agronomy

MDPI

Article

Non-Destructive Estimation of the Leaf Weight and Leaf Area in Common Bean

Juan Carlos Suárez [1,2,*], Fernando Casanoves [3] and Julio Di Rienzo [4]

[1] Programa de Ingeniería Agroecológica, Facultad de Ingeniería, Universidad de la Amazonia, Florencia 18001, Colombia
[2] Centro de Investigaciones Amazónicas CIMAZ Macagual Cesar Augusto Estrada González, Grupo de Investigaciones Agroecosistemas y Conservación en Bosques Amazónicos—GAIA, Florencia 18001, Colombia
[3] CATIE—Centro Agronómico Tropical de Investigación y Enseñanza, Turrialba 30501, Costa Rica; casanoves@catie.ac.cr
[4] Facultad de Ciencias Agropecuarias, Universidad Nacional de Córdoba, Cordoba 5016, Argentina; dirienzo.julio@gmail.com
* Correspondence: ju.suarez@udla.edu.co; Tel.: +57-3202804455

Abstract: Regression models to predict leaf area and leaf weight in common bean (*Phaseolus vulgaris*) were fitted using the three leaflets of the leaves. A total of 1504 leaves from 40 genotypes were collected, covering a large range of leaf sizes. Width, length, area, and weight were measured for each leaflet. The total leaf area and weight was obtained by the sum of left, central, and right leaflets. The dataset was randomly divided into training and validation sets. The training set was used for model fitting and selection, and the validation dataset was used to obtain statistics for model prediction ability. The leaf area and leaf weight were modeled using different linear regression models based on the length and width of the leaflet. Polynomial regressions involving both length and width of the leaflet provided very good models to estimate the expected area (R^2 = 0.978) and weight (R^2 = 0.820) of leaves.

Keywords: multiple regression analysis; leaf morphology; model validation

Citation: Suárez, J.C.; Casanoves, F.; Di Rienzo, J. Non-Destructive Estimation of the Leaf Weight and Leaf Area in Common Bean. *Agronomy* **2022**, *12*, 711. https://doi.org/10.3390/agronomy12030711

Academic Editor: Pasquale Losciale

Received: 6 December 2021
Accepted: 6 March 2022
Published: 16 March 2022

1. Introduction

One of the main crops in the diet of the tropics of Latin America and East Africa is the common bean (*Phaseolus vulgaris* L.). The common bean is produced by smallholders in areas with high climatic vulnerability. Due to their importance, there are currently many issues in relation to the adaptation of improved varieties to different environmental conditions, particularly in the face of imminent challenges posed by global climate change. Leaf area is crucial to evaluating water stress [1], as a larger leaf area leads to greater hydric conductance [2].

To determine the adaptation of bean materials to different conditions, growth should be evaluated and simulated using mathematical models that describe the relationship between plant growth, dry matter production, and expansion of total leaf area. One of the ways to evaluate the interactions between environmental conditions and genotypes is the area and weight of the leaf, which are considered indicators of growth and yield in crops.

Determining leaf area and weight requires the destructive sampling of many leaves, a process that can be time-consuming and labor-intensive, as well as have an impact on plant growth. Regression models allow the prediction of leaf area and weight using variables that can be measured without destructive sampling techniques. In addition, this process allows the repeated measurement of leaves over time, while avoiding the biological alteration characteristic of destructive methods. *Phaseolus vulgaris* leaves are trifoliate. Therefore, the length and width of each leaflet can be used to estimate the leaf total area and weight.

Several models have been reported to predict leaf area in legumes such as Alfalfa (*Medicago sativa*) [3], chestnut (*Castanea sativa*) [4], chickpea (*Cicer arietinum* L.) [5], cowpea

(*Vigna unguiculata* L.) [6], faba bean (*Vicia faba* L.) [7,8], peanut (*Arachis hypogaea* L.) [9,10], pigeon pea (*Cajanus cajan* L.) [11], and soybean (*Glycine max*) [12–14], where equations were developed from the length or width of some of the leaflets; however, no models are reported to predict leaf weight and area from leaflet measurements for a large number of genotypes of *Phaseolus vulgaris*.

Specifically for the common bean, a few models have been generated to predict leaf area. The first linear model was reported by Lakitan [15], whose objective was to predict the area of the central and lateral leaflets for three bean genotypes from the length and width of a leaflet. Muñoz et al. [16] proposed to measure leaf area by multiplying the length by the width of the leaflet corrected by a correction factor; then, Iamauti [17] and De Jesus et al. [18] mentioned an empirical model from leaflet length, a situation that was also reported by Figueiredo et al. [19] for the cultivar Perola type Carioca, Toebe et al. [20] for cultivar "Macarrão" and Queiroga et al. [21] for cultivar UEL-2, but these last two specify only the use of the central leaflet to predict leaf area. Bhatt and Chanda [22] reported different models for bean cultivation under different stress conditions. Likewise, Ramírez-Builes et al. [23] reported models for different bean genotypes (BAT 477, Morales, SER 16 and SER 21), some of them resistant to water-stress conditions, based on length, width, and its product, as well as Hara et al. [24]. Finally, Lakitan et al. [25] proposed a simplified procedure for the non-destructive estimation of the trifoliate leaf area using the length and width of the leaflets. Therefore, the objective of this study was to determine an empirical model to estimate the leaf area and weight of different lines of the common bean from the linear dimensions of the leaflets, with an emphasis on 40 advanced modern lines developed with the capacity to adapt to drought, high temperatures, low soil fertility, aluminum toxicity, resistance to pests, and high nutrient content. These lines were obtained by crossing *P. vulgaris* with other species of genus *Phaseolus* sp.

2. Materials and Methods

2.1. Plant Material and Experimental Design

Evaluation of leaflet morphometric measurements in bean genotypes was carried out in the Centro de Investigaciones Amazonicas CIMAZ Macagual, Universidad de la Amazonia, Colombia (1°37′ N and 75°36′ W) located in Florencia, Caquetá (Colombia) in a tropical rainforest ecosystem. The location has an average annual precipitation of 3800 mm with 1700 h of sunshine per year, an average temperature of 25.5 °C, and an average relative humidity of 84%. A total of 40 improved lines of common bean from the Mesoamerican and Andean gene pools were used developed by CIAT's Bean Breeding Program: 18 advanced lines of *Phaseolus vulgaris* (ALB 349, AMADEUS, BFS 10, BFS 29, ICA QUIMBAYA, ICTA PETEN, SCR 23, SCR 40, SCR 61, SEN 48, SEN 70, SER 323, SER 324, SMC 234, SMN 98, SMR 181, SMR 182, SMR 185), 3 interspecific lines from *P. vulgaris* × *P. coccineus* which resulted in a group of Mesoamerican gene pool lines (ALB 350, ALB 352, ALB 60) and Andean (ALB 267, RRA 57), 2 interspecific lines from *P. vulgaris* × *P. acutifolius* (SER 213, SIN 351-1), 5 interspecific lines from *P. vulgaris* × *P. dumosus* (SMC 101, SMC 33, SMC 41, SMN 24, SMR 84), 4 interspecific lines from *P. vulgaris* × *P. acutifolius* × *P. coccineus* (SEF 1, SEF 10, SEF 49, SEF 70, SEF 73), 1 interspecific lines from *P. vulgaris* × *P. acutifolius* × *P. dumosus* (SMR 195), 4 interspecific lines from *P. vulgaris* × *P. acutifolius* × *P. dumosus* × *P. coccineus* (SMC 199, SMC 209, SMC 216, SMR 140).

