DOI: 10.1007/S11099-011-0048-5 PHOTOSYNTHETICA 49 (3): 405-416, 2011

Application of linear models for estimation of leaf area in soybean [*Glycine max* (L.) Merr]

E. BAKHSHANDEH[*,+], B. KAMKAR[*], and J.T. TSIALTAS[**]

Department of Agronomy, Gorgan University of Agricultural Science and Natural Resources, P.O. Box 49189-43464, Gorgan, Iran[*]
Aristotle University of Thessaloniki, School of Agriculture, Laboratory of Agronomy, 541 24 Greece[**]

Abstract

Leaf area estimation is an important measurement for comparing plant growth in field and pot experiments. In this study, determination of the leaf area (LA, cm^2) in soybean [*Glycine max* (L.) Merr] involves measurements of leaf parameters such as maximum terminal leaflet length (L, cm), width (W, cm), product of length and width (LW), green leaf dry matter (GLDM) and the total number of green leaflets per plant (TNLP) as independent variables. A two-year study was carried out during 2009 (three cultivars) and 2010 (four cultivars) under field conditions to build a model for estimation of LA across soybean cultivars. Regression analysis of LA *vs*. L and W revealed several functions that could be used to estimate the area of individual leaflet (LE), trifoliate (T) and total leaf area (TLA). Results showed that the LW-based models were better (highest R^2 and smallest RMSE) than models based on L or W and models that used GLDM and TNLP as independent variables. The proposed linear models are: LE = 0.754 + 0.655 LW, (R^2 = 0.98), T = –4.869 + 1.923 LW, (R^2 = 0.97), and TLA = 6.876 + 1.813 ΣLW (summed product of L and W terminal leaflets per plant), (R^2 = 0.99). The validation of the models based on LW and developed on cv. DPX showed that the correlation between calculated and measured LA was strong. Therefore, the proposed models can estimate accurately and massively the LA in soybeans without the use of expensive instrumentation.

Additional key words: dry mass; leaf dimensions; number of leaflets; soybean.

Introduction

Soybean is an important source of protein and oil for human and animal consumption (Setiyono *et al*. 2008). LA is a parameter related to radiation interception, photosynthesis, biomass accumulation, transpiration and gas exchange in crop canopies (Jonckheere *et al*. 2004, Kandiannan *et al*. 2009) and it is also important in crop-weed competition (Akram-Ghaderi and Soltani 2007). LA is useful for the analysis of canopy architecture as it allows the determination of leaf area index (LAI) and thus, accurate measurements of LA are essential for understanding the interaction between crop growth and environment (de Jesus *et al*. 2001). Measuring the surface area of a large number of leaves can be both time-consuming and labour costing. Many methods have been evolved to facilitate the measurement of LA. The total leaf area (TLA) of the plant can be obtained by either direct or indirect methods. LA is generally determined by direct methods, which consists of removing and measuring all the leaves of the plant. However, these methods, including those of tracing, blueprinting, photographing, require instruments, tools and machines such as hand scanners and laser optic apparatuses (Peksen 2007). An alternative method to measure LA is to use image analysis with image measurement and analysis software. The capture of an image by a digital camera is rapid, and the analysis using proper software is accurate (Bignami and Rossini 1996, Rico-Garcia *et al*. 2009), but the processing is time-consuming, and the facility is

Received 9 November 2010, *accepted* 23 May 2011.
[+]Corresponding author; Tel (fax): +98-171-4427060, , e-mail: Bakhshandehesmail@yahoo.com
Abbreviations: ALA – actual leaf area; GLDM – green leaf dry matter; L – maximum length terminal leaflet; LA – leaf area; LE – leaflet area; LW – product of length and width of terminal leaflet; OLA – observation leaf area; PLA – prediction leaf area; R^2 – coefficient of determination; R^2_a – adjusted coefficient of determination; RMSE – root mean square error; T – trifoliate area; TLA – total leaf area; TNLP – total number of green leaflets per plant; TV – tolerance values; VIF – variance inflation factor; W – maximum width of terminal leaflet; SRO – stepwise regression option.
Acknowledgments: We thank Mr R. Ghadiryan and Mrs N. Ramzanzadeh for their kind help during the experimentation.

generally expensive and not suitable for nonflat leaves measurement, because pictures taken not exactly perpendicularly can cause erroneous LA estimations. However, these are very expensive and complex devices for basic and simple studies (Serdar and Demirsoy 2006), it is not feasible to conduct successive measurements of the same leaf, and plant canopy can be damaged, which might cause problems to other measurements or experiments (Rouphael *et al*. 2010). LA can be measured quickly, accurately, and nondestructively by using a portable scanning planimeter (Daughtry 1990), but it is suitable only for small plants with few leaves (Nyakwende *et al*. 1997) and not feasible for large leaves.

Indirect methods are useful when these equipments are not available or nondestructive measurements are needed, such as field conditions or low plant density growing in pots of controlled experiments (Peksen 2007). Especially when using unique plants, for example in genetically segregating populations, nondestructive measurements are of great value (Rouphael *et al*. 2010). In nondestructive methods, LA is usually estimated by measuring the number, width or length of plant parts or whole plant, *e.g.,* leaf width, length and number, branch length and number, and plant height (Akram-Ghaderi and Soltani 2007). These measurements can be undertaken without cutting the plants. The indirect methods enable researchers to measure LA on the same plants during the plant growth period and may reduce variability in experiments (Gamiely *et al*. 1991, Nesmith 1991, de Swart *et al*. 2004, Serdar and Demirsoy 2006).

Simple, accurate models eliminate the need for expensive LA meters or time-consuming geometric reconstructions (Gamiely *et al*. 1991). Therefore, an inexpensive, rapid, reliable and nondestructive method for measuring LA is required by the agronomists. If the mathematical relationships between LA and leaf dimensions could be clarified, a nondestructive model can be developed and to be more advantageous than many of the methods mentioned above (Villegas *et al*. 1981, Beerling and Fry 1990). Also, there are numerous of LA prediction models based on the total number of leaves or leaflets per plant, green leaf dry matter or total aboveground biomass dry matter. Various models relating traits mentioned above to LA have been developed for several crops (Bhatt and Chanda 2003, de Jesus *et al.* 2001, Peksen 2007, Kathirvelan and Kalaiselvan 2007, Ma *et al*. 1992, Tsialtas and Maslaris 2005, 2008a,b; Tsialtas *et al.* 2008, Bange *et al*. 2000, Rouphael *et al*. 2007, 2010; Shih *et al*. 1981, Lu *et al*. 2004, Shin and Snyder 1984, Payne *et al*. 1991, Akram-Ghaderi and Soltani 2007, Soltani *et al*. 2006, Lieth *et al*. 1986, Ramos *et al*. 1983, Mokhtarpour *et al*. 2010 and Kumar 2009).

