Prediction of aboveground biomass and carbon stock of *Balanites aegyptiaca*, a multipurpose species in Burkina Faso

Sambo Ouédraogo a,*, Oumarou Ouédraogo a, Kangbéni Dimobe a,b,c, Adjima Thiombiano a, Joseph I. Boussima a

Laboratory of Plant Biology and Ecology, University Joseph Ki-Zerbo, 03 BP 7021 Ouagadougou 03, Burkina Faso
West African Science Service Center on Climate Change and Adapted Land Use (WASCAL), Competence Center, Avenue Muamar Khadba, Ouagadougou BP, 9507, Burkina Faso
Institut des Sciences de l’Environnement et du Développement Rural, Université de Dédougou, BP 176 Dédougou, Burkina Faso

**ARTICLE INFO**

Keywords:
- Biomass allocation
- Predictive equations
- Carbon content
- *Balanites aegyptiaca*
- Land use and landcover
- Agricultural science
- Environmental science

**ABSTRACT**

*Balanites aegyptiaca* (L.) Delile is native to semi-arid regions in Africa where it is a well-known and conspicuous component of savannas. The species is highly preferred by local people because of its high socio-economic, cultural and ecological values. However, the species faces multiple environmental challenges such as desertification and human pressure. This study aimed to develop allometric models to predict aboveground biomass (AGB) of *B. aegyptiaca* in two climatic zones in Burkina Faso. Overall, thirty trees were sampled using destructive method in six study stands along two climatic zones. We assessed the biomass allocation to the different components of trees by computing its fraction. Furthermore, allometric models based on diameter at breast height (dbh) and basal diameter at 20 cm height (D20) were fitted separately as well as combined with crown diameter (CD) and/or tree total height (Ht). For each biomass component, non-linear allometric models were fitted. Branch biomass accounted for 64% of the AGB in the two climatic zones and increased with dbh. No significant difference in carbon content was found. However, biomass allotment (except leaves) varied across climatic zones. Although both dbh and D20 are typically used as independent variables for predicting AGB, the inclusion of the height in the equations did not significantly improve the statistical fits for *B. aegyptiaca*. However, adding CD to dbh improved significantly the equations only in the Sudano-Sahelian zone.

The established allometric models can provide reliable and accurate estimation of individual tree biomass of the species in areas of similar conditions and may contribute to relevant ecological and economical biomass inventories.

1. **Introduction**

The issue of climate change is a major concern at the global level and its effects affect all areas of life (Mbow et al., 2013; Dimobe et al., 2018). Protected and non-protected forest ecosystems, which play a key role in mitigating the effects of global warming, are intensively threatened by global and local population growth, which causes increased needs (in terms of energy) and human occupation that causes irreversible damage to biodiversity. Strategies to mitigate the negative effects of global warming has become a global issue with mainly two strategies: the reduction of Greenhouse Gases (GHGs) release and the sequestration and storage of released CO₂ (Hairiah et al., 2011; Houghton et al., 2012). Therefore there is an increasing interest to convince policy makers of the need for tools to assess plants ability to capture and store the atmospheric carbon (C). There is also a necessity to set reliable, accurate and economical methods for estimating the biomass of trees and shrubs (Djomo et al., 2010, 2016; Laminou Manzo et al., 2015). These tools would help to determine the geographical distribution of C stocks and to understand changes in C stocks in relation to other parameters such as land use and climatic zones.

Estimates of stored C rely mainly on biomass assessments. There are tools such as pantropical allometric equations proposed by Chave et al. (2005) and Chave et al. (2014) for estimating the C of the Sudanian savannas for global Tropical zones, but unfortunately these tools give varied results because of the variability of the methods used to evaluate the aboveground biomass (Oelbermann et al., 2004; Picard et al., 2015;

---

* Corresponding author.
E-mail address: samboenma1@gmail.com (S. Ouédraogo).

10.1016/j.heliyon.2020.e04581
Received 23 January 2020; Received in revised form 18 May 2020; Accepted 28 July 2020
2405-8440/© 2020 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
Chabi et al., 2016; Yuen et al., 2016). These tools did not take into account the climatic conditions of most African tropical savannas. The use of pre-existing equations, rather than those determined on site, is an important source of uncertainty in assessments of biomass and carbon changes (Yuen et al., 2016). The most accurate method for estimating forest biomass is destructive sampling, which consists of harvesting all trees in a given area and measuring the mass of wood and foliage components (Basuki et al., 2009). It is however laborious and expensive. This leads to a general lack of allometric adapted equations to estimate biomass in West African savanna ecosystems (Dimobe et al., 2018). Some studies were carried out in West Africa in arid environments, but these works are unfortunately limited (Sawadogo et al., 2010; Bayen et al., 2016; Dimobe et al., 2018; Balima et al., 2019) in Burkina Faso, Chabi et al. (2016) in Benin, Tredennick et al. (2013) in Mali and Mbou et al. (2013) in Senegal.

Tree cover and tree structure are greatly influenced by the variability of anthropogenic disturbances. That limits the use of general models fitted with data from South American savanna ecosystems and from dry forests in India, which are different from local arid forests in Africa (Yuen et al., 2016; Dimobe et al., 2018). There is an evidence that the application of these equations to targeted species might cause large systematic differences with the local environment data (Mensah et al., 2016, 2018; Dimobe et al., 2018). Currently, the role of each species in climate change mitigation and C sequestration is unknown and is either underestimated or overestimated because it depends heavily on the allometric models chosen for biomass and C estimates (Picard et al., 2012; Chave et al., 2014; Yuen et al., 2016; Dimobe et al., 2018).

Among allometric models recently approved for reducing the uncertainties on biomass prediction are the additive models (Dong et al., 2016; Dimobe et al., 2018). The reliability of these models is fitting accurately total above ground biomass (TAGB) that equals to the sum of tree components biomass (Dong et al., 2016). What matters is therefore to use the additivity method in fitting models to allow an TAGB prediction with less bias as possible. Several methods exist in this area, but seemingly unrelated regression (SUR) is widely use in recent studies in Burkina Faso (Dimobe et al., 2018; Ganamé et al., 2020), because it creates compatibility between the prediction of the tree's components biomass and TAGB (Parresol et al., 2001).

