Research Article

A Method for Calculating the Area of *Zostera marina* Leaves from Digital Images with Noise Induced by Humidity Content

Cecilia Leal-Ramirez and Hector Echavarria-Heras

Centro de Investigación Científica y Educación Superior de Ensenada, Carretera Ensenada-Tijuana No. 3918, Zona Playitas, 22860 Ensenada, BC, Mexico

Correspondence should be addressed to Cecilia Leal-Ramirez; cleal@cicese.mx

Despite the ecological importance of eelgrass, nowadays anthropogenic influences have produced deleterious effects in many meadows worldwide. Transplantation plots are commonly used as a feasible remediation scheme. The characterization of eelgrass biomass and its dynamics is an important input for the assessment of the overall status of both natural and transplanted populations. Particularly, in restoration plots it is desirable to obtain nondestructive assessments of these variables. Allometric models allow the expression of above ground biomass and productivity of eelgrass in terms of leaf area, which provides cost effective and nondestructive assessments. Leaf area in eelgrass can be conveniently obtained by the product of associated length and width. Although these variables can be directly measured on most sampled leaves, digital image methods could be adapted in order to simplify measurements. Nonetheless, since width to length ratios in eelgrass leaves could be even negligible, noise induced by leaf humidity content could produce misidentification of pixels along the peripheral contour of leaves images. In this paper, we present a procedure aimed to produce consistent estimations of eelgrass leaf area in the presence of the aforementioned noise effects. Our results show that digital image procedures can provide reliable, nondestructive estimations of eelgrass leaf area.

1. Introduction

Zostera marina also known as eelgrass is a relevant seagrass species, which supplies significant amounts of organic materials to food webs in shallow coastal environments and provides habitat (in bays, lagoons, or estuaries) for many fishes and their larvae [1]. Eelgrass beds can also help remediate contaminated sediments [2], filter and retain nutrients from the water column [3], help in the stabilization of sediments [4], and reduce erosion forces by stumping wave energy, thus promoting the stabilization of adjacent shorelines [5]. However, the permanence of eelgrass beds—as well as those formed by other seagrass species—is currently being threatened by anthropogenic influences to such an extent that special conservation efforts are needed [6]. This requires the development of accurate and cost-effective procedures aimed at obtaining scientific knowledge about the pertinent growth dynamics. This is particularly relevant in assessments of restoration projects, where the use of

noninvasive data gathering techniques turns out to be of fundamental importance.

The characterization of eelgrass biomass and its dynamics is an important input for the assessment of the overall status of both natural and transplanted eelgrass populations. In eelgrass the basic unit for studying biomass and its production is the shoot, which includes sheaths, leaves, rhizomes, and roots. Biomass consists of an aboveground component formed by sheaths and leaves and a belowground constituent formed by rhizomes and roots. Root emergence occurs at leaf scars, also known as rhizome nodes. The production of leaves and rhizome nodes is connected such that each leaf produced is linked to a rhizome node. Hence, the overall production of shoots can be estimated by measuring the production of leaves [7]; this makes us know the growth rate of leaves fundamental to the assessment of eelgrass populations [8]. Moreover, estimations of leaf biomass and leaf-growth rates are keys to assessing the reestablishment of ecological functioning in restored areas. Nevertheless, traditional methods

for the estimation of eelgrass leaf biomass and the related leaf growth rates are destructive and time consuming. Even though these procedures do not damage natural seagrass populations, they could produce undesirable effects on transplant experiments. Favorably, the conspicuous growth form of eelgrass makes it possible to introduce proxies that allow assessments while avoiding invasive interference. Moreover, estimations of leaf biomass or productivity in eelgrass can be efficiently obtained using allometric alternatives, which state these variables in terms of leaf length or area [9, 10]. But even though leaf architecture in eelgrass makes length a consistent descriptor of area, allometric models that express leaf biomass in terms of linked area perform relatively better than those involving leaf length as an independent variable. Therefore, for consistent allometric estimations of leaf biomass or productivity of eelgrass it is convenient to produce reliable estimations of leaf area. The observed ribbon-like appearance of the leaves in *Zostera marina* is a feature that permits obtaining direct and fairly accurate estimations of blade length l and width h. These variables provide convenient estimations of the corresponding blade area a, through the leaf length times width proxy [1]. If we used the symbol o as a subscript to represent observed values for the above named variables, then estimations of leaf area obtained through this proxy are given by

$$a_o = l_o \cdot h_o \tag{1}$$

which combined with allometric methods could simplify assessments of eelgrass leaf biomass and productivity [9, 10].