Although few models have been developed for nondestructive LA estimation in soybean cultivars (Wiersma and Bailey 1975), we are unaware of any attempt to test these models for their accuracy in PLA in other cultivars. Any proposed model for nondestructive LA predictions should be tested for its accuracy, because leaf shape formation is strongly affected by genetic (Stoppani *et al*. 2003) and environmental factors (Tsialtas *et al*. 2008). The main aim of this study was to develop models for LA prediction from the terminal leaflet dimensions, the total number of green leaflets per plant and the green leaf dry matter per plant in soybean that would be able to compensate for changes in leaf shape between cultivars and which could be used for various cultivars without recalibration. Also, we aimed to assess the reliability of the developed models by comparing them with previously proposed models.

Materials and methods

Data collection: During 2009 and 2010 growing seasons, soybean plants were grown at the experimental field (36°85'N, 54°27΄'E and 13 m a.s.l.) of the Faculty of Plant Production, Gorgan University of Agricultural Sciences and Natural Resources (GUASNR), Gorgan, Iran. The climate is temperate subhumid (Soltani and Hoogenboom 2003). Monthly maximum and minimum temperatures, mean of irradiance and rainfall were measured at a standard weather station located a few hundred meters from the experimental unit. Weather conditions during the experiments (2009 and 2010) and long term ones are given in Table 1.

In the first experiment three soybean cultivars (DPX, Sahar, and Williams) were sown on 4 July 2009. In the

Table 1. Means of maximum and minimum temperature, mean of irradiance and sum of rainfall for monthly periods during the experiments (2009 and 2010) compared with long-term statistics (40 years) at Gorgan weather condition (1967–2007).

Month	Maximum temperature [°C]			Minimum temperature [°C]			Mean of irradiance [MJ m^{-2} day^{-1}]			Sum of rainfall [mm]		
	2009	2010	Long-term	2009	2010	Long-term	2009	2010	Long-term	2009	2010	Long-term
June	29.9	39.2	29.6	18.8	14.0	18.4	19.4	24.6	21.6	13.1	0.00	35.7
July	34.7	40.6	32.0	23.4	20.2	21.9	23.1	23.9	23.2	5.90	15.8	32.1
August	30.9	35.6	32.3	23.3	24.8	22.7	15.3	24.6	21.3	41.2	0.00	27.7
September	29.9	32.5	29.8	20.7	21.6	19.6	19.1	19.7	19.2	82.6	29.3	41.6
October	26.7	29.4	24.8	16.1	18.0	13.9	14.4	14.8	14.5	86.1	35.0	64.9
November	22.9	27.3	21.6	11.1	13.0	11.7	11.3	12.6	11.8	70.5	50.8	75.4

second experiment, four cultivars (DPX, Sahar, Williams, and Hill) were sown on 25 June 2010. These cultivars were selected as the most representative of the cultivars cropped in Northern Iran. Some information of the cultivars properties are given in Table 2. The experimental design was the randomized complete block with 4 replications. Plot sizes were 5 m long with row spacing of 45 cm and included 5 rows. Plant population density was 32 plants m^{-2}. Seeds were planted by hand at the soil depth of 5 cm. Based on a soil analysis results (depth of 0–30 cm), 150 kg ha^{-1} of triple super phosphate and 50 kg ha^{-1} of potassium sulfate was used at sowing time and without using nitrogen fertilizer during growing season. The soil was a deep silty clay loam. Irrigation was supplied when necessary. Weeds were hand-controlled and if necessary appropriate chemicals were applied against pests and diseases, so the effect of diseases, pests and weeds were minimal in both years.

Sampled leaves did not present any damage and deformation caused by diseases, insects or other factors. Leaf samples varied in size from large to small for each cultivar, they were selected randomly from different plant parts and at different growth stages ranging from flowering stage (R1) to full seeds (R6) (Fehr and Caviness 1977). In each sampling, immediately after cutting, leaves were placed in plastic bags and were transferred to the GUASNR laboratory for further measurements. A total of 3,385 leaflets, (~450 and 550 leaflets for each cultivar in 2009 and 2010, respectively) were measured in order to develop the best-fitted models for the LA prediction. Maximum terminal leaflet width (W) (from end-to-end between the widest lobes of the lamina perpendicular to the lamina midrib) and length (L) (from lamina tip to the point of petiole intersection along the midrib) were measured to the nearest 0.1 cm by a ruler. Also, in both experiments a total of 1,048 plants, from all cultivars were selected and their actual leaf area (ALA) related to summed lengths (ΣL), widths (ΣW) and product of length and width terminal leaflets per plant (ΣLW) (244 plants), green leaf dry matter per plant (GLDM) (444 plants) and the total number of green leaflets per plant (TNLP) (360 plants) were determined to plot against each other. In order to calculate GLDM, the leaf samples of each plant were oven-dried at 70°C till constant mass.

ALA of all terminal leaflets (LE), trifoliate (T) and TLA per plant separately were measured with a digital area meter (*WDIGC-2, DELTA-T Devices,* Durham, UK). Means, standard deviations, minimum and maximum values of the L, W and LE area for each soybean cultivar in each year and pooled data are shown in Table 3.

Table 2. Some information about the cultivars used in this study.

Cultivars	Growth type	Maturity	Flower colour	Mean height [cm]
Sahar	Determinate	Late	White	85
Hill	Determinate	Early	White	85
Williams	Indeterminate	Early	White	110
DPX	Indeterminate	Late	Purple	120

Table 3. Mean ± SD (standard deviations), minimum (Min) and maximum (Max) values for the terminal leaflet length (L), width (W) and leaflet area (LA) used for model building and validation of soybean cultivars (data are from pool all the different growth stages (R1 to R6) for each cultivar). *n* – the number of observation data.

Cultivar	*n*	L [cm] Mean ± SD	Min	Max	W [cm] Mean ± SD	Min	Max	LA [cm^2] Mean ± SD	Min	Max
For model building (2009)										
DPX	444	10.2 ± 2.77	2.20	15.30	5.86 ± 2.15	1.30	10.3	43.48 ± 25.74	0.62	101.4
Sahar	437	8.31 ± 2.40	3.10	13.70	4.46 ± 1.53	1.50	7.50	27.19 ± 15.72	1.91	64.69
Williams	328	8.71 ± 2.56	3.00	14.30	5.35 ± 1.77	1.30	9.00	34.86 ± 20.10	3.01	81.97
Pooled data	1,209	9.13 ± 2.72	2.20	15.30	5.22 ± 1.94	1.30	10.3	35.25 ± 22.14	0.62	101.4
For model validation (2009)										
Random sample, DPX	100	10.3 ± 3.03	4.30	15.00	5.97 ± 2.34	1.50	10.0	45.44 ± 28.43	2.25	98.27
For model building (2010)										
DPX	544	8.95 ± 1.92	3.50	13.50	4.60 ± 1.26	1.70	7.80	29.04 ± 11.98	8.77	65.19
Sahar	532	8.93 ± 2.19	4.10	14.00	4.47 ± 1.45	1.70	8.00	38.84 ± 14.73	4.07	68.13
Williams	542	9.18 ± 2.21	3.50	13.20	4.81 ± 1.26	2.20	8.00	31.30 ± 12.64	9.56	63.48
Hill	558	8.71 ± 2.28	3.50	14.20	4.27 ± 1.35	1.50	8.00	25.93 ± 13.34	3.93	64.80
Pooled data	2,176	8.94 ± 2.16	3.50	14.20	4.54 ± 1.35	1.50	8.00	28.76 ± 13.34	3.93	68.13
For model validation (2010)										
Random sample, DPX	200	9.68 ± 2.02	4.00	13.80	4.96 ± 1.66	2.10	8.60	31.62 ± 16.45	6.62	77.05

Model building: The relationships were evaluated by fitting regression models with the linear regression procedure of *SAS* (*SAS* Institute 1992) and the stepwise elimination option as reported by Miranda and Royo (2003). The internal validity of the models was tested by coefficient of determination (R^2) and root mean square error (RMSE) calculated as:

$$\mathrm{RMSE} = \left[\Sigma(\mathrm{P} - \mathrm{O})^2 / (\mathrm{n} - 1)\right]^{0.5} \quad (1)$$

where P is the predicted LA, O is the observed LA and, n is the number of observations.

