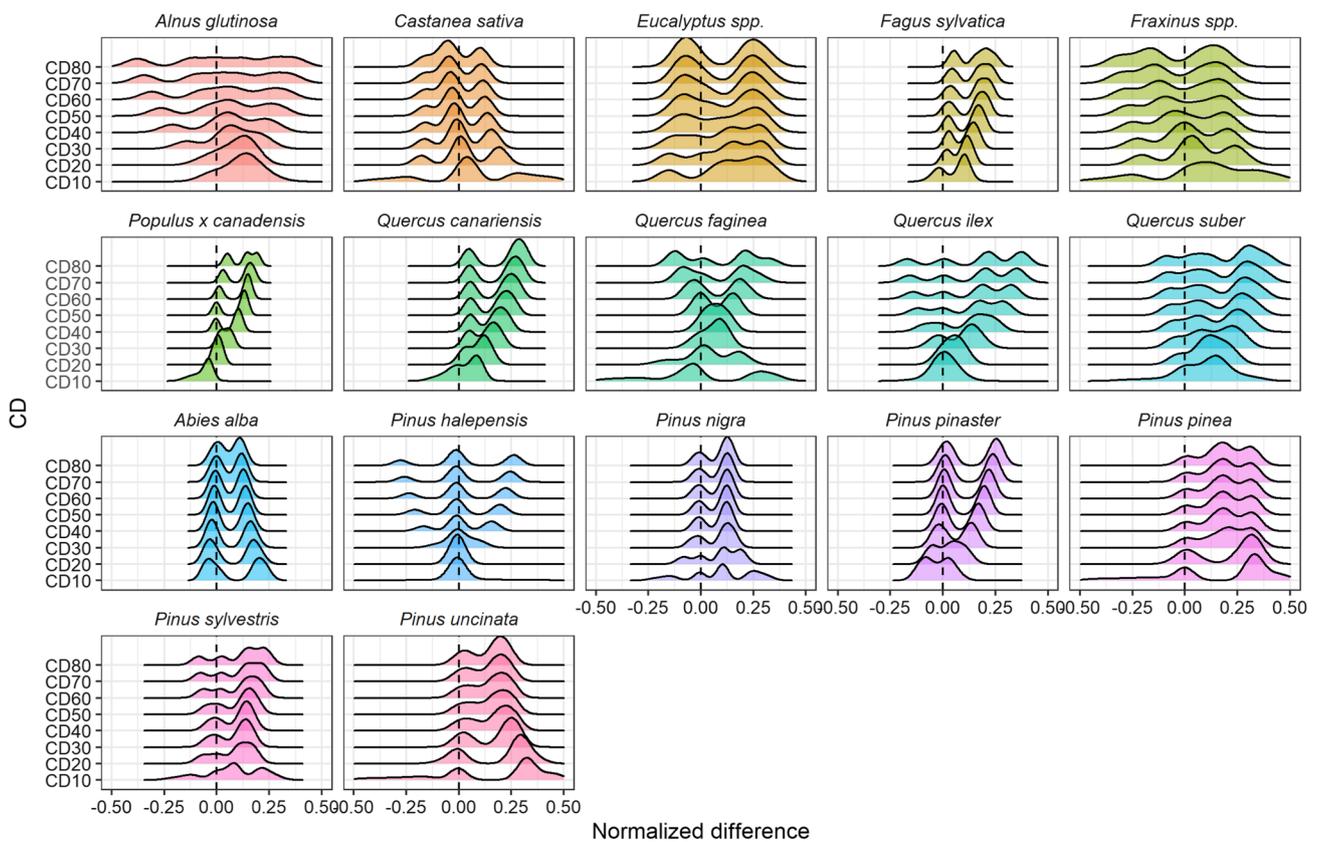
**Fig. 1** Total aerial biomass estimated for each of the main 17 tree species in Spain along a DBH sequence and according to different sources of aboveground biomass equations. Colored lines represent biomass estimations, whereas vertical dashed lines indicate the min-

imum and maximum diameter used for the calibration of the equations. Barplots in the background indicate the diameter distribution of the species according to the Spanish third National Forest Inventory (NFI3)

## Plot-level aboveground biomass

At the national scale, plot-level estimates could be grouped into two large sets that were quite different from each other, confirming the bimodal trend observed when analyzing the differences at the individual tree level. Estimates obtained from INIA and RUIZ were very similar between them, with an average difference of 0.6% (mean biomass content for RUIZ:  $63.8 \text{ Mg}\cdot\text{ha}^{-1}$ ; for INIA:  $62.1 \text{ Mg}\cdot\text{ha}^{-1}$ ), whereas biomass estimates using IEFC's equations—including BEFs—were also very similar and gave the smallest estimate for biomass (IEFC1 =  $48.3 \text{ Mg}\cdot\text{ha}^{-1}$ ; IEFC2 =  $52.4 \text{ Mg}\cdot\text{ha}^{-1}$ ; BEF =  $54.6 \text{ Mg}\cdot\text{ha}^{-1}$ ). However, the mean difference in biomass estimation did not fully capture the variability in discrepancies between sources. In one out of two plots, the differences in biomass estimation between IEFC1 or IEFC2 and those from INIA and RUIZ were larger than 25%, and in one quarter of them they surpassed 50% (Table 3).