Digital image processing techniques were initially aimed to calculate the area of leaves for terrestrial plants [11–13]. These methods provide simplified estimations of biologically relevant variables. For example, Patil and Bodh [14] used area of *sugarcane* leaves for plant growth monitoring to analyze manure scarcity and environmental stress and to assess disease severity. Lü et al. [15] used leaf area measurement to assess long-term influences on yield and because it is a fundamental index in crop growth and nurturing practice. Although, leaf area in eelgrass can be conveniently obtained by means of (1), and both l_o and h_o can be directly measured on most sampled leaves, methods based on digital image processing could be adapted in order to simplify these tasks. Moreover, eelgrass leaf area can be directly estimated from digital imagery by using the Monte Carlo method [1]: if we let a_{mc} denote these estimations, then they are obtained through

$$a_{mc} = \frac{\text{LPN}}{\text{UPN}^2}, \tag{2}$$

where LPN is the number of points placed inside the considered leaf area and UPN^2 stands for the number of points contained in a unit area.

Besides, the Monte Carlo method eelgrass leaf area could be also obtained from digital images by using the length times width proxy of (1). Indeed, if l_d and h_d, respectively, denote leaf length and width obtained from the associated digital image, then these variables can be estimated through

$$l_d = \frac{\text{np}_l}{\text{unp}}, \tag{3}$$

$$h_d = \frac{\text{np}_h}{\text{unp}}, \tag{4}$$

where np_l and np_h are, respectively, obtained by counting the number of points contained over the length and width dimensions of the leaf, and unp is the number of points contained in the appropriate distance measurement unit. Therefore, denoting by means of a_d the associated leaf area, we will correspondingly have

$$a_d = l_d \cdot h_d. \tag{5}$$

Nevertheless, when using either (2) or (5) to produce estimations of leaf area we must be aware that some *Zostera marina* leaves could be very long or present curvatures, among irregularities caused by environmental factors like grazing or drag forces. The influence of these factors could affect image quality, which could produce biased estimations for leaf area. These effects have been partially addressed by Ramfos et al. [16], who proposed a method based on image processing techniques for measurements of a *Zostera marina* leaf by taking into account the effects of curvature on accuracy. Yet another important factor which we address here concerns the effects that the humidity contents of a leaf can originate in image processing. In fact, once leaves are removed from a shoot they begin to lose water and degrade. Hence, if leaves cannot be processed immediately after being collected, it is important to keep them in a manner that reduces changes in shape [17]. Therefore, an efficient digitalizing of a *Zostera marina* blade requires maintenance of an optimal humidity for increased image fidelity. On the other hand, humidity contents in a leaf can induce noise to an image by adding extraneous information, which usually manifests by pixel value misidentification.

Data published by Echavarria-Heras et al. [1], taken over a comprehensive sampling experiment, show that measured maximum width for a *Zostera marina* leaf is 6 mm. Surely, other authors report a variation range from 1.5 to 12 mm for this estimation [18]. A wide variation range in width in conjunction with noise due to humidity content can increase uncertainty in blade width measurements obtained from digitalized leaf images. This makes it necessary to devise a way that allows discriminating the concomitant error spreading over leaf area assessments. So far, an approach that integrates among others techniques, one aimed to handle noise effects induced by the humidity contents on a *Zostera marina* leaf, has not been produced. In this study, we conceived a method which using criteria based on statistical analysis techniques reduces the effects that noise linked to the humidity contents of a *Zostera marina* leaf produces on the accuracy of associated area estimations obtained from a digital image.