Residuals were analyzed to determine the presence of outliers and nonconstant error variance. Outlier is defined as:

$$\mathrm{Outlier} = \begin{cases} 0 & \text{if} \left| r_i \right| \le k\sigma \\ 1 & \text{Otherwise} \end{cases} \quad (2)$$

where, by default k = 3, and scale σ is computed as corrected median of the absolute residuals (Cankaya *et al*. 2006).

LE, T, and TLA were the dependent variables and L, W, LW, ΣL, ΣW, ΣLW, GLDM, and TNLP were also considered as independent variables. For each cultivar in each year, a one-parameter model was fitted, separately. For detecting collinearity using two measurements (*e.g.,* L and W), the variance inflation factor (VIF) (Marquardt 1970, Neter *et al*. 1996) and the tolerance values (TV) (Gill 1986) were calculated as follow:

$$\mathrm{VIF} = \frac{1}{1 - R^2} \quad (3)$$

$$\mathrm{TV} = \frac{1}{\mathrm{VIF}} \quad (4)$$

where R^2 is the coefficient of determination. For the VIF values higher than 10 or TV values smaller than 0.10, then collinearity may have more than a trivial impact on the estimates of the parameters and consequently one of them should be excluded from the model (Cristofori *et al*. 2007, Fallovo *et al*. 2008, Rouphael *et al*. 2010).

Model validation: In order to validate the produced models over pooled data from all cultivars in both years, two validation experiments were conducted in 2009 and 2010 on leaf samples of DPX grown at the experimental farm of GUASNR, Iran. Data used in the model validation were obtained from leaf samples at different parts of plant canopy at different growth stages (R1 to R6). A total of 100 and 200 leaflets were used to determine LE and T, in 2009 and 2010, respectively. Also, 176 plants selected and ALA related to GLDM (56 plants), ΣL, ΣW, ΣLW (40 plants), L and W expressed in cm, and TNLP (81 plants) were determined by the previously described procedures in both years. GLDM and TNLP were determined only in 2010.

The models proposed by Wiersma and Bailey (1975) were evaluated for their accuracy in LA prediction in the present study:

LE = 0.624 + 0.723 LW
T = 0.411 + 2.008 LW
TLA = 6.532 + 2.045 ΣLW (summed product of L and W of terminal leaflets per plant).

Three techniques were used to validate the models:

(*1*) the validation data set was used to produce a validation model by re-estimating the model parameters using the Stepwise Regression Option (SRO) approach to develop the estimation model and the models were compared for consistency (Miranda and Royo 2003, 2004),

(*2*) in order to compare the predicted leaf area (PLA) to the observed leaf area (OLA), graphical procedures were used (Bland and Altman 1986),

(*3*) values for PLA were subtracted from OLA and differences were plotted against the PLA for each model. Lack of agreement was evaluated by calculating the relative bias, estimated by the mean of the differences (d) and the standard deviation (SD) of the differences (Marini 2001).

Results

Preliminary analysis: The results of stepwise regression analysis indicated that among variables involved for the estimation of LE and T (L, W, and LW), the LW explained the biggest part of the variation. The variable for TLA estimation was ΣLW defeating ΣL and ΣW. Also, statistic test results revealed that pooled data of all cultivars showed normal distribution. For this reason, data were pooled and a general one-parameter model (LW or ΣLW) was calculated to develop LA prediction model for soybean.

As a preliminary step to produce and calibrate models, the degree of collinearity between W and L was analyzed by VIF and TV. In the present study, the VIF values ranged from 2.10 to 8.33, and TV values ranged from 0.12 to 0.48, depending on the cultivar. In all cultivars, VIF was <10, and TV was >0.10, showing that the collinearity between L and W can be considered as negligible, and these variables can be both included in the model.

Leaflet, trifoliate and total leaf area estimation models by leaf dimensions: L ranged from 2.20 to 15.30 cm and W ranged from 1.30 to 10.30 cm per across cultivars in both years (Table 3). The fitting model was done for each cultivar in both years, separately. Strong relationships were found between L or W with LE or T. The results of

Table 4. Coefficients of linear regression model (y = a + bx) fitted to data of measured *vs.* leaf area. (a) is constant and (b) is coefficient of the models to estimate the leaf area from product of length and width (LW), L and W were in cm, summed product of length and width per plant (ΣLW), green leaf dry matter per plant [g] (GLDM), the total number of green leaflet per plant (TNLP) in soybean. For all models, $P<0.0001$. n – the number of observation data. RMSE – root means square error, R^2 – coefficient of determination, and SE – standard error.