The spatial representation of differences in biomass estimation showed that RUIZ and INIA overpredicted IEFC1 and IEFC2 almost in all territories, although for INIA the differences were particularly large in the Western part of Spain (Fig. 4). The main pairwise differences were observed either along the northern fringe—when IEFC1 was compared to any other equation—or in the southwestern part of Spain when INIA and/or RUIZ were compared to IEFC1 and IEFC2. On the other hand, differences between IEFC1 and IEFC2, or between RUIZ and INIA were rather slim in most of the territory and only surpassed 25% in around 10% of the plots (Table 3; Fig. 4).



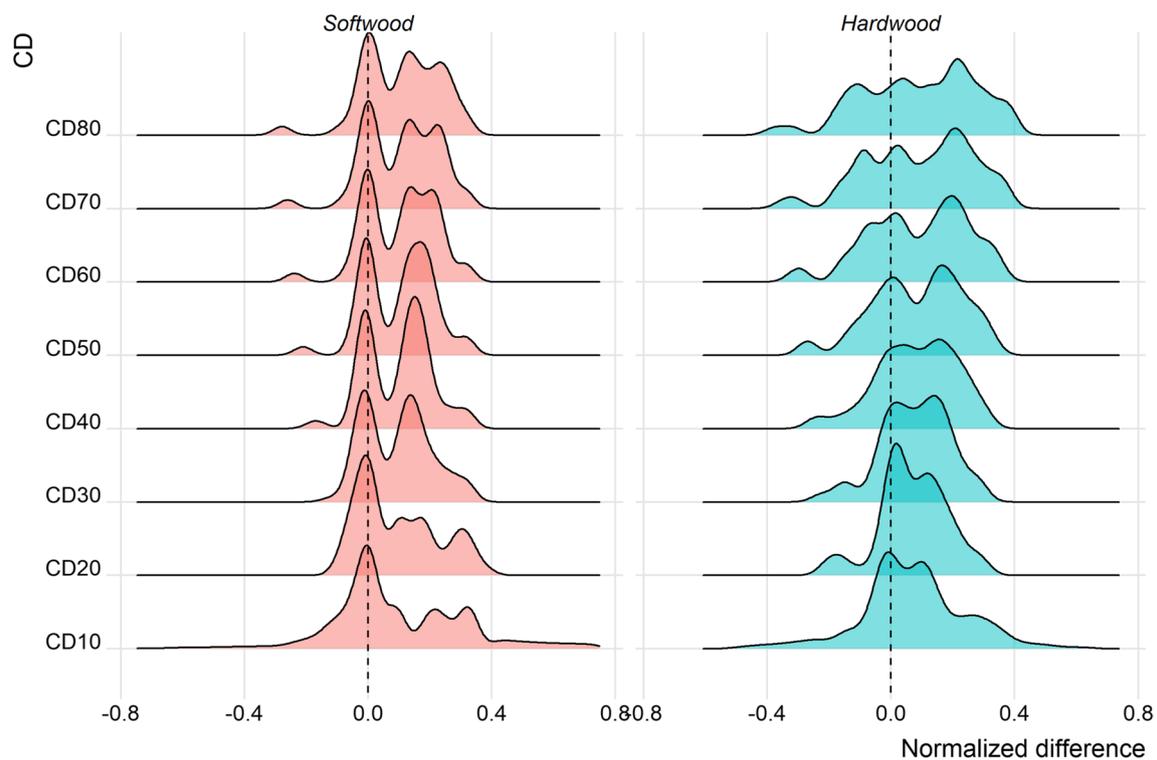
**Fig. 2** Species-specific distribution of pairwise normalized differences in biomass estimation across the five sources tested, along a diameter gradient

## Main factors driving differences in biomass estimation

Random forest models provided insights into the importance of species identity behind the pairwise differences across sources (Fig. 5). The dominance of fast-growing species such as *Eucalyptus spp.* or *Pinus pinaster* was behind the sharp differences detected in the northern fringe between IEFC1 and IEFC2, and between IEFC1 and INIA. However, structural features had a prevalent role in some of the observed differences. In particular, mean tree height was the main driver of the differences when RUIZ was compared to equations not including height as a predictor (i.e., IEFC1 and INIA). Conversely, climatic variables had none or very little importance as predictors of the differences in biomass estimates between sources (Fig. 5). Biomass estimates using BEFs were rather similar to those produced by allometric equations calibrated with the same dataset (IEFC1 and IEFC2). Differences between BEF and IEFC1 were less than 10% in most plots and for most species but *Quercus suber*. Large differences between BEF and IEFC2 were in turn observed for the productive, fast-growing plantations in Northern Spain (Fig. 4).

## Discussion

We found considerable differences between the biomass values predicted using different methods, particularly when escalating the predictions to the plot level. Our findings constitute a warning against the uncritical choice of equations to determine biomass or carbon values. Previous studies already report noteworthy country-specific differences in biomass estimates that have been partially attributed to the use of constant species-specific BEFs (Jalkanen et al. 2005), but also to the use of different sources of data or different mathematical approaches. Our results show that even within one country, the choice of allometric equation can lead to very high disparities—in some cases above 50%—in biomass (and ultimately carbon) estimates at the plot scale. The large discrepancies observed have very relevant implications for the calculation of biomass and carbon accumulated in forests, and for the accurate accounting of carbon balance associated to mitigation initiatives such as the Paris Agreement or the European Union Green Deal (COM/2019/640). Moreover, we could identify some of the main factors behind those disparities, which will eventually help obtain more accurate estimations by considering the