These selected genotypes were resistant to different abiotic and biotic stress factors, high micronutrient content, and other desirable attributes that enhance adaptation and grain yield. The ALB (small red kidney, black kidney) lines have enhanced adaptation to drought and Al toxicity. AMADEUS is a high-temperature tolerant improved line, the BFS lines (small red) have enhanced adaptation to low soil fertility, RRA (various colored) were developed to generate resistance to Andean root rot, genus *Pythium* sp. and *Sclerotium* sp. The SEF (red), SEN (black) and SER (small red) lines were improved to adapt to drought and heat. The SIN (black) lines were interspecific with drought tolerance. SCR (red), SMC (colored), SMR (red) and SMN (black) lines are tolerant to drought with a high mineral

content in seeds (Fe). ICA QUIMBAYA (red) is a commercial variety resistant to bean golden mosaic virus and Al toxicity, and ICTA PETEN (black) is a commercial variety, it has tolerance to rust and golden mosaic virus with high Fe content.

Approximately twenty plants per plot were sown in a greenhouse for each genotype, and the plots were arranged in a nonreplicated complete randomized design. The same substrate was used in each plot. The plot size was 2 m × 2 m, with three furrows at 0.6 m between them. For each plot, three plants were selected from the middle part of the central furrow to control border effects for a total of 120 plants (3 plants × 40 genotypes). For each plant, between 4 and 19 leaves (depending on the plant size) were systematically selected (one out of three) from the 7th leaf developed from the base to the top of the plant in the growth stage of pod filling (R_8), which corresponds to a period of 70 to 80 days after sowing. To predict leaf area and weight, we measured the length and width of each leaflet to see if the morphological relationship between them remains constant for leaves of different ages.

We obtained the measurements of morphometric variables for a total of 4512 leaflets of 1504 leaves. Each leaflet (left, right, and center) on each leaf was weighed independently using an Ohaus Scout electronic scale (100 ± 0.001 g), after which it was scanned using the HP ScanJet Pro 2500 scanner. The length, width, and area of each leaflet was measured from its scanned image using the ImageJ program V1 [26]. From this information, we calculated descriptive statistics for each of the morphometric variables evaluated, as well as calculated the Pearson's correlation coefficients between morphometric variables.

2.2. Data Analysis

Only 28 genotypes from a total of 40 had enough observations ($n \geq 14$) to fit their own complete polynomial model (Model 1, see below). Nevertheless, the observations of these 28 genotypes (n = 1398) represented 93.2% of the complete dataset (n = 1500). To compare the performance of genotype-specific models against a unique one, we used both strategies to obtain predictions. Then, we calculated correlation concordance coefficients for the predictions under the two strategies. For total leaf area, the correlation concordance coefficient was 0.99, whereas for total leaf weight it was 0.93. According to these preliminary results, we decided to fit unique models for all genotypes which are more suitable for practical applications.

We used multiple linear regression models to predict leaf area and weight from leaflets length and width. Due to the curvilinear trends between the dependent (leaf weight and area) and predictor variables (leaflet length and width), multiple linear regression models including second-order polynomials on length, width or both variables, and the interaction (length × width) were used to estimate the expected value of leaflet weight and area. The best model was selected using AIC and BIC criteria. The smaller the better, for both criteria. Because none of them were uniformly the best, we decided to choose a model in which both criteria agree.

We evaluated five different models according to the morphometric measures used (c = central, r = right, and l = left leaflets; W = width, WW = width^2, L = length, LL = length^2 and LW = length × width):

Model 1: to estimate leaf area and weight using all the leaflet data:

$$\begin{aligned} Y_i = \beta_0 + \beta_1\, cL_i + \beta_2\, cW_i + \beta_3\, cLL_i + \beta_4\, cWW_i + \beta_5\, lL_i + \beta_6\, lW_i + \beta_7\, lLL_i + \\ \beta_8\, lWW_i + \beta_9\, rL_i + \beta_{10}\, rW_i + \beta_{11}\, rLL_i + \beta_{12}\, rWW_i + e_i \end{aligned} \quad (1)$$

Model 2: to estimate leaf area and weight using the central leaflet data:

$$Y_i = \beta_0 + \beta_1\, cL_i + \beta_2\, cW_i + \beta_3\, cLL_i + \beta_4\, cWW_i + e_i \quad (2)$$

Model 3: to estimate leaf area and weight using the left leaflet data:

$$Y_i = \beta_0 + \beta_1\, lL_i + \beta_2\, lW_i + \beta_3\, lLL_i + \beta_4\, lWW_i + e_i \quad (3)$$

Model 4: to estimate leaf area and weight using the right leaflet data:

$$Y_i = \beta_0 + \beta_1 \, rL_i + \beta_2 \, rW_i + \beta_3 \, rLL_i + \beta_4 \, rWW_i + e_i \quad (4)$$

Model 5: to estimate leaf area and weight using only the leaflet width data:

$$Y_i = \beta_0 + \beta_1 \, cW_i + \beta_2 \, cWW_i + \beta_3 \, lW_i + \beta_4 \, lWW_i + \beta_5 \, rW_i + \beta_6 \, rWW_i + e_i \quad (5)$$

For all models, e_i ($i = 1, \ldots, 1200$) is the error term assumed to be normal and independently distributed, with zero mean and a common variance.

The models using only length information from the three leaflets were not reported because they were much less predictive than those based only on width.

From the total of the 1504 leaves recorded, we obtained 1500 valid data values. We randomly selected 1200 leaves to conform the training set used to estimate the models. The remaining 300 leaves were used as a validation set to measure the predictive ability of a fitted model, calculating, from this set, the coefficient of determination (R^2) and prediction mean square error (PMSE). We used the adjusted R^2 in multiple linear regression models and R^2 in the case of simple regression models. Adjusted R^2 (or R^2) and PMSE are summary statistics commonly used for the comparison of models. However, they do not say anything about patterns of departure of observed values against predicted. A simple and effective way to visualize any trouble regarding this issue is to draw a scatter plot of observed vs. predicted values and overlap a reference line ($y = x$) as well as the regression line of observed vs. predicted values. Departures of the regression line from the reference line can suggest problems of over or underestimation [27]. All the statistical analyses were done in InfoStat version 2021 [28] and R version 3.6.1 [29].

3. Results

3.1. Summary Measures of Morphometric Variables Evaluated in the Different Bean Lines

The summary statistics are shown for each of the morphometric variables evaluated (Table 1). The maximum and minimum statistics of the regressors are useful to approximate the range of prediction space in which the prediction model is reliable. Pairwise Pearson's correlation coefficients were also calculated among the set of morphometric variables (Figure 1). The correlation between leaf total area and the leaflet measurements (Figure 1) was higher than the correlation between leaf total weight and leaflet measurements (Figure 2) (ranged between 0.886–0.931 and 0.78–0.86, respectively). For leaf total area, the width measurements of the leaflets were more correlated than those of leaflet length (ranged between 0.886–0.9 and 0.91–0.93, respectively). The leaf total weight and the width measurements of the leaflets were more correlated than with those of leaflet length (ranged between 0.78–0.826 and 0.804–0.813, respectively).

Table 1. Summary statistics for morphometric variables of leaflets of improved lines of common bean. The sample was taken to cover a large range of leaf sizes.