*Balanites aegyptiaca* (L.) Delile is a species adapted to Sahelian climate and is encountered in Sahel-Saharan and Middle East regions of Africa (Sagna et al., 2014). With broad ecological amplitude, it is very resistant to drought and relatively indifferent to the type of soil. In Burkina Faso, the species is abundant and well distributed in both communal and protected areas of Sahelian and Sudano-Sahelian zones (Thiombiano et al., 2012). Indeed, the species is widely used in the rural and in urban areas in human and animal diet, medicinal and cosmetic uses. The species has rapid a growth, colonizing various ecosystems. These assets make it a key species in restoration programs and also for C sequestration. Hence the estimate of its C stock can support REDD+ (Reducing Emissions from Deforestation and forest Degradation) initiatives or the impact of conservation program actions in the Sahel.

As contribution, this study aimed to evaluate biomass allocation patterns for aboveground components, to develop allometric equations for *B. aegyptiaca*, a dominant tree species in tropical semi-arid zones, and to assess its C content and C amount in four land use land covers, namely the agroforestry parklands, the wooded grasslands, the shrubby savannas and the shrubby grasslands.

**Figure 1.** Location of the study area (Burkina Faso, West Africa).
2. Materials and methods

2.1. Study zone

The study was conducted in six near-natural stands of *B. aegyptiaca* in Burkina Faso, West Africa (three stands in the Sahelian climatic zone and three in the Sudano-Sahelian zone- Figure 1). The main vegetation types in the Sahelian zone are shrub savannas and grass savannas while in the Sudano-Sahelian there are shrub savannas, tree savannas, woodlands and grass savannas with a grass layer dominated by annual herbaceous species in the Sudano-Sahelian zone. These vegetation types are adopted from the classification commonly used for West Africa (FAO, 1981 and 2002). Agroforestry parklands include cultivated fields and young fellows are widely spread in the two zones, and many fruit species (*Vitellaria paradoxa*, *Parkia biglobosa*, *Tamarindus indica*, *Balanites aegyptiaca* etc) are identified and spared by farmers.

The rainfall pattern at both stands is uni-modal with different rainy season lengths and annual rainfall amounts. In the Sahelian zone, the length of the rainy season is 2–3 months with 487.37 ± 110.07 mm of average annual rainfall and an average annual temperature of 42.3 °C during the last 30 year period (1987–2017). In the Sudano-Sahelian zone which is the largest climatic zone in Burkina Faso, the rainy season lasts for 4–5 months with an annual rainfall of 839.58 ± 151.97 mm and an average annual temperature of 30.7 °C for 1987–2017 period.

2.2. Forest inventory and biomass data

Forest inventories were carried in September 2018 in 65 sample plots, established in six shrubby savannas in the two climatic zones. The objective of the forest inventory was to select sample individuals of the species based on stem diameter classes. The plot size was 50 m × 20 m (1000 m²) with 31 plots in Sahelian zone and 34 plots in Sudano-Sahelian zone. In each stand, a minimum distance of 100 m separated consecutive plots. The maximum distance separating plots in this study was 500 m.

Based on the forest inventory, a total of thirty (30) individual trees of *B. aegyptiaca* in the six stands were sampled. Stands were selected based on the trees abundance, avoiding the destructive sampling impact on species density on the basis of the species distribution (Figure 2). Sudano-Sahelian and Sahelian climatic zones were considered because they are the suitable zones for *B. aegyptiaca* growth in Burkina Faso. Within each plot, tree parameters such as diameter at breast height (dbh), total tree height (H), crown diameter (CD), and basal diameter (diameter at 20 cm above the ground level, D$_{20}$) were measured on the *B. aegyptiaca* individuals.

For biomass sampling, a total of 30 individual trees were grouped into four different classes of dbh size (5–10; 10–15; 15–20 and > 20 cm). The sampled trees’ number to harvest was proportional to their abundance in a particular diameter class (Dimobe et al., 2018). Hence, the number of individual trees selected per diameter class are as follow: 9 individuals for class 5–10 cm 6 individuals for class 10–15 cm, 7 individuals for class 15–20 cm and 8 individuals having dbh greater than 20 cm.

The CD of each tree was measured in two perpendicular directions (east-west and north-south) to avoid bias, and the average value was used (Dimobe et al., 2018). Total tree height was measured from the base of the trunk to the apex of the tree using a tape of 50 m. For trees forking below 1.30 m, the diameters of all branches were measured, and the root mean square diameter was calculated as follows:

$$D = \sqrt{\sum d_i^2}$$

where “$d_i$” is the of dbh of individual stems of the tree.

The trees selected for sampling were cut at the base (20 cm above the ground level) using a chainsaw and each tree was divided into stem, branches and foliage. The fresh weights of the tree components (stem, branches and foliage) were weighed separately in the field using a 100 kg balance for the wood and a 5 kg spring balance for the foliage. Samples of foliage, stems and branches were taken and weighed (fresh weight) using

![Figure 2. Balanites aegyptiaca distribution map (source: National secondary inventory).](image-url)
the 5 kg electronic balance (0.5 g of accuracy). The samples of branch and stem were taken as a pie or a disc of 5 cm thickness depending on the size (stem diameter) of the trunk or branch (Dimobe et al., 2018). Discs and leaf samples were carried to the laboratory and oven-dried to a constant weight at 105 °C for branch and stem samples and at 75 °C for leaf samples. The dry weight of the samples was recorded immediately after removal from the oven to avoid weighing errors. The sample dry weight was divided by corresponding green weight to obtain the dry-to–green weight ratio (DG-ratio) for each tree component (stem, branches and foliage) according to Dimobe et al. (2018) approach. The dry weight of each component was obtained by multiplying the DG-ratio by the green weight of the respective tree component. Total AGB of the tree was computed as the sum of all component weights in kilograms. The samples’ descriptive values are in Table 1.