**Fig. 3** Group-specific distribution of pairwise normalized differences in biomass estimation across the five sources tested, along a diameter gradient

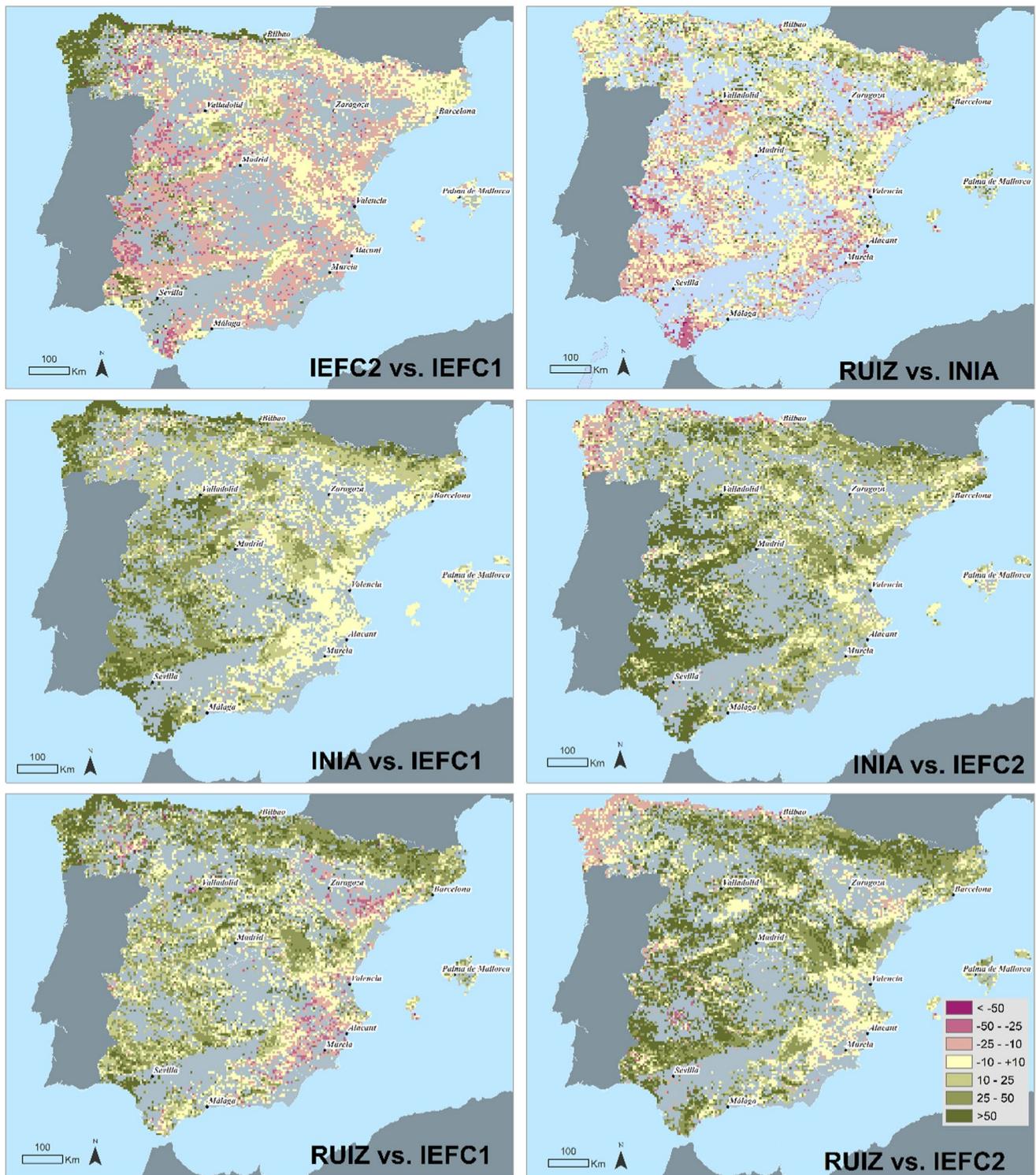
uncertainties associated to the local environment and the main species to assess.

### Effects of functional group and tree size on biomass estimates

The divergences in carbon estimations produced by the different sources increased with tree size for all species, which is common in biomass equations and was expected considering their nonlinear nature (Annighöfer et al., 2016; Zianis and Mencuccini, 2004). However, they increased unevenly depending on the species or the functional group considered. For some species (e.g., *Abies alba*, *Pinus uncinata*, *Pinus nigra*), the differences between equations remained stable along the tree size gradient, whereas for most species the difference grew with size, especially in the case of hardwoods. This trend has also been observed in Sweden, where the errors associated with the biomass calculations at the tree level do not depend so much on the size or age of the trees for pine species as they do for other genera (Jalkanen et al., 2005). Moreover, many hardwood stands in Spain have been managed as coppice systems for centuries and then abandoned in the last decades (Ruiz-Peinado et al., 2012; Valbuena-Carabaña et al., 2010). As a consequence, they often develop large root systems and tortuous growth

patterns in the aerial part, with trunk shapes far from a cylinder or cone. They are also longer-lived than softwoods, often reaching larger sizes, particularly in diameter (Falster et al., 2015). Under these circumstances, it is logical to expect further differences between equations, since each of them is fit to the particularities of the dataset used to calibrate them. A possible solution could be the use of different equations according to the shape of the tree, which is estimated using a shape parameter. This approach is for instance used for the calculation of volumes in the Spanish Forest Inventory and other sources (Alberdi et al. 2017b).