Area	Model No.	Cultivar	Variable	Exp. 1 n	a ± SE	b ± SE	R^2	RMSE [cm^2]	Exp. 2 n	a ± SE	b ± SE	R^2	RMSE [cm^2]
Leaflet area													
	(1)	DPX	LW	444	−1.39 ± 0.26	0.68 ± 0.0035	0.99	2.77	544	2.82 ± 0.17	0.61 ± 0.0037	0.98	1.67
	(2)	Sahar	LW	437	−0.38 ± 0.21	0.68 ± 0.0044	0.98	2.12	532	1.40 ± 0.16	0.64 ± 0.0033	0.99	1.72
	(3)	Williams	LW	328	−0.01 ± 0.25	0.69 ± 0.0042	0.99	2.20	542	3.42 ± 0.18	0.60 ± 0.0036	0.98	1.73
	(4)	Hill	LW	-	-	-	-	-	558	1.73 ± 0.17	0.61 ± 0.0038	0.98	1.98
	(5)	Pooled data	LW	1,029	−0.46 ± 0.14	0.68 ± 0.0022	0.99	2.47	2,176	2.17 ± 0.09	0.62 ± 0.0019	0.98	1.90
Trifoliate area													
	(6)	DPX	LW	444	−8.85 ± 1.07	2.02 ± 0.014	0.98	11.16	544	0.57 ± 0.79	1.88 ± 0.017	0.96	7.54
	(7)	Sahar	LW	437	−7.29 ± 0.73	1.97 ± 0.016	0.97	7.50	532	−3.76 ± 0.73	1.89 ± 0.015	0.97	7.89
	(8)	Williams	LW	328	−7.99 ± 0.99	1.98 ± 0.017	0.98	8.84	542	1.68 ± 0.97	1.72 ± 0.019	0.94	9.22
	(9)	Hill	LW	-	-	-	-	-	558	−0.10 ± 0.65	1.76 ± 0.014	0.96	7.33
	(10)	Pooled data	LW	1,029	−8.58 ± 0.51	2.00 ± 0.008	0.98	9.38	2,176	−0.44 ± 0.41	1.81 ± 0.009	0.95	8.50
Total leaf area in plant^{-1}													
	(11)	DPX	ΣLW	20	26.37 ± 41.73	1.83 ± 0.044	0.99	53.37	21	27.18 ± 20.24	1.84 ± 0.027	0.99	42.24
	(12)	Sahar	ΣLW	35	−3.72 ± 28.69	1.83 ± 0.038	0.99	65.50	31	27.20 ± 24.49	1.78 ± 0.031	0.99	51.46
	(13)	Williams	ΣLW	25	−71.42 ± 28.79	1.95 ± 0.043	0.99	40.33	39	−2.90 ± 27.44	1.79 ± 0.037	0.98	66.66
	(14)	Hill	ΣLW	-	-	-	-	-	33	9.14 ± 20.53	1.74 ± 0.037	0.99	44.40
	(15)	Pooled data	ΣLW	80	−11.38 ± 17.00	1.86 ± 0.022	0.99	56.33	124	12.61 ± 11.67	1.79 ± 0.015	0.99	64.56
Total leaf area in plant^{-1}													
	(16)	DPX	GLDM	47	160.79 ± 66.14	216.87 ± 10.44	0.91	232.47	83	178.87 ± 38.82	130.01 ± 3.43	0.95	175.27
	(17)	Sahar	GLDM	51	128.17 ± 52.13	232.30 ± 10.35	0.91	183.94	65	143.87 ± 55.47	160.69 ± 6.79	0.90	167.09
	(18)	Williams	GLDM	48	139.45 ± 73.21	209.55 ± 15.14	0.81	283.07	69	76.09 ± 49.20	141.92 ± 6.36	0.88	135.84
	(19)	Hill	GLDM	-	-	-	-	-	81	38.89 ± 29.72	159.59 ± 3.54	0.96	123.57
	(20)	Pooled data	GLDM	146	143.64 ± 36.83	219.63 ± 6.79	0.88	236.34	298	163.07 ± 23.01	139.51 ± 2.52	0.91	182.65
Total leaf area in plant^{-1}													
	(21)	DPX	TNLP	26	186.07 ± 148.16	37.26 ± 3.91	0.79	227.27	83	−254.9 ± 101.2	31.51 ± 1.73	0.80	336.81
	(22)	Sahar	TNLP	35	−83.18 ± 114.79	28.59 ± 2.27	0.83	230.53	62	−276.7 ± 68.23	28.94 ± 1.22	0.88	224.40
	(23)	Williams	TNLP	-	-	-	-	-	73	98.51 ± 99.42	22.61 ± 1.75	0.73	247.65
	(24)	Hill	TNLP	-	-	-	-	-	81	−14.60 ± 94.01	26.98 ± 2.09	0.70	230.33
	(25)	Pooled data	TNLP	61	356.72 ± 134.13	24.10 ± 2.94	0.73	368.02	299	−135.7 ± 46.11	28.04 ± 0.85	0.79	289.90

Table 5. Statistics and estimated parameters from regression models for leaflet, trifoliate and total leaf area [cm^2] estimation, product of length and width (LW); L and W were in cm, summed product of length and width per plant (ΣLW), green leaf dry matter per plant [g] (GLDM), the total number of green leaflet per plant (TNLP). The estimation model (Esti. M.) was developed over data pooled from cultivars in 2009 and 2010. Validation models (Valida. M.) were developed from validation data set for the different level of the plant canopy of DPX cultivar sown in both years. To validate the estimation model, leaf samples were taken from the plants at the R1 to R6 stages in 2009 and 2010 (Exp. 1, 2009 and Exp. 2, 2010), n – the number of observation data.

Statistic or parameter estimate	Leaflet area (LW)			Trifoliate area (LW)			Total leaf area (ΣLW)			Total leaf area (GLDM)			Total leaf area (TNLP)		
	Esti. M.	Valida. M.		Esti. M.	Valida. M.		Esti. M.	Valida. M.		Esti. M.	Valida. M.		Esti. M.	Valida. M.	
	(1)	Exp. 1	Exp. 2	(2)	Exp. 1	Exp. 2	(3)	Exp. 1	Exp. 2	(4)	Exp. 1	Exp. 2	(5)	Exp. 1	Exp. 2
Intercept	0.754	–0.887	–1.16	–4.869	–8.149	–14.91	6.876	49.77	61.28	298.47	-	274.32	–5.612	-	–367.88
Standard error of intercept	0.079	0.575	0.292	0.320	2.247	1.561	9.978	39.03	28.29	26.97	-	79.68	47.29	-	85.30
Regression coefficient for LW	0.655	0.681	0.648	1.923	2.014	2.036	1.813	1.826	1.754	136.54	-	134.39	26.53	-	37.23
Standard error for LW	0.001	0.007	0.005	0.006	0.028	0.028	0.013	0.048	0.031	3.33	-	6.27	0.893	-	1.44
Variance inflation factor (VIF) for LW	3.57	8.33	2.56	3.57	8.33	2.56	-	-	-	-	-	-	-	-	-
Root mean square error (RMSE in cm^2)	2.261	2.985	1.836	9.165	11.665	9.827	64.369	37.684	76.036	291.56	-	265.28	320.16	-	277.83
Coefficient of determination R^2	0.983	0.989	0.987	0.968	0.981	0.965	0.988	0.991	0.993	0.791	-	0.895	0.711	-	0.894
Adjusted coefficient of determination R^2_a	0.983	0.989	0.987	0.968	0.981	0.965	0.988	0.991	0.993	0.791	-	0.895	0.711	-	0.894
$P>F$	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	-	<0.0001	<0.0001	-	<0.0001
n	3385	100	200	3385	100	200	244	15	25	444	-	56	360	-	81