Variable	Leaflet	Mean	Std. Dev	Minimum	Maximum
Area (cm^2)	Right	38.24	20.62	1.70	132.78
	Central	39.65	20.09	2.33	132.76
	Left	39.50	20.58	2.48	146.99
	Total	117.40	59.34	8.94	412.53
Weight (g)	Right	0.57	0.35	0.02	2.17
	Central	0.60	0.35	0.03	2.13
	Left	0.58	0.34	0.03	2.28
	Total	1.75	1.01	0.09	6.58

Table 1. *Cont.*

Variable	Leaflet	Mean	Std. Dev	Minimum	Maximum
Length (cm)	Right	8.77	2.30	2.51	16.84
	Central	9.13	2.31	2.53	17.88
	Left	8.91	2.34	2.40	17.14
	Mean	8.93	2.21	2.48	16.94
Width (cm)	Right	6.27	1.89	1.02	12.86
	Central	6.59	1.95	1.18	12.65
	Left	6.40	1.93	1.03	13.34
	Mean	6.42	1.84	1.20	12.95

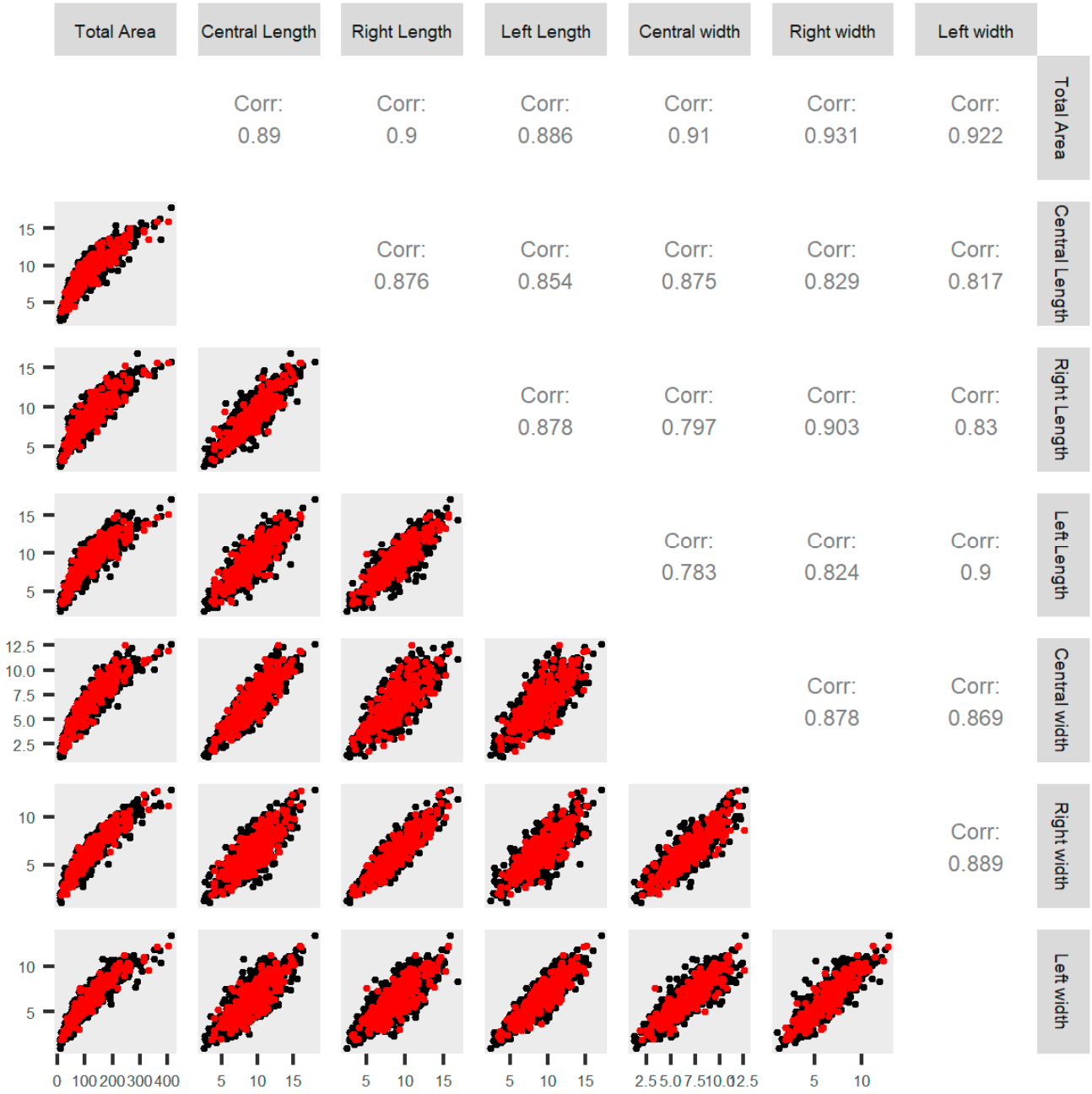

Figure 1. Scatter plots matrix between leaf total area versus length and width of the central, right, and left leaflet of leaves in common bean. Validation observations in red, training observations in black. Corr: Pearson correlation coefficients.