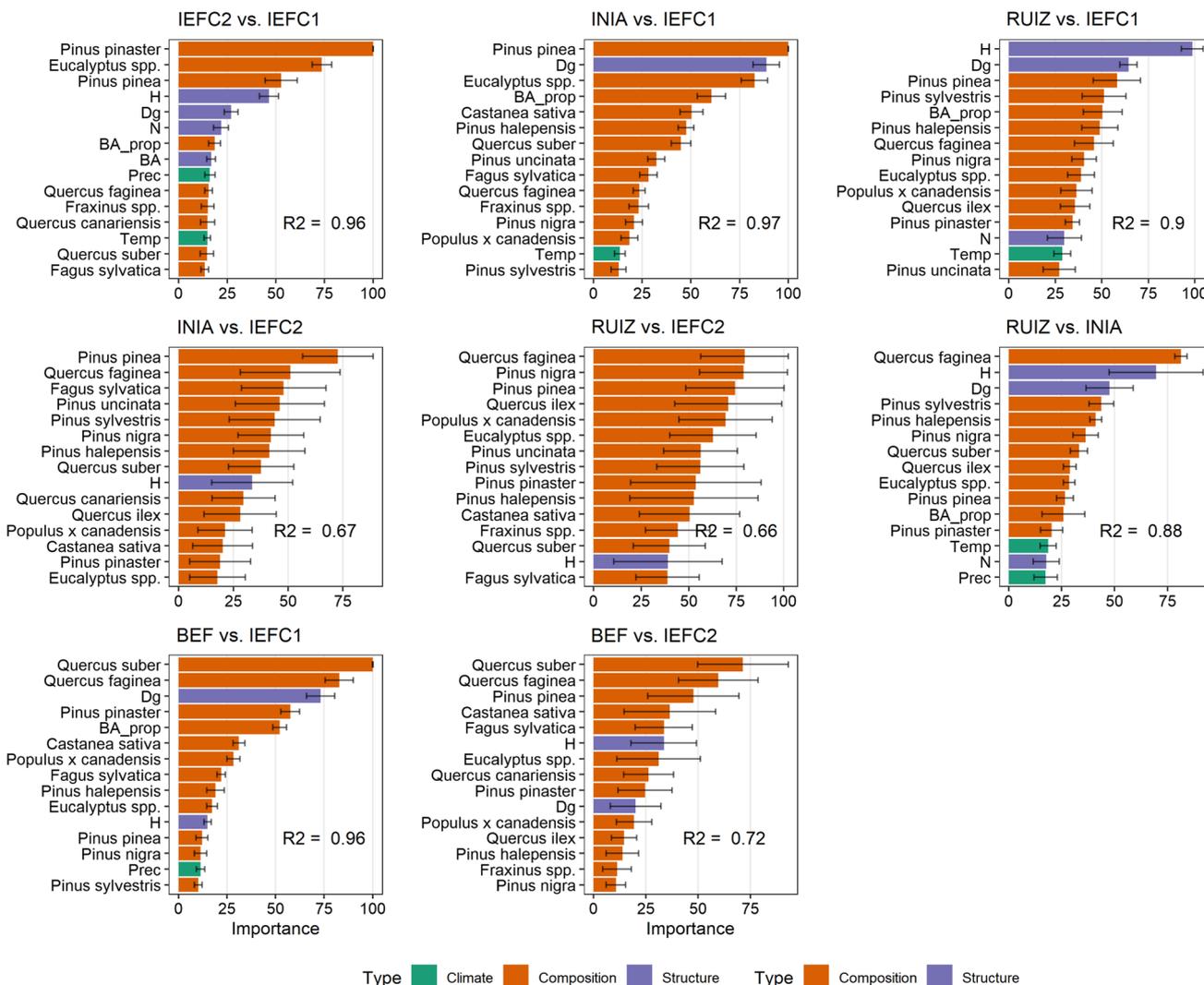
The discrepancies in biomass estimations were particularly high for diameters beyond the range used to calibrate the equations. It is fundamental that the databases used to obtain regression equations contain trees with large diameters, since they can account for a large proportion of stand aboveground biomass, particularly in mature forests (Brown, 2002). Extrapolating beyond the measured size range is obviously a practice to avoid, but for the majority of species analyzed here the size range used to calibrate the equations was wide enough to encompass the vast majority of the trees measured in the National Forest Inventory. However, special attention should be paid when estimating the biomass or carbon content in highly capitalized, fast-growing or mature forests, which contain trees with large diameters that can cause disparities to be more pronounced (Brown, 2002), but



**Fig. 4** Spatial representation of the pairwise differences in plot-level estimations of biomass for the plots of the Third National Forest Inventory (NFI3)

also on very dense stands, where the accumulation of many small discrepancies can also lead to sharp differences in total biomass.

Many allometric equations include tree height (H)—in addition to diameter (D)—as a covariate for tree biomass estimation, and some studies even include more complex



**Fig. 5** Main predictors of the pairwise differences between biomass estimates according to the variable importance as determined by the random forest models

formulations such as  $D^2 \cdot H$ . Although including height is considered to generally improve biomass estimations (Ruiz-Peinado et al. 2012), it is a variable more complicated to measure and the gains are generally low, so some authors argue it does not pay-off the effort (Annighöfer et al. 2016). In this study we did not observe any clear trend derived from the inclusion of height as a parameter in biomass estimates. The two sources that include this variable (IEFC2 and RUIZ) did not predict systematically higher or lower biomass values as compared to the other sources.