2.3. Determination of the carbon content

We used the ash method to determine the proportion of organic C of each individual tree of B. aegyptiaca (Dimobe et al., 2018). To determine the C content of each component, composite samples were constituted for each of them by crushing stem, branches and leaves samples separately, after oven drying, using mortar and pestle (Bayen et al., 2016). Thereafter, five samples of 2 g (g) dry powder of each component were put in a porcelain crucible and incinerated in an oven at 550 °C for 2 h, until complete calcination. This analysis was conducted in the Laboratory of Plants and Soils at the University Joseph Ki-ZERBO, in Burkina Faso. The samples were then slowly cooled in a desiccator and weighed. After cooling, the ash-containing crucibles were weighed and the percentage of organic C is calculated as indicates in this formula (Bayen et al., 2016).

\[ \text{Ash} = \frac{(W3-W1)}{(W2-W1)} \times 100 \]  
\[ \text{Carbon} = \frac{\text{Biomasse} \times \text{Carbon} \%}{100} \]  
\[ \text{Carbon} = 0.58 \times \text{Carbon} \% \]

where 0.58 is the content of carbon in the organic matter, WI is the weight of the ashes, W2 is the weight of the oven-dried grind samples + ashes, and W3 is the weight of the ash with crucibles.

The average of C content was calculated for each tree component based on the number of replicates and applied for each individual tree. The average C content of the three components (leaf, branch and stem) was considered as the C content at tree level. We computed the mean C content for each climatic zone based on the average C content stored in individual trees of the species.

2.4. Data analysis

A generalized linear model (GLM) was performed to assess the effects of climatic zones and tree size C content. In order to determine whether the effects of tree size would vary by climatic zone, we tested the interaction effects between tree size and climatic zones. Analysis of variance (ANOVA) was applied to assess if tree components biomasses are different between the two climatic zones, using multiple comparison test. Additive allometric equations have been developed for each climatic zone to predict tree biomass component (leaf, branch and stem) and the total aboveground biomass (TAGB). A scatterplot between predictive and predicted variables was plotted to have an overview of data and to identify outliers and the nature of correlation (Figure 3A&B). To identify the accurate nature of the correlation between components biomass and predictors (dbh, basal diameter, height, crown diameter), we explored their relationship using linear (\( Y = a + bx \)), exponential (\( Y = ae^{bx} \)) and power (\( Y = ax^b \)) equations. The power function was found as the accurate equation and was then selected and presented in its linear form as:

\[ \ln y = \ln a + \beta \ln x + \epsilon \]

\[ \ln y = \ln a + \beta x + \gamma \ln x + \epsilon \] where \( y \) is the dependent variable (biomass), \( x \) and \( z \) are the predictors; \( a \) and \( \beta \) are regression coefficients, and \( \epsilon \) the random error. We used this form because the biomass data usually exhibit non constant residuals in variance in model residuals (Parresol et al., 2001). So, variance can be stabilized either by providing a weight function or by using the log-transformation.

Seemingly unrelated regression (SUR), a generalization of a linear regression model that consists of several regression models, each having its own dependent variable and potentially different sets of exogenous explanatory variables (Zellner, 1962) were used to fit models and to apply additivity propriety between AGB and components’ biomass. Each model is a valid on its own and can be estimated separately.

We computed the equations for each component and TAGB based on seven non-linear equations with dbh, height (Ht), D20 and CD as independent variables. Dbh and D20 were used as the lone predictors variable in equation (4), equation (5) and Eq. (7). In Eq. (6), dbh and height were used compound variables; and in Eqs. (8), (9), and (10), dbh or D20 was fitted with height or CD as additional predictor variables (Dimobe et al., 2018). Equations are presented as follows:

\[ \ln y = \ln a + \beta \ln (dbh) \]  
\[ \ln y = \ln a + \beta \ln (D20) \]  
\[ \ln y = \ln a + \beta \ln (Dh^2) \ln (Ht) \]  
\[ \ln y = \ln a + \beta \ln (dbh^2) \ln (Ht) + \gamma \ln (CD) \]  
\[ \ln y = \ln a + \beta \ln (dbh^2) \ln (CD) + \gamma \ln (Ht) \]  
\[ \ln y = \ln a + \beta \ln (Dh^{2.2}) + \gamma \ln (CD) \]  
\[ \ln y = \ln a + \beta \ln (Dh^{2.2}Ht) + \gamma \ln (CD) \]

where \( Y \) is the component biomass, and \( a \) and \( \beta \) are the regression coefficients.

To correct the bias introduced due to logarithmic transformation of response variable, the correction factors (CF) were used in each equation and was calculated as follows:

\[ \text{CF} = \exp \left( \frac{\text{RSE}^2}{2} \right) \]

Additive allometric equations are widely used to reduce uncertainties due to additive property between biomass components (Dimobe et al., 2018). TAGB models were fitted by using additive models from the best component models as follow:

\[ \text{TABG} = \alpha_1 \text{X}_{\text{L}}(\text{CF}_1) + \alpha_2 \text{X}_{\text{B}}(\text{CF}_2) + \alpha_3 \text{X}_{\text{S}}(\text{CF}_3) \]

\[ \text{TABG} = \alpha_3 \text{X}_{\text{L}}(\text{CF}_1) + \alpha_4 \text{X}_{\text{B}}(\text{CF}_2) + \alpha_5 \text{X}_{\text{S}}(\text{CF}_3) \]

\[ \alpha_3 \text{X}_{\text{L}}(\text{CF}_1) = \text{Leaf biomass}, \alpha_4 \text{X}_{\text{B}}(\text{CF}_2) = \text{branch biomass} \]

\[ \alpha_5 \text{X}_{\text{S}}(\text{CF}_3) = \text{stem biomass} \]
The best models were selected based on the adjusted coefficient of determination (Adj.R²), the standard residual error (RSE), the Akaike information criterion (AIC), the global significance of the model (p-value), the root mean square error (RMSE) and the average systematic bias (Vahedi et al., 2014; Bayen et al., 2016; Dimobe et al., 2018). Hence, models with higher R², smaller values of AIC, RSE, RMSE and significant coefficients with p < 0.001 were selected as the best fitted model to predict the biomass of the components and the TAGB of the species (Dimobe et al., 2018). In order to validate selected models, Student t-test was also performed to check difference between predicted and observed biomass.

The second national inventory (IFN2) of 2012 has produced a national land use, land cover (LULC) map for Burkina Faso (BDOT) using a modified LULC type’s classification scheme of FAO (2002). We used the TAGB additive equations to estimate the amount of C stored in these main LULC types from the mean stem density of the species in each land cover type, the stem density potential was estimated as follows:

\[
LULC_{trees} = Dm \times LULC_{area} \quad (13)
\]

\[
Clim\_Zone\_trees = \sum LULC\_trees \quad (14)
\]

\[
Dm = \frac{1}{N} \times \sum (ni \times 10000) / Sp \quad (15)
\]

where LULC_trees is the number of trees in each land use and land cover type, Clim_Zone_trees is the number of trees in each climatic zone; Dm is the density (trees/ha) of trees in a given land cover type; N = total number of plots in the i-th LULC type; ni is the number of B. aegyptiaca individual trees in each plot and Sp is the area of the plot. LULC area is the area of each land cover type at climate zone level, obtained from a national land cover dataset using ArcGis 1.2 software.