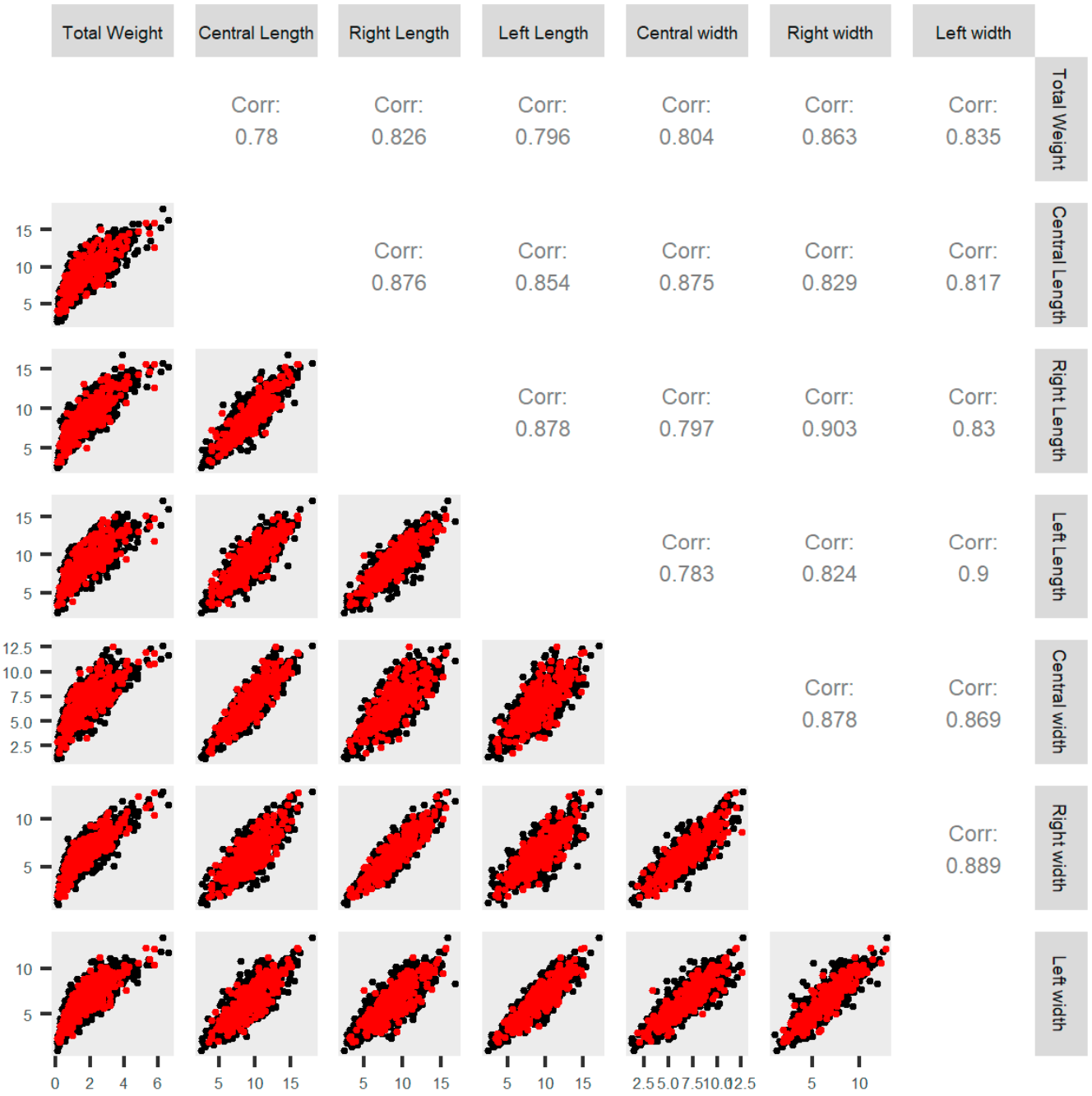

Figure 2. Scatter plots matrix between leaf total weight versus length and width of the central, right, and left leaflet of leaves in common bean. Validation observations in red, training observations in black. Corr: Pearson correlation coefficients.

3.2. Regression Models

In all the models, evaluating the inclusion of interaction (length × width) was not necessary ($p > 0.10$), but the second-order polynomial term on leaflet length and leaflet width was included. In addition, the inclusion of an interaction term in any model increases both AIC and BIC values. Regression coefficients for the five models described above were estimated for both dependent variables: weight and leaf area (Table 2). In total area estimation, Model 1, which is a second-order polynomial on all leaflet length and leaflet width, without their interaction, has the lowest PMSE and the greatest adjusted R^2 (90.95 and 0.9781, respectively) and has the least values for both AIC and BIC. The models to estimate the total area from the individual leaflets (Models 2 to 4) have greater PMSE and

lesser adjusted R^2 (ranged between 345.65 and 368.13, and ranged between 0.9108 and 0.9154, respectively). The model to estimate leaf area using only the width information (Model 5) has a PMSE = 177.72 and adjusted R^2 = 0.9571 (Table 2).

Table 2. Regression coefficients of the five models evaluated to predict leaf area and weight from leaflet measures of improved lines of common bean. Prediction mean square error (PMSE) and determination coefficient (adjusted R^2) are calculated on the validation dataset.

	Leaf Area			Leaf Weight			
Parameter	**Coefficient**	**t**	***p*-Value**	**Coefficient**	**t**	***p*-Value**	**Model**
β_0	1.360385	0.39	0.6958	0.1151	0.71	0.4752	1. All the leaflets' measures
β_1 cL	−1.217674	−0.78	0.4375	−0.0025	−0.03	0.9730	
β_2 cW	−1.851119	−1.14	0.2540	0.0310	0.41	0.6800	
β_3 cLL	0.192913	2.48	0.0134	−0.0004	−0.11	0.9114	
β_4 cWW	0.719912	6.67	0.0000	0.0044	0.87	0.3828	
β_5 lL	2.264534	1.33	0.1848	−0.1135	−1.44	0.1512	
β_6 lW	1.858530	0.99	0.3209	0.1704	1.97	0.0496	
β_7 lLL	0.057770	0.66	0.5113	0.0081	2.00	0.0460	
β_8 lWW	0.351000	2.68	0.0075	−0.0050	−0.82	0.4113	
β_9 rL	1.732512	0.97	0.3303	−0.0600	−0.73	0.4670	
β_{10} rW	−3.661254	−1.92	0.0553	0.0144	0.16	0.8706	
β_{11} rLL	0.095602	1.03	0.3039	0.0077	1.79	0.0736	
β_{12} rWW	0.845674	6.17	0.0000	0.0128	2.02	0.0439	
R^2	0.9781			0.8205			
PMSE	90.9593			0.2033			
AIC	8447.72			1492.26			
BIC	8518.09			1562.63			
β_0	13.614836	2.16	0.0313	0.2187	1.20	0.2317	2. Only central leaflet
β_1 cL	−6.117789	−2.63	0.0087	−0.1971	−2.93	0.0035	
β_2 cW	6.701626	2.59	0.0097	0.2465	3.29	0.0010	
β_3 cLL	0.866885	7.51	0.0000	0.0175	5.23	0.0000	
β_4 cWW	0.837329	4.71	0.0000	0.0036	0.70	0.4834	
R^2	0.9164			0.6988			
PMSE	345.6574			0.3412			
AIC	9930.04			1944.81			
BIC	9960.20			1974.97			
β_0	−3.350078	−0.52	0.6007	0.0972	0.57	0.5669	3. Only left leaflet
β_1 lL	2.210764	0.82	0.4139	−0.1692	−2.36	0.0185	
β_2 lW	2.094504	0.67	0.5002	0.2575	3.12	0.0018	
β_3 lLL	0.319671	2.28	0.0230	0.0149	4.01	0.0001	
β_4 lWW	1.364022	6.09	0.0000	0.0056	0.95	0.3434	
R^2	0.9153			0.7545			
PMSE	350.0354			0.2784			
AIC	10,045.55			1858.97			
BIC	10,075.51			1889.13			
β_0	−8.177508	−1.42	0.1562	−0.0085	−0.05	0.9570	4. Only right leaflet
β_1 rL	5.809652	2.39	0.0170	−0.0901	−1.36	0.1734	
β_2 rW	−1.468532	−0.54	0.5879	0.1679	2.28	0.0229	
β_3 rLL	0.152891	1.19	0.2348	0.0114	3.25	0.0012	
β_4 rWW	1.676578	8.44	0.0000	0.0133	2.46	0.0140	
R^2	0.9108			0.8092			
PMSE	368.1351			0.2160			
AIC	9755.55			1616.08			
BIC	9785.71			1646.24			

Table 2. *Cont.*

	Leaf Area			Leaf Weight			
Parameter	**Coefficient**	**t**	***p*-Value**	**Coefficient**	**t**	***p*-Value**	**Model**
β_0	7.928129	−7.16	0.0000	−0.0813	−4.67	0.0000	
β_1 cW	1.895257	3.54	0.0004	0.0458	2.47	0.0136	
β_2 rW	−3.845477	0.57	0.5674	−0.0339	0.54	0.5915	
β_3 lW	2.634387	2.14	0.0326	0.0507	0.71	0.4752	
β_4 cWW	0.543651	10.16	0.0000	0.0022	−0.10	0.9212	
β_5 rWW	1.212421	11.62	0.0000	0.0230	6.21	0.0000	5. All the leaflets width
β_6 lWW	0.608981	9.14	0.0000	0.0076	2.92	0.0035	
R^2	0.9571			0.8109			
PMSE	177.72			0.2142			
AIC	9001.98			1556.72			
BIC	9042.19			1596.93			

β_0 = intercept, $\beta_1, \ldots, \beta_{12}$ = correlation coefficients, c = central, r = right, and l = left leaflets; W = width, WW = width^2, L = length, LL = length^2 and LW = length × width, AIC = Akaike information criteria, BIC = Bayesian information criteria.