### Plot-level biomass estimations across different equation sources

Upscaling from tree-level to plot-level estimates often reveals patterns that cannot be appreciated when using individual tree results (Aguirre et al. 2021; Cosmo et al. 2016). As seen above (Fig. 3), the differences in biomass estimates for young trees are very small in absolute values, but still the total biomass results of young stands can

overcome that from older stands. This is due to the larger number of trees per unit area of younger stands (Pretzsch and Biber, 2005), which have an amplifying effect on the differences on forest stand level (Neumann et al. 2016). Moreover, it is known that the height–diameter relationship often varies with other variables such as tree density, site index and degree of competition, among others, which are usually not included when developing tree height allometric equations (Cysneiros et al. 2020). Therefore, the use of height as a predictor of biomass or carbon can become a source of uncertainty not visible using individual tree results (Neumann et al. 2016).

In our study, upscaling the biomass prediction to the plot level revealed interesting geographical patterns related to the presence of certain dominant species in the plots, but also to some structural features, as shown by the random forest models. Differences in plot-level estimates were particularly high in the northern fringe, mainly due to the presence of fast-growing species such as *Eucalyptus* or *Pinus pinaster*, which quickly reach large diameters. Due to their economic importance—these two fast-growing species represent 60% of the total volume harvested in Spain (Bravo et al. 2017)—discrepancies in biomass estimation for these species can lead to misestimating the revenues and the carbon balance of forest plantations (Paquette and Messier, 2010). Therefore, extreme caution should be paid when deciding which equation to use for each project. However, other not so productive areas also showed important differences in biomass estimations, notably *dehesas* in southwestern Spain. In this case, *dehesas* are sabana-like landscapes with only a few scattered trees, so total biomass values are lower than dense forests. Consequently, even small deviations in total biomass can lead to high percentage differences. However, their large extension may lead to important disparities in the estimation of their absorption/mitigation capacity.

### Good practices for biomass estimates

Although the aim of this work never was to identify the best available equation—since there is no “ground truth” to compare with—we did find some patterns in similarities/discrepancies that deem helpful to choose a suitable allometric equation. In general, equations fitted from the same dataset led to similar estimates, regardless of the method used. Hence, IEFC1 and IEFC2 reached similar results, and so did INIA and RUIZ—though generally yielding higher estimates. Previous research has found that part of the observed differences in biomass and carbon estimation across different sources can be attributed to the environmental growing conditions (Neumann et al. 2016). In our dataset, there are indeed species that have been parameterized in different places for each source, and this can undoubtedly foster

differences in the results. In this sense, if local equations are available for the target area, these should be given priority over “more generalist” equations (in the sense of those calibrated on data acquired over a large area), at least if the aim is to obtain the most accurate estimates. If generality is sought, it may be in turn preferable to use equations that cover larger territories. However, these are often fitted from datasets collected by different teams, and several authors have emphasized the need to unify methodologies to ensure that the results are truly comparable (Annighöfer et al. 2016; Neumann et al. 2016).

Biomass estimations using BEFs are often considered of poorer quality, only recommended when no estimates at the tree level are available (Castedo-Dorado et al. 2012; Jalkanen et al. 2005; Neumann et al. 2016; Petersson et al. 2012). The reasons are varied and include a double source of uncertainty—from volume equations and from biomass expansion factors themselves—and the fact that constant BEF are commonly used, as was the case in our study. BEFs are known to vary with tree age (Jalkanen et al. 2005), total stand volume (Aguirre et al. 2021; Soares and Tomé, 2004) and climate (Aguirre et al. 2021), among other factors. Nevertheless, we found that the BEF equations produced similar estimates than the allometric equations derived from the same dataset (IEFC1 and IEFC2; see Supplementary Material S3). In this regard, the use of BEF equations to provide nation-level estimations of carbon balance for international agencies seems like a valid and unbiased option, although further research would be required to confirm this extreme.

### Conclusions

We observed contrasting differences in plot-level biomass estimates depending on the source of allometric equations used. The differences were mostly related to tree size, main species and source of data used to calibrate the models. Our results show that special attention on the choice of allometric equations should be put for capitalized forests, where large trees can trigger large differences between estimations, especially if they hold trees bigger than the sample used to calibrate the equations. The main differences were found when using equations parameterized using different datasets, which constitutes a warning against the uncritical choice of equations to determine biomass or carbon values. Our results can provide insights into the use of allometric equations for biomass and carbon estimation and inform carbon accountability in restoration programs and in nation-wide monitoring initiatives for climate action.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s10342-022-01444-w>.

**Acknowledgements** This work was funded by the LIFE CLIMARK project (LIFE16 CCM/ES/000065) and by the Spanish State Research Agency through the UMBRACLIM project (PID2019-111781RB-I00). The authors would like to thank Lluís Coll for helpful discussions during the conceptualization of the study and Pere Gelabert for his help in map design and creation. Special thanks to Miquel de Caceres, Jordi Vayreda and Jordi Martínez-Vilalta for the creation of the Catalan Forestry Lab and the gathering and release of the IEF data.