The analyses were carried out in R (R 2.15.3, Rcmdr) and Minitab 16 version 1.0.01.

3. Results

3.1. Above ground biomass and carbon allocated in tree components

The analysis on biomass allocation and C content of tree components showed that branches had the highest biomass fraction (64.19 ± 9.95% in the Sahelian zone and 63.88 ± 9.83% in the Sudano-Sahelian zone), followed by stems (33.78 ± 10.09% in the Sahelian zone and 34.76 ± 9.83% in the Sudano-Sahelian zone). Foliage had the lower allocation in the two zones with 2.02 ± 1.50% and 1.36 ± 0.60% in the Sahelian zone and Sudano-Sahelian zone, respectively (Figure 4A). The results showed no significant difference when comparing tree foliage component biomass between the climatic zones (P > 0.05), but significant difference was found between the other component and the TAGB (P < 0.05).

The C content showed a same trend in both climatic zones with a lower content in foliage (Figure 4B). In the Sahelian zone, the C content of foliage was 50.70 ± 1.02% followed by branch (57.01 ± 0.18%) and stem (57.03 ± 0.27%). In the Sudano-Sahelian zone, the C content in foliage was slightly superior (52.23 ± 0.93%) to the one in Sahelian zone while stem and branch C contents (55.93 ± 0.26% and 56.74 ± 0.72% respectively) were found slightly low. No significant effect of climate and tree size on C content was found in tree components while using GLM (Table 2). Analysis on TAGB, showed that it ranges from 11.5 kg to 227.6 kg in the Sahelian zone and from 10.80 to 589.5 kg in the Sudano-Sahelian zone (Table 3). When comparing, TABG as well as the C content mean values across climatic zones, the results showed no significant effect of climate on both biomass and C content in tree components (P > 0.05). In the same line, C storage is similar in the two zones, and C content is the same trough dbh range in the two zones.

3.2. AGB allometric models

In the Sahelian zone, the fitted allometric models to predict biomass components showed AIC values that ranged between 10.57 and 25.03; RSE between 0.314 and 0.554 and RMSE between 0.251 and 4.61 (Table 4). For the Sudano-Sahelian zone, AIC values range between 13.18 and 34.56, RSE between 0.311 and 0.584, and RMSE ranges between 0.290 and 0.813 (Table 5).

In the Sahelian zone, the higher Adj. R² values were recorded for the stem biomass (ranged from 0.75–0.86) while leaf biomass had the lowest Adj. R² (0.66–0.93), and the equations for stem biomass prediction presented the lowest Adj. R² values ranged between 0.70 and 0.72.

The best equation for leaf biomass was fitted using dbh in the Sahelian zone Mod. 1, Table 4) while it showed a better prediction with dbh and crown area (dbh×CD) used as compound variable (Mod. 14), for the Sudano-Sahelian. The basal diameter (Dbh) provided a better estimation of both branch and stem biomasses in the sahelian zone (Mod. 5 and Mod. 9).

In the Sudano-Sahelian zone, branch biomass was better predicted using dbh and crown diameter (dbh×CD) as compound predictor (Mod. 18) while dbh alone was the best predictor for stem biomass (Mod. 21).
While predicting the TAGB, the additive TAGB model was obtained by combining the best allometric equation of each biomass component with 86% of explained variance. The model developed for Sahelian zone was:

\[
\text{TAGB} = 1.106e^{-3.77(dbh)^{1.41}} + 1.122e^{-3.29(D20)^{2.42}} + 1.163e^{-3.74(D20)^{2.12}}
\]

whereas in the Sudano-Sahelian zone it was

\[
\text{TAGB} = 1.070e^{-3.86(dbh^2CD)^{0.54}} + 1.049e^{-2.30(dbh^2CD)^{2.02}} + 1.041e^{-2.37(dbh)^{2.02}}
\]

with Adj. \( R^2 = 0.92 \) (Table 6).

The selected equations in this study are presented in Table 6. Leaf biomass was better predicted using dbh in the Sahelian zone and dbh combined with crown diameter (dbh^2CD) in the Sudano-Sahelian zone. Basal diameter was a good predictor for branch and stem biomasses in the Sahelian zone while in the Sudano-Sahelian zone, dbh alone was the good predictor for stem biomass and dbh^2CD for the branch biomass.

The comparison between the observed and predicted values of each component biomass in both climatic zones showed no significant deviation. No significant deviation was observed between the observed and predicted biomass values for leaf (\( t = 5.569e-17, df = 12, p\text{-value} = 0.99 \)), branch (\( t = -3.686e-17, df = 16, p\text{-value} = 0.992 \) for leaf; \( t = 8.7382e-16, df = 16, p\text{-value} = 0.892 \) for branch \( t = 0, df = 16, p\text{-value} = \)) in the Sahelian zone. In the Sudano-Sahelian zone, these statistics were: \( t = -5.290e-16, df = 12, p\text{-value} = 1 \), and stem (\( t = 4.101e-16, df = 12, p\text{-value} = 0.96 \) in the Sahelian zone. In the Sudano-Sahelian zone, these statistics were: \( t = -3.686e-17, df = 16, p\text{-value} = 0.992 \) for leaf; \( t = 8.7382e-16, df = 16, p\text{-value} = 0.892 \) for branch \( t = 0, df = 16, p\text{-value} = \))

**Figure 4.** Biomass allocation (A) and C content (B) in tree components of *Balanites aegyptiaca* in Sahelian zone and Sudano-Sahelian zone.