In total weight estimation, Model 1 has the lowest PMSE and the greater adjusted R^2 (0.2033 and 0.8205, respectively) and has the least values for both AIC and BIC. The models to estimate the total weight from the individual leaflets (Models 2 to 4) have the greatest PMSE and lower adjusted R^2 (PMSE ranged between 0.2160 and 0.3412, and R^2 ranged between 0.6988 and 0.8092, respectively). The model to estimate leaf weight using only the width information (Model 5) has a PMSE = 0.2142 and adjusted R^2 = 0.8109 (Table 2).

In order to evaluate the prediction quality, we used scatter plots of observed against predicted values, adding the reference line y = x. A departure of the regression line of observed vs. predicted from the reference line is evidence of bias. For leaf area estimation (Figure 3), the model with the least bias is Model 2. However, because it only considers the central leaflet, its R^2 is lower than the R^2 in Model 1. Model 1 shows a little departure from the reference line for large values of area, showing a little underestimation. Model 4 is similar in bias to Model 1 but shows a smaller R^2. Finally, Model 5 exhibits a similar behavior to Model 1, also underestimating for large-area leaves but showing a high R^2 in both cases. When modeling the area and considering the PMSE (smaller is better), these two models have lower PMSE (90.95 and 177.72, respectively) and higher R^2 than the other models. For this reason, we recommend using Model 1 or Model 5, but it is important to consider that Model 5 only needs width measures of leaflets (more parsimonious model).

For leaf weight estimation (Figure 4), the model with the least bias is Model 4. However, because it only considers the right leaflet, its R^2 is lower than the Model 1 R^2 (considering the three leaflets). In the same way as for area estimation, Model 1 shows a little departure from the reference line for large values of weight, showing a little underestimation. Model 2 and Model 3 show a little bias for large weight values but underestimate and overestimate, respectively. Finally, Model 5 shows a similar behavior to Model 1, also underestimating for large leaf weights but showing a high R^2 in both cases (0.81 and 0.82, respectively). Considering the PMSE, these two models have lower PMSE (0.20 for Model 1 and 0.21 for Model 5) and higher R^2 than the other models. In weight estimation, Model 5 has almost the same PMSE and R^2 as Model 1, but only uses width measures of leaflets (more parsimonious model).

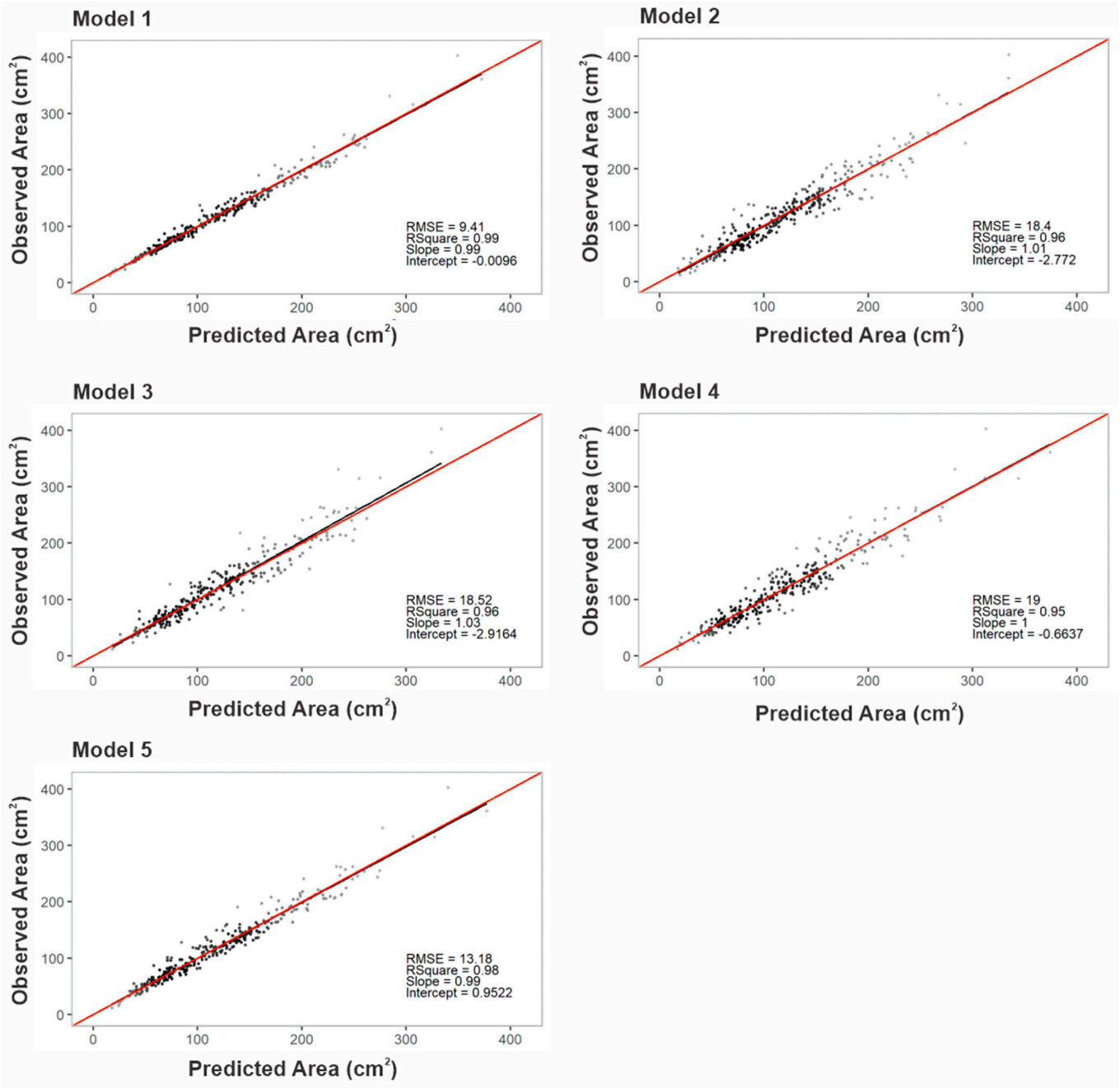

Figure 3. Relationship between the observed leaf area and the predicted leaf total area of improved lines of common bean for the five evaluated models. Regression line of observed versus predicted values: red line, reference line (x = y): black line.