**Funding** Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Aguirre A, del Río M, Ruiz-Peinado R, Condés S (2021) Stand-level biomass models for predicting C stock for the main Spanish pine species. *For Ecosyst* 8:29. <https://doi.org/10.1186/s40663-021-00308-w>
- Alberdi I, Cañellas I, Bombín RV (2017a) The Spanish national forest inventory: history, development, challenges and perspectives. *Pesqui Florest Bras* 37:361–368. <https://doi.org/10.4336/2017.pfb.37.91.1337>
- Alberdi I, Vallejo R, Álvarez-González JG, Condés S, González-Ferreiro E, Guerrero S, Hernández L, Martínez-Jauregui M, Montes F, Oliveira N, Pasalodos-Tato M, Robla E, Ruiz-González AD, Sánchez-González M, Sandoval V, San Miguel A, Sixto H, Cañellas I (2017b) The multi-objective Spanish national forest inventory. *For Syst* 26:e04S. <https://doi.org/10.5424/fs/2017262-10577>
- Annighöfer P, Ameztegui A, Ammer C, Balandier P, Bartsch N, Bolte A, Coll L, Collet C, Ewald J, Frischbier N, Gebereyesus T, Haase J, Hamm T, Hirschfelder B, Huth F, Kändler G, Kahl A, Kawaletz H, Kuehne C, Lacoine A, Lin N, Löf M, Malagoli P, Marquier A, Müller S, Promberger S, Provendier D, Röhle H, Sathornkitch J, Schall P, Scherer-Lorenzen M, Schröder J, Seele C, Weidig J, Wirth C, Wolf H, Wollmerstädt J, Mund M (2016) Species-specific and generic biomass equations for seedlings and saplings of European tree species. *Eur J for Res* 135:313–329. <https://doi.org/10.1007/s10342-016-0937-z>
- Bravo, F., Guijarro, M., Cámara, A., Díaz Balteiro, L., Fernández-Rebollo, P., Pajares, J.A., Pemán, J., Ruiz-Peinado, R., 2017. Informe de Situación de los Bosques y del Sector Forestal en España. Sociedad Española de Ciencias Forestales
- Breidenich C, Magraw D, Rowley A, Rubin JW (1998) The kyoto protocol to the united nations framework convention on climate change. *Am J Int Law* 92:315–331. <https://doi.org/10.2307/2998044>
- Brown S (2002) Measuring carbon in forests: current status and future challenges. *Environ Pollut* 116:363–372. [https://doi.org/10.1016/S0269-7491\(01\)00212-3](https://doi.org/10.1016/S0269-7491(01)00212-3)
- Castedo-Dorado F, Gómez-García E, Diéguez-Aranda U, Barrio-Anta M, Crecente-Campo F (2012) Aboveground stand-level biomass estimation: a comparison of two methods for major forest species in northwest Spain. *Ann for Sci* 69:735–746. <https://doi.org/10.1007/s13595-012-0191-6>
- Chave J, Riéra B, Dubois M-A (2001) Estimation of biomass in a neotropical forest of French Guiana: spatial and temporal variability. *J Trop Ecol* 17:79–96. <https://doi.org/10.1017/s0266467401001055>
- Condés S, McRoberts RE (2017) Updating national forest inventory estimates of growing stock volume using hybrid inference. *For Ecol Manag* 400:48–57. <https://doi.org/10.1016/j.foreco.2017.04.046>
- Cosmo LD, Gasparini P, Tabacchi G (2016) A national-scale, stand-level model to predict total above-ground tree biomass from growing stock volume. *For Ecol Manag* 361:269–276. <https://doi.org/10.1016/j.foreco.2015.11.008>
- Cysneiros VC, Pelissari AL, Gauí TD, Fiorentin LD, de Carvalho DC, Silveira Filho TB, Machado S, do A., (2020) Modeling of tree height–diameter relationships in the Atlantic Forest: effect of forest type on tree allometry. *Can J for Res* 50:1289–1298. <https://doi.org/10.1139/cjfr-2020-0060>
- Dirección General para la Biodiversidad (2007) Tercer inventario forestal nacional (1997–2007). Ministerio de Medio Ambiente, Madrid
- Dubayah R, Blair JB, Goetz S, Fatoyinbo L, Hansen M, Healey S, Hofton M, Hurtt G, Kellner J, Luthcke S, Armston J, Tang H, Duncanson L, Hancock S, Jantz P, Marselis S, Patterson PL, Qi W, Silva C (2020) The global ecosystem dynamics investigation: high-resolution laser ranging of the earth's forests and topography. *Sci Remote Sens* 1:100002. <https://doi.org/10.1016/j.srs.2020.100002>
- Eggleston S, Buendia L, Miwa K, Ngara T, Tanabe K (2006) IPCC guidelines for national greenhouse gas inventories. Institute for Global Environmental Strategies, Hayama
- Falster DS, Duursma RA, Ishihara MI, Barneche DR, FitzJohn RG, Vårhammar A, Aiba M, Ando M, Anten N, Aspinwall MJ, Baltzer JL, Baraloto C, Battaglia M, Battles JJ, Bond-Lamberty B, van Breugel M, Camac J, Claveau Y, Coll L, Dannoura M, Delagrèze S, Domec J-C, Fatemi F, Feng W, Gargaglione V, Goto Y, Hagihara A, Hall JS, Hamilton S, Harja D, Hiura T, Holdaway R, Hutley LS, Ichie T, Jokela EJ, Kantola A, Kelly JWG, Kenzo T, King D, Kloeppel BD, Kohyama T, Komiyama A, Laclau J-P, Lusk CH, Maguire DA, le Maire G, Mäkelä A, Markesteijn L, Marshall J, McCulloh K, Miyata I, Mokany K, Mori S, Myster RW, Nagano M, Naidu SL, Nouvellon Y, O'Grady AP, O'Hara KL, Ohtsuka T, Osada N, Osunkoya OO, Peri PL, Petritan AM, Poorter L, Portsmouth A, Potvin C, Ransijn J, Reid D, Ribeiro SC, Roberts SD, Rodríguez R, Saldaña-Acosta A, Santa-Regina I, Sasa K, Selaya NG, Sillett SC, Sterck F, Takagi K, Tange T, Tanouchi H, Tissue D, Ume-hara T, Utsugi H, Vadeboncoeur MA, Valladares F, Vanninen P, Wang JR, Wenk E, Williams R, de Aquino Ximenes F, Yamaba A, Yamada T, Yamakura T, Yanai RD, York RA (2015) BAAD: a biomass and allometry database for woody plants. *Ecology* 96:1445–1445. <https://doi.org/10.1890/14-1889.1>
- Ibáñez J, Vayreda J, Gracia C (2002) Metodología complementaria al inventario forestal nacional en Catalunya. El inventario forestal nacional: elemento clave para la gestión forestal sostenible. Fundación General de la Universidad de Valladolid, Valladolid, pp 67–77
- Jalkanen A, Mäkipää R, Ståhl G, Lehtonen A, Petersson H (2005) Estimation of the biomass stock of trees in Sweden: comparison of biomass equations and age-dependent biomass expansion factors. *Ann for Sci* 62:845–851. <https://doi.org/10.1051/forest:2005075>