**Table 2.** Results of the GLM presenting the effects of tree size and climate on C content of *B. aegyptiaca.*

| Estimate       | Standard error | z-value | Pr (>|z|) |
|----------------|----------------|---------|----------|
| Intercept      | 1.304          | 0.012   | 13.234   | 0.023    |
| Climatic zone  | 0.452          | 0.008   | 3.015    | 0.351    |
| Tree size      | 1.120          | 0.031   | 3.545    | 0.245    |
| Tree size: climatic zones | 0.956 | 0.016 | 2.145 | 0.183 |
0.964 for stem. Using the equations to predict TAGB, we found that the AGB did not vary significantly between the predicted and observed values in the Sahelian zone ($t = 0.001 \text{ df} = 12$, $p$-value $= 0.998$) as well as in the Sudano-Sahelian zone ($t = -3.011 \text{e-16}, \text{ df} = 16$, $p$-value $= 0.923$). These statistics showed the good fit and reliability of the selected allometric equations. Graphically, the plots of observed versus predicted values of biomass showed a linear trend ($y = x$), and confirm the good accuracy log–transformation equations used in both climatic zones (Figure 5A&B).

### 3.3. Carbon potential in different land covers for each climatic zone

In each climatic zone, the potential of C for LULC types estimated showed valuable results. Mean dbh did not varied significantly across LULC types (Table 7). The results of the average biomasses and the corresponding quantity of C per LULC types and climatic zone were estimated and reported in Table 8. There were a significant difference between the quantities of stored C in different climate zones ($p < 0.05$). The Sudano-Sahelian zone which is the largest in terms of area had a high potential of *B. aegyptiaca* trees and therefore was identified as the climatic zone that stores more C by this species. Considering its populations in the LULC types of the two climatic zones, the total amount of C stored by *B. aegyptiaca* was estimated at 3 547 980.21 tons (Table 8). In this study, the LULC types with high C potential were wooded grassland and the shrubby grassland, respectively in the Sahelian zone and the Sudano-Sahelian zone (Table 8). The Sahelian zone has an estimated C potential storage of 3.57 tons/ha while the Sudano-Sahelian zone stores 1.82 tons/ha following the dynamics of estimated stem density potentials. Carbon storage amount trend is summarized across LULC types based on their respective areas in Table 8. Sudano-Sahelian zone LULCs stored more C compared to those in Sahelian zone (Table 8).

### 4. Discussion

#### 4.1. Balanites aegyptiaca TAGB and C content allotment within tree across two climatic zones

The results in this study showed that woody component biomass (stem and branch) had the largest contribution to aboveground biomass. Branch contributed for approximately 64% of TAGB in the two climatic zones.
zones, whereas leaf represented the smallest portion (<3%) in both zone. These results are consistent with several published studies in the tropical African savanna species as Vitellaria paradoxa, Combretum glutinosum, Terminalia laxiflora and Afzelia africana (Mensah et al., 2016; Balima et al., 2019; Dimobe et al., 2019).

Previous studies on biomass distribution within tree, reported that the share of components on the aboveground biomass varies within tree and across species (Dimobe et al., 2018, 2019; Meng et al., 2019). However, other studies found that stem had greatest biomass than branches and foliage (Sawadogo et al., 2010; Henry et al., 2011; Luo et al., 2014; Xue et al., 2016). These differences might be linked to the species specific responses to their site conditions. Scientist considered that allocating more resources to stem and branch when the size increases is to promote height and crown growth to increase the competitiveness for light relative to neighbor trees (Dimobe et al., 2018; Meng et al., 2019). Leaf biomass proportion decreased with increasing dbh in both climatic zones suggesting that the species has the same specific ecological and physiological responses towards environmental conditions (Balima et al., 2019).

When considered at the species level, this study showed that the average C content of the different parts varies significantly between the foliage and the woody parts (branches and the trunk). These results are consistent with those of Wani and Qaisar (2014), and Bayen et al. (2016) who also found a significant difference of the C content in the different parts of Desodara cedrus, Prunus floribunda, Ulmus wallichiana and Jatropha curcas, in the tropical zone. Also, works conducted by Negi and Durgapal (2003) and Pala et al. (2013) had explained earlier this difference indicating that the C content in the different parts of a species depends on the ash content in these parts, the ash being related to the quantity of the structural components. This difference could also be explained by the presence of the lignin in the woody parts and absent in the foliage. These results support the hypothesis stating that the more biomass is allocated to a given tree component, the more the C will be stored in that component, and lower C is shared in the parts that physiological functions are intense (Negi and Durgapal, 2003; Bayen et al., 2016; Dimobe et al., 2018; Balima et al., 2019; Meng et al., 2019).

As this study revealed that biomass fractions at tree level were statistically similar between the Sahelian zone and the Sudano-Sahelian zones. The C content stays similar with all tree dbh size. This trend being observed at the components and tree levels between the two climatic zones, suggests that climate does not affect aboveground biomass distribution within B. aegyptiaca tree component, and the C content. Similar results were recently found by Balima et al. (2019) on Afzelia Africana, and these authors concluded that this might be related to a low climatic variation between two neighboring climatic zone as the case of

| Table 5. Regression coefficients with goodness-of-fit statistics of the biomass allometric models for the leaf, branch, and stem components of B. aegyptiaca in the Sudano-Sahelian zone. |
| Models | Predictors | α(SE) | β (SE) | γ | Adj.R² | AIC | RSE | RMSE | CF |
|--------|------------|------|-------|---|-------|-----|------|-------|----|
| Leaf   | Mod. (13)  | Dbh | -4.27 (0.58)*** | 1.50 (0.21)*** | 0.75 | 19.28 | 0.383 | 0.361 | 1.074 |
|        | Mod. (14)  | dbh²*CD | -3.86 (0.52)*** | 0.54 (0.07)*** | 0.75 | 19.09 | 0.370 | 0.360 | 1.070 |
|        | Mod. (15)  | dbh²*Ht | -4.31 (0.63)*** | 0.57 (0.08)*** | 0.71 | 21.82 | 0.413 | 0.411 | 1.087 |
|        | Mod. (16)  | dbh²*CD; *Ht | -3.66 (0.54)*** | 0.69 (0.16)*** | 0.76 | 19.69 | 0.371 | 0.342 | 1.070 |
| Branch | Mod. (17)  | Dbh | -2.85 (0.89)*** | 2.44 (0.22)*** | 0.88 | 20.25 | 0.391 | 0.370 | 1.079 |
|        | Mod. (18)  | dbh²*CD | -2.30 (0.44)*** | 0.89 (0.06)*** | 0.92 | 13.18 | 0.313 | 0.291 | 1.049 |
|        | Mod. (19)  | dbh²*Ht | -2.80 (0.66)*** | 0.94 (0.09)*** | 0.86 | 23.47 | 0.430 | 0.384 | 1.025 |
|        | Mod. (20)  | dbh²*CD; Ht | -2.14 (0.45)*** | 1.03 (0.13) | -0.71 (0.61) | 0.93 | 13.63 | 0.311 | 0.290 | 1.096 |
| Stem   | Mod. (21)  | Dbh | -2.37 (0.88)*** | 2.02 (032)*** | 0.72 | 31.61 | 0.581 | 0.541 | 1.041 |
|        | Mod. (22)  | dbh²*CD | -1.87 (0.77) | 0.74 (0.11)*** | 0.72 | 32.62 | 0.560 | 0.521 | 1.183 |
|        | Mod. (23)  | dbh²*Ht | -2.40 (0.90)*** | 0.78 (0.13)*** | 0.70 | 33.75 | 0.584 | 0.550 | 1.169 |
|        | Mod. (24)  | dbh²*CD; Ht | -1.93 (0.85)*** | 0.68 (0.25)*** | 0.23 (1.13) | 0.70 | 34.56 | 0.581 | 0.813 | 1.183 |