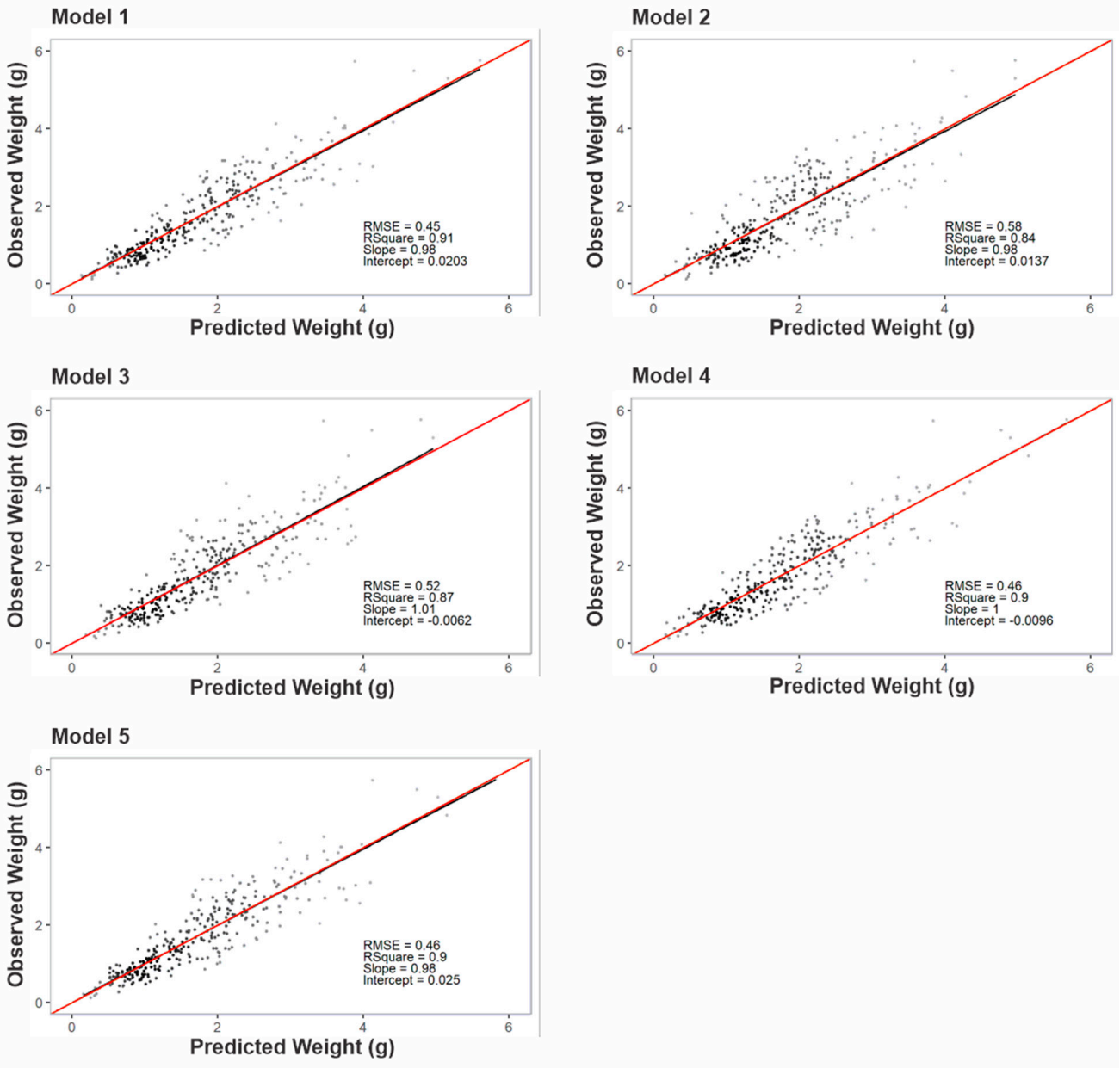

Figure 4. Relationship between the observed weight and the predicted leaf total area of improved lines of common bean for the five evaluated models. Regression line of observed versus predicted values: red line, reference line (x = y): black line.

4. Discussion

The estimation of the leaf total area and weight by non-destructive methods is widely used in studies of physiology (photosynthetic capacity) and management intensity (fertilization levels, water availability), among others [30–33]. The use of destructive methods has a disadvantage when the aim is to evaluate growth, because the use of different leaves through time increases the experimental error [21,34].

Some research proposes statistical models to estimate leaf area and weight in legumes [3–14]. In all cases, the linear regressions used involve measurements of length and width (L and W) of one or more leaflets of the leaf. Furthermore, due to the non-linear relationship between these measurements and leaf area or weight, their squares (LL and WW) or their interactions (LW) are incorporated (e.g., Lakitan et al. [25]).

In the case of *Phaseolus vulgaris*, the leaf is trifoliate, therefore, the length and width of each leaflet can be obtained to estimate the leaf's total area and weight. A few models have

been proposed to estimate the leaf area in beans using linear regressions [20–22,25]. These models were made using one, two or three cultivars. The major novelty of our research lies in the dataset covering a large variety of genotypes and leaf sizes and ages, which makes it a general-use model. In addition, we propose models to estimate leaf weight based on L and W measurements of the leaflets. In our models, estimates were made using a training dataset and a validation dataset, which allowed us to obtain independent measurements of the PMSE and R^2 statistics [35]. Among the models found for beans, only Lakitan et al. [25] and Queiroga et al. [21] used a validation dataset.

The number of leaves measured in this research (1500) far exceeds the sample size used by other authors. Due to the number of intervening genotypes, compared to other research, the range of variation of the length and width measurements of each leaflet was much greater, increasing the range of validity of the models that we propose in this research. The minimum and maximum W (1.02 cm; 13.34 cm) and the minimum and maximum L (2.40 cm; 17.88 cm) in this study have both a larger range than reported by Lakitan et al. [25] (W: 0.9 cm; 9 cm, L: 1.5 cm; 15.5 cm). The range of regressor values should always be reported [36], however, sometimes the ranges are not provided (e.g., Queiroga et al. [21]).

Several studies have determined that some species present more regularity in the area of the central leaf compared with the lateral ones, finding that there is more correlation between length and width of the central leaf with respect to the other two [21]. In our research, the correlation between L and W for central leaflet was 0.87, while the correlation between L and W for right and left leaflets was 0.90 in both cases, showing a similar behavior for the shape of the three leaflets. This different result may be due to the fact that in our study, there were 40 genotypes instead of one, as with Queiroga et al. [21]. Marshall [37] and Chanda et al. [38] point out that this correlation can change according to the plant age and environmental conditions.

The best models in our research do not include the LW interaction, but they include the second-order polynomial term. Lakitan et al. [25] developed models using L, W, and LW. However, they use zero intercept in their models, because for them "it makes sense to force intercept to zero, since if L, W or LW is zero, then leaf area should also be zero". This is a common error in regression models. When the zero-intercept regression model is fitted, it should not be based only on conceptual issues but mainly on the data. If there are not near-zero-regressor values available, when assuming zero intercept, we are also assuming the linearity of the model in a range of values were the regressors were not observed. Forcing the model to have zero-intercept because of a non-sense estimated intercept (for example negative area) leads, at the same time, to miss-fitting the model for observed values.

Pompelli et al. [39] and Lakitan et al. [25] found that the power model predicted the leaf area with good accuracy but increasing heteroscedastic residual dispersion compared with the use of LW. We could not verify such a behavior.

We obtained the best model to estimate leaf area and weight using all the leaflet information (L and W) and its second-order polynomial (LL and WW, Model 1). These models showed only a little underestimation for large leaf area and weight values. In addition, estimates of leaf total area and weight are provided based on the length and width measurements of leaflets, achieving that with only two measurements, there are models with R^2 greater than 0.9 in the case of total area. In addition, for total area estimation, the models using the central, left, or right leaflet measurements are very similar in terms of their PMSE and R^2.

For weight, there is a marked difference between the estimates obtained from the individual leaflets. Those from the length and width of the right leaflet shows the highest R^2 and lowest PMSE than the central and left leaflets. The correlation between leaf area and weight with W measurements was higher than the correlation with L measures for all leaflets. Queiroga et al. [21] estimated the leaf area by measuring only the width of the central leaflet. Using a logarithmic regression with zero intercept they obtained a good percentage of explained variability (R^2 = 0.98), but they do not evaluate the presence of

under or overestimation of their models. The selection of the best model must be based not only on R^2 and PMSE but also use some measure of accuracy [27].