- Jenkins JC, Chojnacky DC, Heath LS, Birdsey RA (2003) National scale biomass estimators for United States tree species. *For Sci* 49(12):35
- Kuhn M (2008) Building predictive models in R using the caret package. *J Stat Softw* 28:1–26. <https://doi.org/10.18637/jss.v028.i05>
- Liaw A, Wiener M (2002) Classification and Regression by random Forest. *R News* 2:18–22
- MITECO, 2019. Guía para la estimación de absorciones de dióxido de carbono. Ministerio de Transición Ecológica
- Montero G, Ruiz-Peinado R, Muñoz M (2005) Producción de biomasa y fijación de CO<sub>2</sub> por los bosques españoles, monografías INIA: serie forestal. Instituto Nacional de Investigación y Técnica Agraria y Alimentaria, Madrid
- Muukkonen P (2007) Generalized allometric volume and biomass equations for some tree species in Europe. *Eur J for Res* 126:157–166. <https://doi.org/10.1007/s10342-007-0168-4>
- Neumann M, Moreno A, Mues V, Härkönen S, Mura M, Bouriaud O, Lang M, Achten WMJ, Thivolle-Cazat A, Bronisz K, Merganič J, Decuyper M, Alberdi I, Astrup R, Mohren F, Hasenauer H (2016) Comparison of carbon estimation methods for European forests. *For Ecol Manag* 361:397–420. <https://doi.org/10.1016/j.foreco.2015.11.016>
- Paquette A, Messier C (2010) The role of plantations in managing the world's forests in the anthropocene. *Front Ecol Environ* 8:27–34. <https://doi.org/10.1890/080116>
- Petersson H, Holm S, Ståhl G, Alger D, Fridman J, Lehtonen A, Lundström A, Mäkipää R (2012) Individual tree biomass equations or biomass expansion factors for assessment of carbon stock changes in living biomass—a comparative study. *For Ecol Manag* 270:78–84. <https://doi.org/10.1016/j.foreco.2012.01.004>
- Pretzsch H, Biber P (2005) A re-evaluation of Reineke's rule and stand density index. *For Sci* 605:304–320
- Roelfsema M, van Soest HL, Harmsen M, van Vuuren DP, Bertram C, den Elzen M, Höhne N, Iacobuta G, Krey V, Kriegler E, Luderer G, Riahi K, Ueckerdt F, Després J, Drouet L, Emmerling J, Frank S, Fricko O, Gidden M, Humpenöder F, Huppmann D, Fujimori S, Fragkiadakis K, Gi K, Keramidas K, Köberle AC, Aleluia Reis L, Rochedo P, Schaeffer R, Oshiro K, Vrontisi Z, Chen W, Iyer GC, Edmonds J, Kannavou M, Jiang K, Mathur R, Safonov G, Vishwanathan SS (2020) Taking stock of national climate policies to evaluate implementation of the Paris agreement. *Nat Commun* 11:2096. <https://doi.org/10.1038/s41467-020-15414-6>
- Rojas-García F, De Jong BHH, Martínez-Zurimendi P, Paz-Pellat F (2015) Database of 478 allometric equations to estimate biomass for Mexican trees and forests. *Ann for Sci* 72:835–864. <https://doi.org/10.1007/s13595-015-0456-y>
- Ruiz-Peinado R, Del Rio M, Montero G (2011) New models for estimating the carbon sink capacity of Spanish softwood species. *For Syst* 20:176. <https://doi.org/10.5424/fs/2011201-11643>
- Ruiz-Peinado R, Montero G, Rio MD (2012) Biomass models to estimate carbon stocks for hardwood tree species. *For Syst* 21:42. <https://doi.org/10.5424/fs/2112211-02193>
- Ruiz-Peinado R, Bravo-Oviedo A, López-Senespleda E, Bravo F, Río MD (2017) Forest management and carbon sequestration in the mediterranean region: a review. *For Syst* 26:04. <https://doi.org/10.5424/fs/2017262-11205>