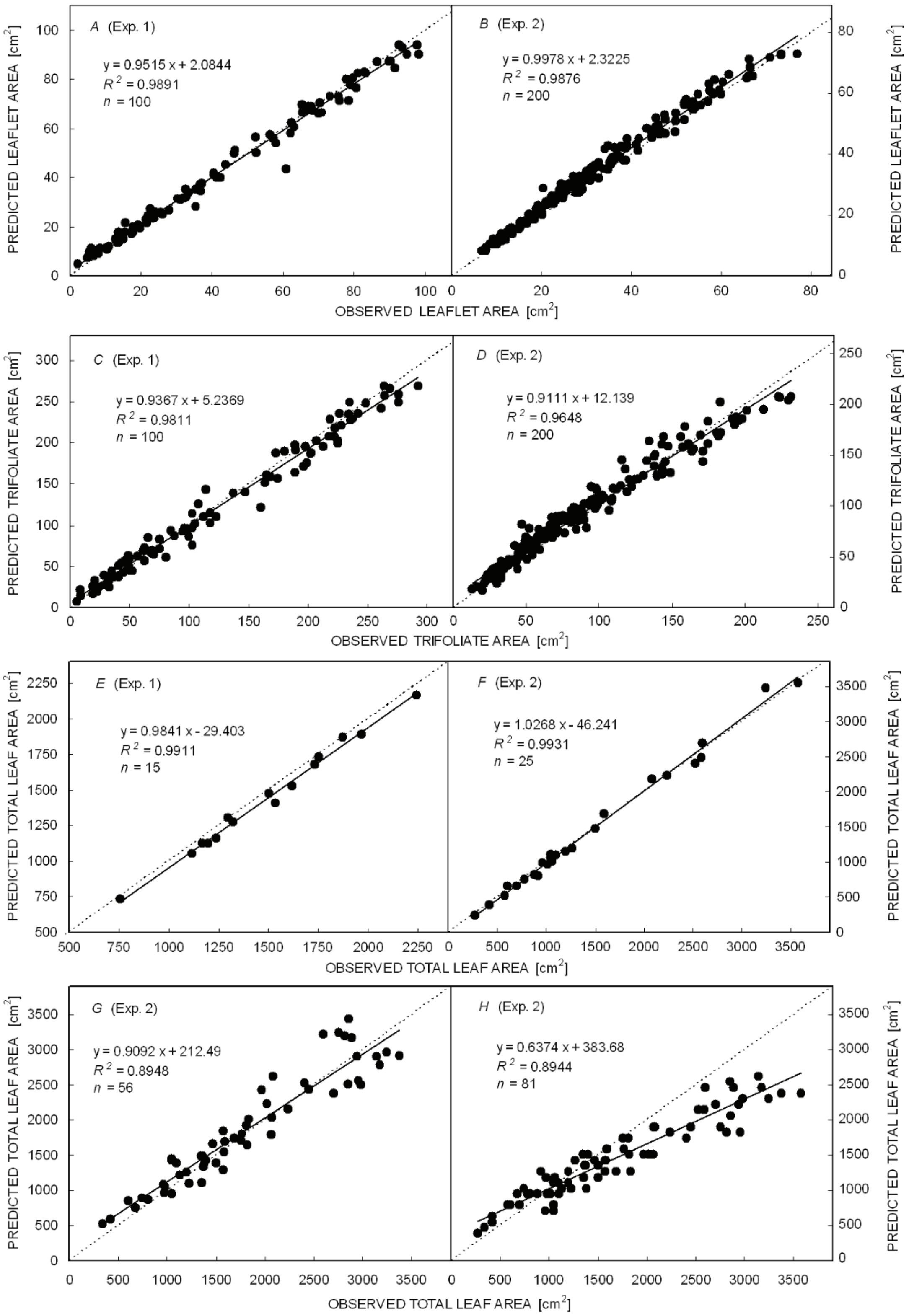

Fig. 1. Predicted leaf area (PLA) by the model of pooled data (all cultivars and in both years) *vs.* the observed leaf area (OLA) for cv. 'DPX'. Leaflet area (*A*, Exp. 1 and *B*, Exp. 2), trifoliate area (*C*, Exp. 1 and *D*, Exp. 2), total leaf area by ΣLW (*E*, Exp. 1 and *F*, Exp. 2), Total leaf area by GLDM (*G*, Exp. 2) and total leaf area by TNLP (*H*, Exp. 2) in the validation data set. Exp. 1 and Exp. 2 related to independent data set 2009 and 2010, respectively. *Solid line* represents linear regression lines of pooled data models. *Dotted lines* represent the 1:1 relationship between the predicted and observed values. Note differences in the observed and predicted total leaf area in each figure.

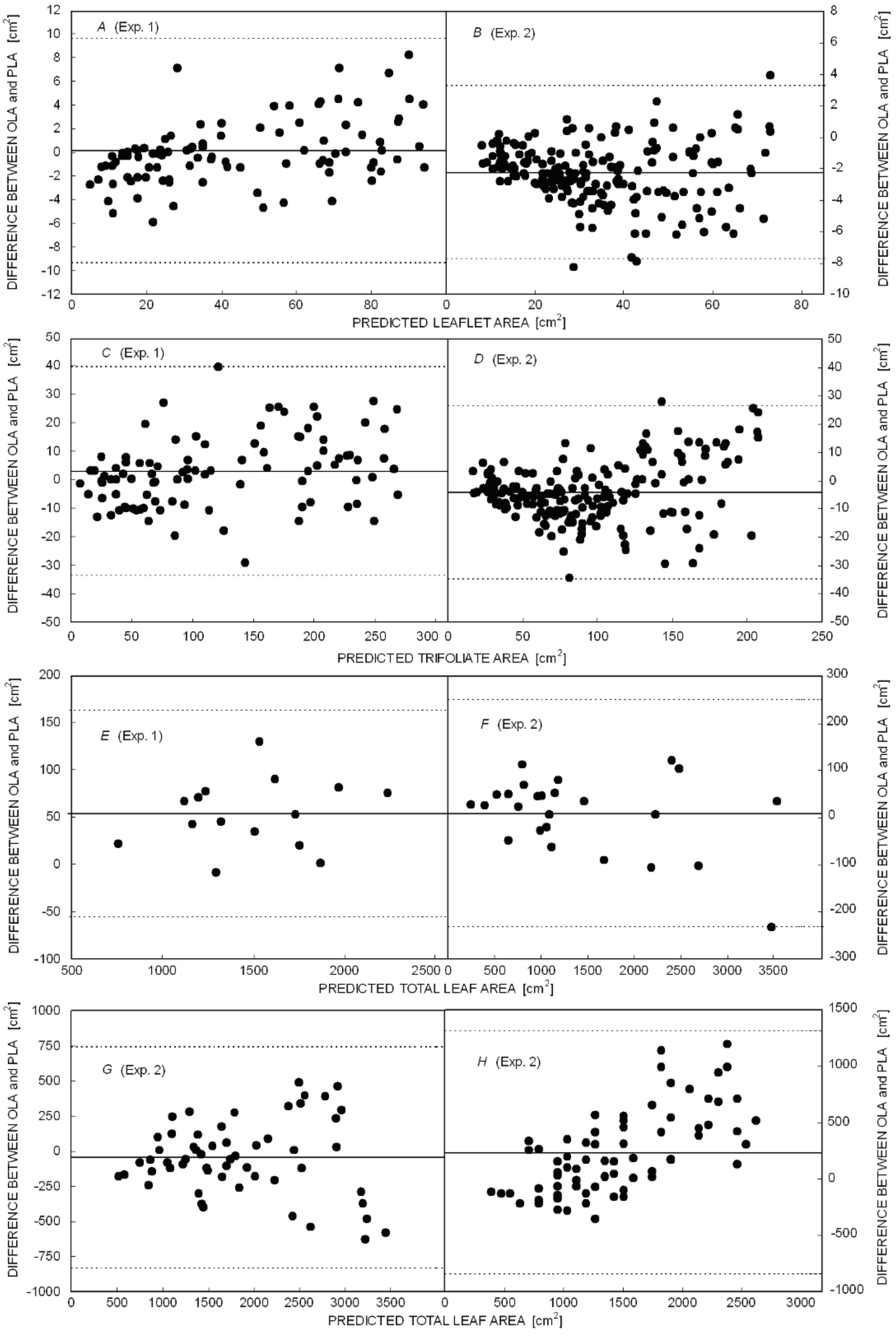

Fig. 2. The difference (d) between predicted leaf area (PLA) by the model of pooled data (all cultivars and in both years) and observed leaf area (OLA) of cv. 'DPX' *vs*. the PLA of DPX sampled in 2009 and 2010. The *solid line* is the mean of the differences. The *broken lines* are the limits of agreement, calculated as d ± 3SD; where d is the mean of the differences, and SD is the standard deviation of the differences. If the differences are normally distributed, 97% of the differences in a population will lie between the limits of agreement. Leaflet area (*A*, Exp. 1 and *B*, Exp. 2), trifoliate area (*C*, Exp. 1 and *D*, Exp. 2), total leaf area by ΣLW (*E*, Exp. 1 and *F*, Exp. 2), total leaf area by GLDM (*G*, Exp. 2) and total leaf area by TNLP (*H*, Exp. 2) in the validation data set. Exp. 1 and Exp. 2 related to independent data set 2009 and 2010, respectively.