As in Queiroga et al. [21], in our research, we found a good estimation of leaf total area and weight only based on the measurement of the width of the three leaflets. Other authors found more correlation between width and leaf area than length (for instance, Zaffaroni [40] in *Vigna unguiculata* L. and Bianco [41] in *Amarantus retroflexus* L.). The width of the leaflets on the leaf is easier to measure than the length, due, among other things, to the boundary between the lamina and the petiole of the leaflet. The results using only the width of the leaflets gave estimates with R^2 values very similar to those corresponding to the use of the two dimensions for both total area and weight, but unlike the area estimation, the weight showed a very similar PMSE between Models 1 and 5, meaning that only measuring the width of the leaflets is better in the case of leaf weight estimation. We also evaluated a model using only central W as in Queiroga et al. [21], but the estimated model (Area = 1.19 + 5.86 W + 1.66 WW) showed an R^2 = 0.86. This result is a consequence of the number of genotypes in our study, and the corresponding differences between W in the central, leaf, and right leaflets are because of the genotype effect.

5. Conclusions

Second-order polynomial regression involving both the length and width of each leaflet provided a very good fit to estimate the expected area and weight of leaves in *Phaseolus vulgaris*. However, it is possible to estimate total leaf area and weight with good accuracy using a second-order polynomial regression involving only the width of each leaflet. The loss of accuracy due to the use of this simplified model is less when estimating weight than when estimating area. Because the purpose of the predictive models we are presenting is not specific to a given genotype, using any model depending on only one leaflet is less accurate than a model using all of them.

Author Contributions: Conceptualization, J.C.S. and F.C.; methodology, J.C.S., F.C. and J.D.R.; software, J.C.S., F.C. and J.D.R.; validation, J.C.S., F.C. and J.D.R.; formal analysis, J.C.S., F.C. and J.D.R.; investigation, J.C.S.; resources, J.C.S.; data curation, J.C.S. and F.C.; writing—original draft preparation, J.C.S. and F.C.; writing—review and editing, J.C.S. and F.C.; visualization, J.C.S., F.C. and J.D.R.; supervision, J.C.S.; project administration, J.C.S.; funding acquisition, J.C.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available from the authors upon request.

Acknowledgments: We acknowledge the financial support from the CGIAR Research Program on Grain Legumes and Dryland Cereals for the development of breeding lines of common bean. We would also like to thank all donors who supported this work through their contributions to the CGIAR Fund.

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

References

1. Zhang, C.; Zhang, J.; Zhang, H.; Zhao, J.; Wu, Q.; Zhao, Z.; Cai, T. Mechanisms for the relationships between water-use efficiency and carbon isotope composition and specific leaf area of maize (*Zea mays* L.) under water stress. *Plant Growth Regul.* **2015**, *77*, 233–243. Available online: https://link.springer.com/article/10.1007/s10725-015-0056-8 (accessed on 12 November 2021). [CrossRef]
2. Sellin, A.; Õunapuu-Pikas, E.; Kaurilind, E.; Alber, M. Size-dependent variability of leaf and shoot hydraulic conductance in silver birch. *Trees* **2011**, *26*, 821–831. Available online: https://link.springer.com/content/pdf/10.1007/s00468-011-0656-5.pdf (accessed on 12 November 2021). [CrossRef]
3. Bianco, S.; De Carvalho, L.B.; Bianco, M.S. Determination of alfalfa leaf area by non-destructive method. *Commun. Plant Sci.* **2011**, *1*, 17–20. Available online: https://www.researchgate.net/publication/224884475 (accessed on 12 November 2021). [CrossRef]

4. Serdar, Ü.; Demirsoy, H. Non-destructive leaf area estimation in chestnut. *Sci. Hortic.* **2006**, *108*, 227–230. [CrossRef]
5. Nahas, S.B.; Arce, O.E.A.; Ricci, M.; Romero, E.R. Leaf area estimation of individual leaf and whole plant of chickpea (*Cicer arietinum* L.) by means of regression methods. *Rev. Agron. Noroeste Argent.* **2019**, *39*, 99–106. Available online: https://www.researchgate.net/publication/338867825 (accessed on 12 November 2021).
6. de Oliveira, R.L.L.; Moreira, A.R.; da Costa, A.V.A.; de Souza, L.C.; Lima, L.G.S.; da Silva, R.T.L. Modelos de determinação não destrutiva de área foliar de feijão caupi *Vigna unguiculata* (L.). *Glob. Sci. Technol.* **2015**, *8*, 17–27. [CrossRef]
7. Erdoğan, C. A Leaf Area Estimation Model for Faba Bean (*Vicia faba* L.) Grown in the Mediterranean Type of Climate. *SDU J. Fac. Agric. Ziraat Fakültesi Derg.* **2012**, *7*, 58–63.
8. Peksen, E. Non-destructive leaf area estimation model for faba bean (*Vicia faba* L.). *Sci. Hortic.* **2007**, *113*, 322–328. [CrossRef]
9. Ma, L.; Gardner, F.P.; Selamat, A. Estimation of Leaf Area from Leaf and Total Mass Measurements in Peanut. *Crop Sci.* **1992**, *32*, 467–471. Available online: https://acsess.onlinelibrary.wiley.com/doi/full/10.2135/cropsci1992.0011183X003200020036x (accessed on 14 November 2021). [CrossRef]
10. Kathirvelan, P.; Kalaiselvan, P. Groundnut (*Arachis hypogaea* L.) Leaf Area Estimation Using Allometric Model. *Res. J. Agric. Biol. Sci.* **2007**, *3*, 59–61.
11. Pezzini, R.V.; Filho, A.C.; Alves, B.M.; Follmann, D.N.; Kleinpaul, J.A.; Wartha, C.A.; Silveira, D.L. Models for leaf area estimation in dwarf pigeon pea by leaf dimensions. *Bragantia* **2018**, *77*, 221–229. Available online: http://dx.doi.org/10.1590/1678-4499.2017106 (accessed on 19 November 2021). [CrossRef]
12. Nehbandani, A.; Soltani, A.; Zeinali, E.; Raeisi, S.; Najafi, A. Allometric relationships between leaf area and vegetative characteristics in soybean. *Int. J. Agric. Crop. Sci.* **2013**, *6*, 1127–1136.
13. Bakhshandeh, E.; Kamkar, B.; Tsialtas, J.T. Application of linear models for estimation of leaf area in soybean [*Glycine max* (L.) Merr]. *Photosynth* **2011**, *49*, 405. [CrossRef]
14. Bakhshandeh, E.; Ghadiryan, R.; Kamkar, B. A rapid and non-destructive method to determine the leaflet, trifoliate and total leaf area of soybean. *Aust. J. Plant Sci. Biotechnol.* **2010**, *4*, 19–23.
15. Lakitan, B. Statistical Design and Analysis for Intercropping Experiments. *Annu. Rep. Bean. Improv. Coop.* **1989**, *32*, 19–21.
16. Muñoz, G.; Giraldo, G.; Fernández de Soto, J. *Descriptores Varietales: Arroz, Fríjol, Maíz, Sorgo*; Publication No. 177; Ciat: Cali, Colombia, 1993.
17. Iamauti, M.T. Avaliação de Danos Causados por Uromyces appendiculatus No Feijoeiro. Ph.D. Thesis, Universidade de São Paulo, Piracicaba, Brazil, 1995.
18. de Jesus, W.C.; Vale, F.X.R.D.; Coelho, R.R.; Costa, L.C. Comparison of two methods for estimating leaf area index on common bean. *Agron. J.* **2001**, *93*, 989–991. Available online: https://acsess.onlinelibrary.wiley.com/doi/full/10.2134/agronj2001.935989x (accessed on 1 November 2021). [CrossRef]
19. da Silva Figueiredo, E.; dos Santos, M.E.; Garcia, A. Modelos de determinação não destrutivo da área foliar do feijoeiro comum (*Phaseolus vulgaris* L.). *Nucleus* **2012**, *9*, 79–84. Available online: https://dialnet.unirioja.es/servlet/articulo?codigo=3988133&info=resumen&idioma=POR (accessed on 19 November 2021). [CrossRef]
20. Toebe, M.; Filho, A.C.; Loose, L.H.; Heldwein, A.B.; Zanon, A. Área foliar de feijão-vagem (*Phaseolus vulgaris* L.) em função de dimensões foliares. *Semin. Agrar.* **2012**, *33* (Suppl. S1), 2491–2500. Available online: https://www.researchgate.net/publication/278121722 (accessed on 19 November 2021). [CrossRef]
21. Queiroga, J.L.; Romano, E.D.U.; Souza, J.R.P.; Miglioranza, É. Estimativa da área foliar do feijão-vagem (*Phaseolus vulgaris* L.) por meio da largura máxima do folíolo central. *Hortic. Bras.* **2003**, *21*, 64–68. [CrossRef]
22. Bhatt, M.; Chanda, S.V. Prediction of leaf area in Phaseolus vulgaris by non-destructive method. *Bulg. J. Plant Physiol.* **2003**, *29*, 96–100.
23. Ramírez-Builes, V.H.; Porch, T.G.; Harmsen, E.W. Development of linear models for estimation of leaflet area in common bean (*Phaseolus vulgahs* L.). *J. Agric. Univ. Puerto. Rico.* **2008**, *92*, 171–182. Available online: https://pubag.nal.usda.gov/catalog/31098 (accessed on 22 November 2021).
24. Hara, A.T.; Gonçalves, A.C.A.; Maller, A.; Hashiguti, H.T.; de Oliveira, J.M. Ajuste de modelo de predição de área foliar do feijoeiro em função de medidas lineares. *Rev. Eng. Agric.* **2019**, *27*, 179–186. Available online: https://www.researchgate.net/publication/332774019 (accessed on 4 November 2021). [CrossRef]
25. Lakitan, B.; Widuri, L.I.; Meihana, M. Simplifying procedure for a non-destructive, inexpensive, yet accurate trifoliate leaf area estimation in snap bean (*Phaseolus vulgaris*). *J. Appl. Hortic.* **2017**, *19*, 15–21. Available online: http://horticultureresearch.net/jah/2017_19_1_15_21.PDF (accessed on 4 November 2021). [CrossRef]
26. Ferreira, T.; Ferreira, T.; Rasb, W. ImageJ User Guide: IJ 1.46r. 2012. Available online: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.731.215 (accessed on 19 November 2021).
27. Suárez, J.C.; Melgarejo, L.M.; Durán Bautista, E.H.; Di Rienzo, J.A.; Casanoves, F. Non-destructive estimation of the leaf weight and leaf area in cacao (*Theobroma cacao* L.). *Sci. Hortic.* **2018**, *229*, 19–24. [CrossRef]
28. Di Rienzo, J.; Casanoves, F.; Balzarini, M.; Gonzalez, L.; Tablada, M.; Robledo, C.W. InfoStat versión 2021. Grupo InfoStat, FCA, Universidad Nacional de Córdoba, Argentina. Available online: http://www.infostat.com.ar.2021 (accessed on 10 November 2021).
29. R Development Core Team. R: The R Project for Statistical Computing. Foundation for Statistical Computing, V., Austria. 2021. Available online: https://www.r-project.org/ (accessed on 9 November 2021).