Mod. = model; dbh = diameter breast height; D20 = basal diameter at 20 cm above the ground level, CD = crown diameter; Ht = total tree height, BP = p-value of test Breusch-Pagan of homoscedascity; DW = value of residuals autocorrelation test (Durbin et Watson test), Adj.R² = coefficient of determination; AIC: Akaike Information Criterion; C = correction factor.

| Table 6. Selected biomass allometric equations for B.aegyptiaca in Burkina Faso. |
| Components | Prediction model | Adj.R² | AIC | RSE | RMSE | CF | P-value |
|------------|-----------------|-------|-----|------|-------|----|---------|
| Sahelian zone | Leaf: Mod. (1) Ln(Leaf biomass) = -3.77 + 1.41Ln(dbh) | 0.71 | 20.44 | 0.372 | 1.106 | <0.001 |
|              | Branch: Mod. (5) Ln(branch biomass) = -3.29 + 2.42Ln(D20) | 0.77 | 21.97 | 0.391 | 1.122 | <0.001 |
|              | Stem: Mod. (9) Ln(stem biomass) = -3.74 + 2.12Ln(D20) | 0.86 | 10.57 | 0.251 | 1.163 | <0.001 |
|              | TAGB = -3.77(dbh) + 2.41(D20) | 0.86 | 13.08 | 0.341 | <0.001 |
| Sudano-Sahelian zone | Leaf: Mod. (14) Ln(Leaf biomass) = -3.86 + 0.543Ln(dbh²CD) | 0.75 | 19.09 | 0.360 | 1.070 | <0.001 |
|              | Branch: Mod. (18) Ln(branch biomass) = -2.30 + 0.89In(dbh²CD) | 0.92 | 13.18 | 0.291 | 1.049 | <0.001 |
|              | Stem: Mod. (21) Ln(stem biomass) = -2.37 + 2.02In(dbh) | 0.72 | 31.61 | 0.541 | 1.041 | <0.001 |
|              | TAGB = -3.86(dbh²CD) + 2.41(D20) | 0.92 | 12.82 | 0.281 | <0.001 |

Mod. = model; dbh = diameter breast height; D20 = basal diameter at 20 cm above the ground level; CD = crown diameter; Ht = total height.
Accurate estimates of woody biomass are important and serve to address issues related to forest C sequestration capacity, renewable products and fuels; and to provide comparable and verifiable information to policymakers and stakeholders (MacFarlane, 2015; Xiang et al., 2016). Assessing of vegetation unities biomass (forests, savannas, species, etc.) is thus an effective tool to inform and convince forest resource managers, decision-makers and scientists of the need linked to woody resources protection. Allometric equations were fitted to predict aboveground biomass (and components biomass) of B. aegyptiaca in Burkina Faso in this study.

Log-transformed equations (power models) were fitted with variables such as dbh, basal diameter (D20) and crown mean diameter (CD) as variables with high biomass predictive ability, and showed good performance with high explained variance (p < 0.05; Adj. R² > 0.65) for all tree components in the two climatic zones. This type of equation is widely reported by many authors in the tropical zones of Africa (Henry et al., 2011; Antin et al., 2013; Laminou Manzo et al., 2015; Xiang et al., 2016; Dimobe et al., 2018; Mensah et al., 2018). In addition, the power models are simple and practical in estimating the biomass of several woody species, and this fact lead several authors to prefer them than polynomial and logarithmic equations that have high elasticity (Xiang et al., 2016; Yuen et al., 2016). In this study, the tree size (dbh, CD and D20) is revealed as an important parameter in predicting aboveground biomass of B. aegyptiaca. In the Sahelian zone, D20 was the accurate predictor while dbh was found as best predictor in the Sudano-Sahelian zone. This finding could be explained by the fact that in the Sahelian zone, trees are generally multistemmed under 1.30 m aboveground and quadratic dbh might be less accurate than D20. At tree level, including CD as additional predictor improved the equations in the Sudano-Sahelian zone, but this had no significant improvement on Sahelian zone equations. This suggests that equation improvement through incorporation of additional dendrometric parameters is not a general case for trees (Dimobe et al., 2018; Mensah et al., 2018), but also stand while climatic parameters affect predictors contribution in fitting equations. The important findings are that tree trunk size (dbh or D20) is the best predictor for aboveground biomass estimates. This is consistent to several studies that reported that tree diameter is the best predictor of biomass that generally explains more than 75% of biomass variability (Henry et al., 2011; Mbow et al., 2013). This greatly convinced authors to defend its performance in predicting biomass, and concluded that dbh is the only most commonly used variable which better explains biomass, given the ease and the accuracy of its measurement (Henry et al., 2011; Vahedi et al., 2014; Xiang et al., 2016). We found also in this study that using tree height as an additive predictor did not improve equations parameters. In opposition of that, other authors such as Vahedi et al. (2014), Alvarez et al. (2012) and (MacFarlane, 2015) supported the theory that allometric equations based solely on dbh would underestimate the assessment of aboveground biomass for mature trees. In contrast with that (Bayen et al., 2016), related that using an allometric equation with only the dbh variable would increase the value of the predicted biomass. If for some authors, the dbh overestimates the predicted biomass, and underestimates it for others, this proves that either of these two situations is strongly related to the conditions of its use and not to the variable itself. Thus, it is the most used variable in the prediction equations and some authors have strongly recommended it, in addition with other parameters (Mbow et al., 2013). However, for branch biomass, the CD was the best predictor variable in this study. This is explained by the fact that the size of the crown is determined by the number and/or length of the branches, but

![Graph showing trend plots for predicted and observed values of total aboveground biomass (TAGB) in Sahelian zone (A) and Sudano-Sahelian zone (B).](image)

Figure 5. 1:1 trend plots for predicted and observed values of total aboveground biomass (TAGB) in Sahelian zone (A) and Sudano-Sahelian zone (B).