the regression analysis demonstrated that models based on a single dimension (L or W) were less acceptable than the models based on LW (data not shown).

A strong, linear function between LW and LE or T accomplished the main criteria of a reliable, nondestructive LE and T prediction model for soybean. The R^2 values for the relationships between LE and LW, and T and LW were higher than 0.98 and 0.94, and RMSE values ranged from 1.67 to 2.77 cm^2, and from 7.50 to 11.16 cm^2 for all cultivars in both years, respectively. There was no significant difference between cultivars and years based on confidence intervals for the coefficients of the linear model (Table 4). Therefore, a general one-parameter model (LW-based) can be used for all cultivars in both years instead of individual models. The following linear models are proposed as more accurate (highest R^2 and smallest RMSE values) ones: LE = 0.754 + 0.655 LW, (R^2 = 0.98, model 1) and T = –4.869 + 1.923 LW, (R^2 = 0.97, model 2). The respective, proposed model for TLA prediction was the following: TLA = 6.876 + 1.813 ΣLW, (R^2 = 0.99, model 3) (Table 5).

Total leaf area estimation models by green leaf dry matter and total number of green leaflets per plant: In both experiments, GLDM, TNLP, and plant LA were measured simultaneously. GLDM, TNLP and plant LA ranged from 0.20 to 24.64 g $plant^{-1}$, 15 to 100 leaflets per plant and 258.5 to 3557 cm^2 $plant^{-1}$, respectively. There was a simple, linear relationship between LA with GLDM and TNLP.

After fitting simple, linear models for each cultivar in each year, results indicated that there was no significant difference between cultivars and years based on confidence intervals for the coefficients of the linear models (Table 4). Also, for the relationships between LA-GLDM and LA-TNLP, higher R^2 values from 0.81 and 0.70 were found, respectively (Table 4). Therefore, one general model can be used for all cultivars in both years instead of individual models. The proposed linear models were TLA = 298.47 + 136.54 GLDM (R^2 = 0.79, model 4) and TLA = –5.612 + 26.53 TNLP (R^2 = 0.71, model 5) (Table 5).

Model validation: Estimated parameters and statistics obtained from *SAS* outputs are presented for the LA estimation and model validation (Table 5). The results demonstrated that the intercept, the regression coefficients and RMSE values of the estimation and validation models were rather similar. Also, R^2 and R^2_a values were similar in all models (LA estimation and validation models) to one independent variable that indicates good accuracy of the models for LA prediction in soybean (Table 5). Moreover, comparisons between OLA *vs.* PLA were done using the best models (1 to 5 models, Table 5) for the validation data set derived from 2009 and 2010 experiments.

Results indicated a close correlation (r>0.94 and p<0.0001), for all the best models and the PLA values were close to the OLA values, with an exception for model 5, giving underestimation of the LA prediction, on the other hand, with increased plant ALA, predictied LA values were lower than the observed ones (Fig. 1).

Although, there was high correlation between OLA and PLA values, the correlation coefficient is insufficient to validate the relationship between PLA and OLA. Thus, a plot of the difference between PLA and OLA against PLA may be more informative (Bland and Altman 1986, Marini 2001). Plotting differences against PLA values allows also the investigation of possible relationships between measurement error and the true values. Lack of agreement between estimated PLA and OLA can be evaluated by calculating the bias, estimated by the mean of the differences (d) and the SD of the differences. In Fig. 2, a solid line represents the mean of the differences for the validation data set derived from both experiments. If the differences are normally distributed, 97% of the differences will be between d ± 3SD, which is the case in the current study, where a few plots were out of these lines while the rest of the plots were placed between lines.

Discussion

LA is an important variable for the most physiological and agronomical studies and it must be recorded for the effective monitoring of the growth and development of plant during the experimentation. Also, plant yield and

Table 6. Estimated and proposed models and linear functions, cofficient of determiation (R^2) and root mean square error (RMSE) for the calculated by models and measured leaf area DPX cultivars (all independent data cv. 'DPX', 2009 and 2010). For all functions, P<0.0001. L and W were in cm.

Source	Model	Linear function	R^2	RMSE [cm^2]
Wiersma and Bailey 1975	Leaflet area = 0.624 + 0.723 LW	y = 3.093 + 1.057 x	0.987	5.958
Wiersma and Bailey 1975	Trifoliate area = 0.411 + 2.008 LW	y = 16.27 + 0.958 x	0.975	16.10
Wiersma and Bailey 1975	Total leaf area = 6.532 + 2.045 ΣLW	y = –64.4 + 1.152 x	0.992	141.9
In this study	Leaflet area = 0.754 + 0.655 LW	y = 2.991 + 0.958 x	0.987	2.987
In this study	Trifoliate area = –4.869 + 1.923 LW	y = 10.32 + 0.917 x	0.975	11.55
In this study	Total leaf area = 6.876 + 1.813 ΣLW	y = –56.0 + 1.021 x	0.992	72.35

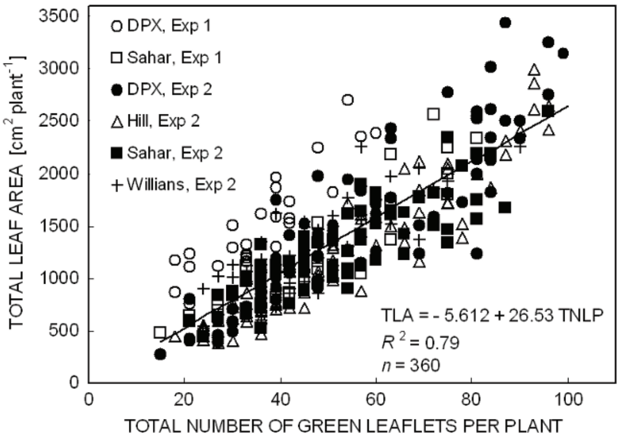

Fig. 3. Linear regression between the total leaf area (TLA) and the total number of green leaflets per plant (TNLP) for all data in both experiments). The slope of the linear regression model was 26.53 cm^2 that indicates a total average leaflet size for soybean cultivars.

quality are affected by photosynthesis and transpiration rate, which are closely related to plant TLA (Serdar and Demirsoy 2006, Peksen 2007).