30. Blanco, F.F.; Folegatti, M.V. Leaf area for greenhouse cucumber under salinity and grafting. *Sci. Agric.* **2005**, *62*, 305–309. [CrossRef]
31. De Swart, E.A.M.; Groenwold, R.; Kanne, H.J.; Stam, P.; Marcelis, L.F.M.; Voorrips, R.E. Non-destructive estimation of leaf area for different plant ages and accessions of *Capsicum annuum* L. *J. Hortic. Sci. Biotechnol.* **2004**, *79*, 764–770. Available online: https://www.tandfonline.com/doi/abs/10.1080/14620316.2004.11511840 (accessed on 19 November 2021). [CrossRef]
32. Rouphael, Y.; Rivera, C.M.; Cardarelli, M.; Fanasca, S.; Colla, G. Leaf area estimation from linear measurements in zucchini plants of different ages. *J. Hortic. Sci. Biotechnol.* **2006**, *81*, 238–241. Available online: https://www.tandfonline.com/doi/abs/10.1080/14620316.2006.11512056 (accessed on 19 November 2021). [CrossRef]
33. Lizaso, J.I.; Batchelor, W.D.; Westgate, M.E. A leaf area model to simulate cultivar-specific expansion and senescence of maize leaves. *Field Crops Res.* **2003**, *80*, 1–17. [CrossRef]
34. Silva NFDa Ferreira, F.A.; Fontes, P.C.R.; Cardoso, A.A. Modelos para estimas a área foliar de abóbora por meio de medidas lineares. *Revista Ceres.* **1998**, *45*, 287–291. Available online: http://www.ceres.ufv.br/ojs/index.php/ceres/article/download/2488/482 (accessed on 19 November 2021).
35. Draper, N.R.; Smith, H. *Applied Regression Analysis Applied Regression Analysis*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2014; pp. 1–716. Available online: https://books.google.es/books?hl=es&lr=&id=d6NsDwAAQBAJ&oi=fnd&pg=PR13&dq=DRAPER,+N.+R.%3B+SMITH,+H.+A.+Applied+regression+analysis.+3.+ed.+New+York:+Wiley,+1998.+736+p&ots=Byo8lbj2PR&sig=z3y9N4qbsQqWP5OLAFfLXjtQ1Gg (accessed on 19 November 2021).
36. Balzarini, M.; Di Rienzo, J.; Tablada, M.; Gonzalez, L.; Bruno, C.; Córdoba, M.; Casanoves, F. *Estadística y Biometrías. Ilustraciones del Uso de Infostat en Problemas de Agronomía*; Segunda Edición; Editorial Brujas: Córdoba, Argentina, 2012; 380p.
37. Marshall, J.K. Methods for leaf area maeasurement of large and small leaf samples. *Photosynthetica* **1968**, *2*, 41–47.
38. Chanda, S.; Joshi, A.; Vaishnav, P.; Singh, Y. Leaf area determination in pearl millet using linear measurements-area and matter-area relationships. *Photosynthetica* **1985**, *19*, 424–427.
39. Pompelli, M.F.; Antunes, W.C.; Ferreira, D.T.R.G.; Cavalcante, P.G.S.; Wanderley-Filho, H.C.L.; Endres, L. Allometric models for non-destructive leaf area estimation of Jatropha curcas. *Biomass Bioenergy* **2012**, *36*, 77–85. [CrossRef]
40. Zaffaroni, E. Estimativa da área foliar em feijão macassar (*Vigna unguiculata* (L.) Walp) a partir de medidas dos folíolos. *Agropecuária Técnica* **1981**, *2*, 80–85.
41. Bianco, S.; Pitelli, R.; Pavani, M.; Silva, R. Estimativa da área foliar de plantas daninhas. XIII-Amaranthus retroflexus L. *Ecossistema* **1995**, *20*, 5–9.