### Table 7. Comparison of dbh mean values across land cover types and climate zones.

| Climatic Zones          | LULC types       | dbh (±SD) (cm) | Mean dbh (cm) |
|-------------------------|------------------|----------------|---------------|
| Sahelian zone           | Agroforestry parklands | 12.45 (±2.93) a | 11.65 (±0.47) a |
|                         | Wooded grasslands | 14.02 (±7.69) b |               |
|                         | Shrubby and grasslands | 11.35 (±1.11) a |               |
| Sudano-Sahelian zone    | Agroforestry parklands | 13.80 (±5.00) b |               |
|                         | Shrubby savanas   | 12.52 (±3.51) a | 12.32 (±2.28) a |
|                         | Wooded grasslands | 11.98 (±2.87) a |               |
|                         | Shrubby grasslands | 10.64 (±2.90) a |               |

LULC = Land use Land Cover, SD = Standard deviation.

Within each climatic zone and across zones, LULCs marked with same capital letters are not significantly different (p > 0.05).
et al., 2003). The lack of light needed for photosynthesis produces more biomass as it is a full-light tree species benefitting from the light conditions which make it grow faster like those in the low-density plant formations, and therefore contribute to a better understanding of the C dynamics in these ecosystems. If research on C storage and the C cycle in different ecosystems is needed to establish plans on climate change management and improvement of agroforestry parklands in semi-arid areas.

The selection of single variable in the Sahelian zone was based on two reasons: first the Ht in the equation did not improve the fitting quality (Table 4) and secondly, it was necessary to avoid Adj. R² inflation induced by the association of several variables in the same equation (Mbow et al., 2013). Some authors recommended the association ofdbh with total height for equation fitting (Sawadogo et al., 2010; Bayen et al., 2016; Dimobe et al., 2018). However, other studies consider that height is an undesirable variable that should not be included in equations because it presents difficulties in having accurate measurements (Jenkins et al., 2003). The lack of fitting improvement by adding tree height might be due to the shrubby status of this species which prefers having more branches than total height (Arbonnier, 2009).

It has been reported that additivity in biomass is recognized as a desirable way to fit models to predict biomass of tree component (Riofrío et al., 2015; Dimobe et al., 2018). In fact, simultaneous fits with related equations and using additive procedures are statistically more efficient, as they take into account statistical dependencies among biomass components recorded from the same tree (Parresol, 1999, 2001).

The adjusted coefficient of R² (Adj.R²) obtained in these equations (Adj.R² ranged from 71% to 92 %) despite climatic zone indicated the reliability of the predictions, especially sustained by the low values of RSME and AIC. Considering the tree main LUCCs in the Sahelian zone, an estimated C potential of 3.57 t/ha was found while the Sudano-Sahelian zone stored 1.83 t/ha following the dynamics of estimated stem density. This suggests that there is a substantial potentiality for C storage if the species of savannas and grasslands are integrated in local plantations and associated with farmers’ preferred species.

Oelbermann et al. (2004) reported that this can reach 2.1 Gt C/year in tropical zone. This reveals the need for evaluation of the stored C of each ecosystem through the specific allometric equations of each species. This opinion has been supported by many studies and still need to be sustained to ensure the accuracy of the C estimates (Djomo et al., 2010, 2016; Mbow et al., 2013; Tumweba et al., 2013) and to make practicable the REDD + programs in tropical Africa.

5. Conclusions and application

Reliable and accurate biomass and C estimation methods can be used to determine the geographic distribution of C stocks and to understand changes in C stocks related to tree communities. In this study, it is shown that tree biomass is more allocated to branch share and leaf share of biomass is lower. C content is higher in wood (branch and stem) than in the foliage. Across the two climatic zones, dbh, biomass and C content keep the same trend; their values are not statically different. This study revealed the high correlation between AGB and dbh of B aegyptiaca. Specific equations have been developed on the basis of dbh and D20 to predict its biomass allocation in different parts of tree. The equations fitted in this study are highly reliable and can be used to predict the biomass of this species in similar tropical ecosystems to our study environment. If research on C storage and the C cycle in different ecosystems contributes to a better understanding of the C dynamics in these ecosystems, the development of specific equations for each agroforestry species is needed to establish plans on climate change management and management of agroforestry parklands in semi-arid areas.

Declarations

Author contribution statement

Sambo Ouédraogo: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

Oumarou Ouédraogo: Conceived and designed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data.

Kangbén Dimobe: Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data.

Adjima Thiombiano, Joseph I. Boussim: Conceived and designed the experiments; Contributed reagents, materials, analysis tools or data.

Table 8. Mean biomass and carbon quantities per LULC types in the two climatic zones.

| LULC types          | LULC trees number | Mean Biomass (tons) | Biom min (tons) | Biom max (tons) | Mean C (tons) | C potential (tons/ha) |
|---------------------|-------------------|---------------------|-----------------|-----------------|--------------|------------------------|
| Sahelian zone       | 4 157 059         | 441 729             | 44 729          | 2 450 503       | 242 951 a    | 3.57                   |
| Agroforestry parklands | 1 052 682       | 111 858             | 11 326          | 620 535         | 61 521       | 0.41                   |
| Wooded grassland   | 1 141 908         | 121 339             | 12 286          | 673 131         | 66 736       | 2.9                    |
| Shrubby and grassland | 1 962 469       | 208 532             | 21 116          | 1 156 836       | 114 692      | 0.26                   |
| Sudano-Sahelian zone | 56 551 327   | 6 009 144           | 608 492         | 33 335 876      | 3 305 029 b | 1.83                   |
| Agroforestry parks | 16 414 059        | 1 744 158           | 176 615         | 96 75 759       | 95 286       | 0.25                   |
| Shrubby savannas   | 33 869 999        | 3 599 026           | 364 441         | 19 965 687      | 1 979 464   | 0.53                   |
| Wooded grassland   | 27 240 059        | 2 894 293           | 293             | 16 057 1 591    | 1 591 029   | 0.17                   |
| Shrubby grassland  | 6 240 029         | 663 065             | 67 142          | 3 678 372       | 364 686     | 0.88                   |
| Total C (t)        |                   |                     |                 |                 | 3 547 980    |                        |

LULC = land use land cover type, Biom min = Minimum value of biomass, Biom max = maximum value of biomass. Different letters connected to C amount at climate level mean there is are significant difference between theme (P < 0.05).
Funding statement

This work was supported by the International Foundation for Science with the Committee on Scientific and Technological Cooperation (COM-SECH) GRANT NUMBER: IFS-76E3A177, the Grand Duchy of Luxembourg through Lux-development & the Swedish Embassy-Burkina Faso.