The lack of accurate models (destructive and non-destructive) predicting LA is a limiting factor for the use of this parameter in physiological studies. Both destructive methods (measuring green leaf dry matter and total aboveground biomass dry matter) and nondestructive methods (measuring leaf width, length and number, branch length and number, and plant height) have been developed for the estimation of LA, both having advantages and disadvantages (Akram-Ghaderi and Soltani 2007). However, Lu *et al.* (2004) and Peksen (2007) reported that the establishment of mathematical and especially linear relationships between LA and leaf dimensions is an advantageous way to determine LA under field conditions.

The results of our study were consistent with results of Wiersma and Bailey (1975) who proposed that linear relationships between LA and one or more dimensions of the terminal leaflet on a soybean plant could be used to estimate LA, accurately. Therefore, in this study we tested accuracy of proposed models by Wiersma and Bailey (1975) for soybean cultivars cultivated in the northern Iran.

Table 6 presents the developed models in this study and the proposed models (Wiersma and Bailey 1975) for 12 soybean cultivars and the linear functions between LA calculated and measured for soybean plant in the northern Iran. All the relationships were highly significant (p<0.0001) with R^2 values higher than 0.97. On the other hand, R^2 values were closed for all the developed and proposed models (Table 6). A rather similar RMSE shows that a calculated LA is close to the measured one and thus RMSE should be the main criterion for selecting the LA prediction model when a precise estimation of the ALA is necessary. Consequently, the proposed models by Wiersma and Bailey (1975) are appropriate for the LA estimation of soybean in the northern Iran.

The results indicated that, close relationships were found between TLA and GLDM, and TLA and TNLP by the linear models (Table 4). Tsialtas and Maslaris (2008c) for sugar beet and Akram-Ghaderi and Soltani (2007) for cotton proposed quadratic and power models for TLA prediction, respectively. Sivakumar (1978) and Ogbuehi and Brandle (1981) reported a linear relationship between TNLP and the plant LA in soybean. Also, Birch *et al.* (1998) reported that in maize TLA can be derived from the total green leaf number. Hemmer *et al.* (1993) and Soltani *et al.* (2006) used a power function ($y = x^b$) to predict plant LA from the number of green leaves in grain sorghum and chickpea, respectively, where y is LA, x is plant green leaf number, and b is coefficient of power equation to estimate LA from plant green leaf number.

In this study, leaflet size was not measured on soybean plant separately. Therefore, average leaflet size was calculated by dividing green LA of the plant by TNLP. The slope of the linear regression model (model 5, Table 5) was 26.53 cm^2 per leaflet, indicating a total average leaflet size for all soybean cultivars in both years (Fig. 3). Quantitative evaluation of organ shape is often needed for many field researches in agronomy, genetics, ecology and taxonomy, due to its inherent character, leaf shape could be used for plant species identification or cultivars classification (Tsialtas and Maslaris 2007).

Conclusions: Results of this study indicated that there were very close relationships between OLA and PLA. Also, the validation models indicated that soybean LA could be measured quickly and accurately by using the developed models in this study (especially, models 1–3, Table 5). These models were chosen for their simplicity, generating results with the same level of accuracy as other more complex estimation models or expensive equipments. Because L and W are dimensions that can be easily measured in the field, greenhouse and pot experiments, using these models would enable researchers to make nondestructive measurements or repeated measurements on the same leaves. Such models can estimate accurately and in large quantities the LA of soybean in many experimental comparisons without the use of any expensive instrument.

References

Akram-Ghaderi, F., Soltani, A.: Leaf area relationships to plant vegetative characteristics in cotton (*Gossypium hirsutum* L.) grown in a temperate sub-humid environment. – Int. J. Plant Prod. **1**: 63-71, 2007.

Bange, M.P., Hammer, G.L., Milroy, S.P., Rickert, K.G.: Improving estimates of individual leaf area of sunflower. –