- Schepaschenko D, Chave J, Phillips OL, Lewis SL, Davies SJ, Réjou-Méchain M, Sist P, Scipal K, Perger C, Herault B, Labrière N, Hofhansl F, Affum-Baffoe K, Aleinikov A, Alonso A, Amani C, Araujo-Murakami A, Armston J, Arroyo L, Ascarrunz N, Azevedo C, Baker T, Bařaly R, Bedeau C, Berry N, Bilous AM, Bilous SY, Bissengou P, Blanc L, Bobkova KS, Braslavskaya T, Brienen R, Burslem DFRP, Condit R, Cuni-Sanchez A, Danilina D, del Torres D, C., Derroire, G., Descroix, L., Sotta, E.D., d'Oliveira, M.V.N., Dresel, C., Erwin, T., Evdokimenko, M.D., Falck, J., Feldpausch, T.R., Foli, E.G., Foster, R., Fritz, S., Garcia-Abril, A.D., Gornov, A., Gornova, M., Gothard-Bassébé, E., Gourlet-Fleury, S., Guedes, M., Hamer, K.C., Susanty, F.H., Higuchi, N., Coronado, E.N.H., Hubau, W., Hubbell, S., Ilsted, U., Ivanov, V.V., Kanashiro, M., Karlsson, A., Karminov, V.N., Killeen, T., Koffi, J.-C.K., Konovalova, M., Kraxner, F., Kreyza, J., Krisnawati, H., Krivobokov, L.V., Kuznetsov, M.A., Lakyda, I., Lakyda, P.I., Licona, J.C., Lucas, R.M., Lukina, N., Lussetti, D., Malhi, Y., Manzanera, J.A., Marimon, B., Junior, B.H.M., Martinez, R.V., Martynenko, O.V., Matsala, M., Matyashuk, R.K., Mazzei, L., Memiaghe, H., Mendoza, C., Mendoza, A.M., Morozziuk, O.V., Mukhortova, L., Musa, S., Nazimova, D.I., Okuda, T., Oliveira, L.C., Ontikov, P.V., Osipov, A.F., Pietsch, S., Playfair, M., Poulsen, J., Radchenko, V.G., Rodney, K., Rozak, A.H., Ruschel, A., Rutishauser, E., See, L., Shchepashchenko, M., Shevchenko, N., Shvidenko, A., Silveira, M., Singh, J., Sonké, B., Souza, C., Stereńczak, K., Stonozhenko, L., Sullivan, M.J.P., Szatniewska, J., Taedoumg, H., Steege, H. ter, Tikhonova, E., Toledo, M., Trelova, O.V., Valbuena, R., Gamarra, L.V., Vasiliev, S., Vedrova, E.F., Verhovets, S.V., Vidal, E., Vladimirova, N.A., Vleminckx, J., Vos, V.A., Vozmitel, F.K., Wanek, W., West, T.A.P., Woell, H., Woods, J.T., Wortel, V., Yamada, T., Hajar, Z.S.N., Zo-Bi, I.C., (2019) The forest observation system, building a global reference dataset for remote sensing of forest biomass. *Sci Data* 6:1–11. <https://doi.org/10.1038/s41597-019-0196-1>
- Scipal, K., Arcioni, M., Chave, J., Dall, J., Fois, F., LeToan, T., Lin, C.-C., Papathanassiou, K., Quegan, S., Rocca, F., Saatchi, S., Shugart, H., Ulander, L., Williams, M., 2010 The BIOMASS mission—An ESA Earth explorer candidate to measure the BIOMASS of the Earth's forests. 2010 IEEE international geoscience and remote sensing symposium 52–55. <https://doi.org/10.1109/igars.2010.5648979>
- Soares, P., Tomé, M., 2004 Analysis of the effectiveness of biomass expansion factors to estimate stand biomass. In: Proceedings of the international conference on modeling forest production
- Ter-Mikaelian MT, Korzukhin MD (1997) Biomass equations for sixty-five North American tree species. *For Ecol Manag* 97:1–24. [https://doi.org/10.1016/S0378-1127\(97\)00019-4](https://doi.org/10.1016/S0378-1127(97)00019-4)
- Valbuena-Carabaña M, de Heredia UL, Fuentes-Utrilla P, González-Doncel I, Gil L (2010) Historical and recent changes in the Spanish forests: a socio-economic process. *Rev Palaeobot Palynol* 162:492–506. <https://doi.org/10.1016/j.revpalbo.2009.11.003>
- Vayreda J, Ibáñez J, Gràcia C (2005) El Inventario Ecológico y Forestal de Catalunya y su consultamediante la aplicación MiraBosc “online.” *Cuad Soc Esp Cienc for* 19:217–227
- Zianis D, Mencuccini M (2004) On simplifying allometric analyses of forest biomass. *For Ecol Manag* 187:311–332. <https://doi.org/10.1016/j.foreco.2003.07.007>
- Zianis D, Muukkonen P, Mäkipää R, Mencuccini M (2005) Biomass and stem volume equations for tree species in Europe. *Silva Fenn Monogr* 4:63

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.