Competing interest statement

The authors declare no conflict of interest.

Additional information

No additional information is available for this paper.

Acknowledgements

Authors are grateful to Grand Duchy of Luxembourg through Lux-development, the Swedish Embassy-Burkina Faso, for supplementary material support and commodity.

References

Alvarez, N., Manel, S., Schmidt, T., 2012. Contrasting distortion of Quaternary gene pools across Europe: the case of the arctic-alpine Gentiana nivalis L. (Gentianaceae). Flora - morphological Distribution. Functional Ecology of Plants 207, 408-413.

Artin, C., Pelletier, R., Vincent, G., Couturon, P., 2013. Crown allometric equations are less responsive than stem allometry to tree size and habitat variations in an Indian monsoon forest. Trees (Berl.) 27, 1485-1495.

Arbonnier, M., 2009. Arbres, arbustes et lianes des zones sèches d’Afrique. Versailleslas: Éd. Quae. MININ, 574 p.

Balma, L.H., Nacoulma, B.M.I., Bayen, P., Dimobe, K., Kouame, F.N., Thiomibano, A., 2019. Aboveground biomass allometric equations and distribution of carbon stocks of the African oak (Afzelia africana Sm.) in Burkina Faso. J. For. Res. Banuki, T.M., van Laake, P.E., Skidmore, A.K., Hussin, Y.A., 2009. Allometric equations for biomass data from Africa. For. Ecol. Manag. 257, 1684–1694.

Bayen, P., Bognounou, P., Zhike, A.M., Ouédraogo, A., Thiomibano, A., 2016. The use of biomass production and allometric models to estimate carbon sequestration of Jatropha curcas L. plantations in western Burkina Faso. Environ. Dev. Sustain. 18, 143–156.

Bong, C., Ouédraogo, O., Thiomibano, I., Boussim, J.I., 2018. Variability and estimating in fruiting of shea tree (Vitellaria paradoxa C.F. Gaertner) associated to climatic conditions in West Africa: implications for sustainable management and development. Plant Prod. Sci. 22, 1–16.

Chair, J., Lourenbach, S., Orekan, V.O.A., Kyi-Baffour, N., 2016. Allometric models and aboveground biomass stocks of a West African Sudan Savannah watershed in Benin. Carbon Bal. Manag. 11, 16.

Chave, J., Andalo, C., Brown, S., Cairns, M.A., Chambers, J.Q., Eamus, D., Folster, H., Fromard, F., Higuchi, N., Kira, T., Lescure, J.P., Nelson, B.W., Ongawa, H., Puig, H., Riera, B., Yamakura, T., 2005. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. Oecologia 145, 78–99.

Chave, J., Réjou-Meuch, M., Bûrquez, A., Chidumayo, E., Colgan, M., Delimiti, W., Duque, A., Eid, T., Feurolle, P., Goodman, R., Al E., 2014. Improved allometric models to estimate the aboveground biomass of tropical trees. Global Change Biol. 20, 3177–3190.

Dimobe, K., Mensah, S., Goetze, D., Ouédraogo, A., Kuyah, S., Poorembbi, S., Thiomibano, A., 2018. Aboveground biomass partitioning and additive models for Combretum glutinosum and Terminalia laxiflora in West Africa. Biomass Bioenergy 115, 151–159.

Dimobe, K., Kuyah, S., Dabré, Z., Ouédraogo, A., Thiomibano, A., 2019. Diversity-carbon stock relationship across vegetation types in W National park in Burkina Faso. For. Ecol. Manag. 438, 243–254.

Djomo, A.N., Ibirahima, A., Saborowski, J., Gravenhorst, G., 2010. Allometric equations for biomass estimations in Cameroon and pan moist tropical ecologies including biomass data from Africa. For. Ecol. Manag. 260, 1873–1885.

Djomo, A.N., Picard, N., Fayolle, A., Henry, M., Ngomanda, A., Ploton, P., McLellan, J., Saborowski, J., Adamou, I., Lejeune, P., 2016. Tree allometry for estimation of carbon stocks in African tropical forests. Forestry 89, 446–455.

Dong, L., Zhang, L., Li, F., 2016. Developing two additive biomass equations for three coniferous plantationspecies innermost China. Forests 11–1163.

FAO, 1981. 62p. Manuel D’inventaire Foréster, 31. Etude FAO-Forre no, Rome.

FAO, 2002. Inventaire forestier national - Manuel de terrain – Modele. FAO, 70pp. http://www.fao.org/3/a9758e/ae578f04.htm.
Xiang, W., Zhou, J., Ouyang, S., Zhang, S., Lei, P., Li, J., Deng, X., Fang, X., Forrester, D.I., 2016. Species-specific and general allometric equations for estimating tree biomass components of subtropical forests in southern China. Eur. J. For. Res. 135, 963–979.

Xue, Y., Yang, Z., Wang, X., Lin, Z., Li, D., Su, S., 2016. Tree biomass allocation and its model additivity for Casuarina equisetifolia in a tropical forest of hainan island, China. PloS One 11 (3), e0151856.

Yuen, J.Q., Fung, T., Ziegler, A.D., 2016. Review of allometric equations for major land covers in SE Asia: uncertainty and implications for above- and below-ground carbon estimates. For. Ecol. Manag. 360, 323–340.

Zellner, A., 1962. An efficient method of estimating seemingly unrelated regression equations and tests for aggregation bias. J. Am. Stat. Assoc. 57 (298), 348–368.