Agron. J. **92**: 761-765, 2000.
Beerling, D.J., Fry, J.C.: A comparison of the accuracy, variability and speed of five different methods for estimating leaf area. – Ann. Bot. **65**: 483-488, 1990.
Bhatt, M., Chanda, S.V.: Prediction of leaf area in (*Phaseolus vulgaris* L.) by non-destructive method. – Bulgarian J. Plant Physiol. **29**: 96-100, 2003.
Bignami, C., Rossini, F.: Image analysis estimation of leaf area index and plant size of young hazelnut plants. – Sci. Hort. **71**: 113-121, 1996.
Bland, J.M., Altman, D.G.: Statistical methods for assessing agreement between two methods of clinical measurements. – Lancet **1**: 307-310, 1986.
Birch, C.J., Hammer, G.L., Rickert, K.G.: Improved methods for predicting individual leaf area and leaf senescence in maize (*Zea mays* L.). – Aust. J. Agr. Res. **49**: 249-262, 1998.
Cankaya, S., Kayaalp, G.Y., Sangun, L., Tahtali, Y., Akar, M.: A comparative study of estimation methods for parameters in multiple linear regression model. – J. Appl. Anim. Res. **29**: 43-47, 2006.
Cristofori, V., Rouphael, Y., Mendoza-de Gyves, E., Bigniami, C.: A simple model for estimating leaf area of hazelnut from linear measurements. – Sci. Hort. **113**: 221-225, 2007.
Daughtry, C.: Direct measurements of canopy structure. – Remote Sens. Rev. **5**: 45-60: 1990.
de Jesus, W.C., Do Vale, F.X.R., Coelho, R.R., Costa, L.C.: Comparison of two methods for estimating leaf area index on common bean. – Agron. J. **93**: 989-991, 2001.
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. – Hort. Sci. Biotechnol. **79**: 764-770, 2004.
Fallovo, C., Cristofori, V., Mendoza-de Gyves, E., *et al*.: Leaf area estimation model for small fruits from linear measurements. – Hort. Sci. **43**: 2263-2267, 2008.
Fehr, W.R., Caviness, C.E.: Stage of Soybean Development, – Iowa State University Cooperative Extension Service Special Report 80, Iowa City 1977.
Gamiely, S., Randle, W.M., Milks, W.A., Smittle, D.A.: A rapid and nondestructive method for estimating leaf area of onions. – Hort. Sci. **26**: 206-212, 1991.
Gill, J.L.: Outliers, residuals, and influence in multiple regressions. – J. Anim. Breed. Genet. **103**: 161-175, 1986.
Hemmer, G.L., Carberry, P.S., Muchow, R.C.: Modeling genotype and environmental control of leaf area dynamics in grain sorghum. I. Whole plant level. – Field Crops Res. **33**: 293-310, 1993.
Jonckheere, I., Fleck, S., Nackaerts, K., Muys, B., Coppin, P., Weiss, M., Baret, F.: Review of methods for in situ leaf area index determination. I: Theories, sensors and hemispherical photography. – Agr. Forest Meteorol. **121**: 19-35, 2004.
Kandiannan, K., Parthasarathy, U., Krishnamurthy, K.S., Thankamani, C.K., Srinivasan, V.: Modeling individual leaf area of ginger (*Zingiber officinale* Roscoe) using leaf length and width. – Sci. Hort. **120**: 532-537, 2009.
Kathirvelan, P., Kalaiselvan, P.: Peanut (*Arachis hypogaea* L.) leaf area estimation using allometric model. – Res. J. Agr. Biol. Sci. **3**: 59-61, 2007.
Kumar, R.: Calibration and validation of regression model for non-destructive leaf area estimation of saffron (*Crocus sativus* L.). – Sci. Hort. **122**: 142-145, 2009.
Lieth, J.H., James, F., Reynolds, H., Rogers, H.: Estimation of leaf area of soybeans grown under elevated carbon dioxide levels. – Field Crops Res. **13**: 193-203, 1986.
Lu, H.Y., Lu, C.T., Wei, M.L., Chan, L.F.: Comparison of different models for nondestructive leaf area estimation in taro. – Agron. J. **96**: 448-453, 2004.
Ma, L., Gardener, F.P., Selamat, A.: Estimation of leaf area from leaf and total mass measurements in peanut. – Crop Sci. **32**: 461-471, 1992.
Marini, R.P.: Estimating mean fruit weight and mean fruit value for apple trees: comparison of two sampling methods with the true mean. – J. Am. Soc. Hort. Sci. **126**: 503-510, 2001.
Marquardt, D.W.: Generalized inverse, ridge regression, biased linear estimation, and nonlinear estimation. – Technometrics **12**: 591-612, 1970.
Miranda, C., Royo, J.B.: A statistical model to estimate potential yields in peach before bloom. – J. Am. Soc. Hort. Sci. **128**: 297-301, 2003.
Miranda, C., Royo, J.B.: Statistical model estimates potential yield in ''*Golden Delicious*'' and ''*Royal Gala*'' apples before bloom. – J. Am. Soc. Hort. Sci. **129**: 20-25, 2004.
Mokhtarpour, H., Teh, C.B.S., Saleh, G., Selmat, A.B., Asadi, M.E., Kamar, B.: Non-destructive estimation of maize leaf area, fresh weight, and dry weight using leaf length and leaf width. – Commun. Biom. Crop Sci. **5**: 19–26, 2010.
Nesmith, D.S.: Non-destructive leaf area estimation of rabbiteye blueberries. – Hort. Sci. **26**: 13-32, 1991.
Neter, J., Kutner, M.H., Nachtshein, C.J., Wasserman, W.: Applied Linear Regression Models. 3rd Ed. Homewood III – Irwin, 1996.
Nyakwende, E., Paull, C.J., Atherton, J.G.: Non-destructive determination of leaf area in tomato plants using image processing. – J. Hort. Sci. **72**: 225-262, 1997.
Ogbuehi, S.N., Brandle, J.R.: Limitations in the use of dry weight and leaf number for predicting leaf area of soybeans. – Crop Sci. **21**: 344-346, 1981.
Payne, W.A., Went, C.W., Hossner, L.R., Gates, C.E.: Estimating pearl millet leaf area and specific leaf area. – Agron. J. **83**: 937-941, 1991.
Peksen, E.: Non-destructive leaf area estimation model for faba bean (*Vicia faba* L.). – Sci. Hort. **113**: 322-328, 2007.
Ramos, J.M., Delmoral, L.F.G., Recalde, L.: Dry matter and leaf area relationship in winter barley. – Agron. J. **75**: 308-310, 1983.
Rico-García, E., Hernández-Hernández, F., Soto-Zarazúa, G.M., Herrera-Ruiz, G.: Two new methods for the estimation of leaf area using digital photography. – Int. J. Agr. Biol. **11**: 397-400, 2009.
Rouphael, Y., Colla, G., Fanasca, S., Karam, F.: Leaf area estimation of sunflower leaves from simple linear measurements. – Photosynthetica **45**: 306-308, 2007.
Rouphael, Y., Mouneimne, A.H., Ismail, A., Mendoza, E., Rivera, C.M., Colla, G.: Modeling individual leaf area of rose (*Rosa hybrida* L.) based on leaf length and width measurement. – Photosynthetica **48**: 9-15, 2010.
SAS Institute: SAS/STAT user's guide. – SAS Institute Inc., Cary 1992.
Serdar, U., Demirsoy, H.: Non-destructive leaf area estimation in chestnut. – Sci. Hort. **108**: 227-230, 2006.
Setiyono, T.D., Weiss, A., Specht, J.K., Cassman, K.G., Dobermann, A.: Leaf area index simulation in soybean grown under near-optimal conditions. – Field Crops Res. **108**: 82-92, 2008.
Shih, S.F., Gascho, G.J., Rahi, G.S.: Modeling biomass produc-

tion of sweet sorghum. – Agron. J. **73**: 1027-1032, 1981.
Shin, S.F., Snyder, G.H.: Leaf area index and dry biomass of taro. – Agron. J. **76**: 750-753, 1984.
Sivakumar, M.V.K.: Prediction of leaf area index in soybean. – Ann. Bot. **42**: 251-253, 1978.
Soltani, A., Hoogenboom, G.: A statistical comparison of stochastic weather generators WGEN and SIMMETEO. – Climate Res. **24**: 215-230, 2003.
Soltani, A., Robertson, M.J., Mohammad-Nejad, Y., Rahemi-Karizaki, A.: Modeling chickpea growth and development: Leaf prediction and senescence. – Field Crops Res. **138**: 14-23, 2006.
Stoppani, M.I., Wolf, R., Francescangeli, N., Martí, H.R.: A non-destructive and rapid method for estimating leaf area of broccoli. – Adv. Hort. Sci. **17**: 173-175, 2003.
Tsialtas, J.T., Maslaris, N.: Leaf area estimation in a sugar beet cultivar by linear model. – Photosynthetica **43**: 477-479, 2005.
Tsialtas, J.T., Maslaris, N.: Leaf shape and its relationship with Leaf Area Index in a sugar beet (*Beta vulgaris* L.) cultivar. – Photosynthetica **45**: 527-532, 2007.
Tsialtas, J.T., Maslaris, N.: Evaluation of a leaf area prediction model proposed for sunflower. – Photosynthetica **46**: 294-297, 2008a.
Tsialtas, J.T., Maslaris, N.: Leaf area prediction model for sugar beet (*Beta vulgaris* L.) cultivars. – Photosynthetica **46**: 291-293, 2008b.
Tsialtas, J.T., Maslaris, N.: Leaf allometry and prediction of specific leaf area (SLA) in a sugar beet (*Beta vulgaris* L.) cultivar. – Photosynthetica **46**: 351-355, 2008c.
Tsialtas, J.T., Koundouras, S., Zioziou, E.: Leaf area estimation by simple measurements and evaluation of leaf area prediction models in Cabernet-Sauvignon grapevine leaves. – Photosynthetica **46**: 452-456, 2008.
Villegas, C.D., Bautista, A.T., Cotejo-Jr., F.R.: Accurate and rapid techniques for leaf area measurement in cassava and sweet potato. – Radix **3**: 10-15, 1981.
Wiersma, J.V., Bailey, T.B.: Estimation of leaflet, trifoliate and total leaf area of soybean. – Agron. J. **67**: 26-30, 1975.