2. Conceptual Framework for Image Processing

Our arrangement depends in a fundamental way on the concept of the peripheral or bordering contour of a bidimensional enclosure or domain. Several definitions of peripheral contour exist, being each one appropriate for different settings. Our interpretation is similar to the perimeter definition of a regular pattern in geometry. More accurately, a peripheral contour in the present settings will be defined as the sequence of boundary pixels of a digitalized eelgrass leaf. Moreover, for a reasonable identification of the area of the pertinent blade it is imperative that in the extents of the corresponding image minimal changes of color levels occur, even though around its outer contour abrupt changes of a color levels could be shown. The effect projected by humidity adds to leaf area pixels placed between the pixels captured by the digital image of the leaf itself and others belonging to its background. Hence, a reliable imbedding of the area of a leaf into an image requires the unambiguous identification of the pixels on its surrounding contour.

Our design is aimed to the aforesaid identification in the presence of noise due to the humidity content in the leaf. For the incumbent characterization, our system uses a quantitative setup developed on the basis of the concepts of adjacency, vicinity, connectivity, and tolerance of similarity between pixels. We briefly describe these notions in what follows.

Two pixels are adjacent if and only if they share one of their borders, or at least one of their corners. Two pixels are neighbors if they fulfill the definition of adjacency. Formally, the vicinity $V_p(x, y)$ of the point $P(x, y)$ is defined through

$$V_p(x, y) = (x + 1, y), (x - 1, y), (x, y + 1), (x, y - 1),$$
$$(x + 1, y + 1), (x + 1, y - 1), (x - 1, y + 1),$$
$$(x - 1, y - 1).$$
$$(6)$$

Without loss of generality, we explain the notion of tolerance of similarity by referring to the RGB format description of a color. This allows quantifying tonality in terms of the intensities of the constituting primary colors: red, green, and blue. To indicate at which amount each one of these colors is mixed to produce a given tonality a value is assigned to each prime color; for example, the value 0 means that a given basic color does not appear in the mix, but if a chief color component is nonvanishing it means that it contributes to the mix in a given intensity. We set C_{max} which identifies the colors number to be used through the whole image processing task; for an RGB color space we have $C_{max} = 256$. Usually, the intensity of each of the primary colors appearing in a mix is measured on a scale ranging from 0 to $C_{max} - 1$. The set of all color intensities can be represented in the form of a cube in the cartesian coordinate system, where each color is a point on the surface or in its interior. Given points $P = (p_1, p_2, \ldots, p_n)$ and $Q = (q_1, q_2, \ldots, q_n)$ in an RGB color space, we will define the distance $d_E(P, Q)$ between them through

$$d_E(P, Q) = \sqrt{\sum_{i=1}^{n} (p_n - q_n)^2}. \qquad (7)$$

Moreover, given a point P in an RGB color space, a second one Q with the greatest similarity to P is the one placed at the smallest distance $d_E(P, Q)$. Furthermore, let $ST(x) = [0, x]$ be a color tonality range, with x being the number of different colors included. Then, we must have $1 \leq x \leq C_{max} - 1$ and we will say that two pixels P and Q are similar to a tolerance limit $ST(x)$ if the inequality

$$d_E(P, Q) \leq x \qquad (8)$$

is satisfied. In what follows the range $ST(x)$ will be simply called "tolerance of similarity" and the upper bound x can be interpreted as the maximum distance that two points located within the extent of an object can attain in an RGB color space in order to be considered similar. Connectivity between pixels is used to identify the limits in objects and regions in an image. We will say that two pixels P and Q are connected with tolerance of similarity $ST(x)$ if they fulfill the definition of adjacency and also if inequality (8) holds.

3. The Image Selection Method

The procedure to obtain efficient estimations of l_d, h_d, and a_{mc} requires two stages. On a first one we create a digital image for each one of the collected leaves. Then, we set C_{max} and continue by choosing an interval of tolerance of similarity $ST(x)$ with $0 \leq x \leq C_{max} - 1$; we use this to obtain the peripheral contour of each one of the available leaf images and from them the linked l_d, h_d, and a_{mc} values. Different intervals $ST(x)$ will produce different estimations for l_d, h_d, and a_{mc}, and consequently we must rely on a criterion for the selection of the $ST(x)$ range that produces the most accurate estimations. To carry out this task in a second stage of the method we arrange leaf length data into groups of leaves whose size differences are bounded by a preferred tolerance q and use that arrangement to obtain related statistics β_a and λ_a that are used to implement what we call the IS_x selection index. In what follows we describe pseudo-codes for the above referred stages. Detailed formulae are presented in the appendices. Tables 1, 2, 3, 4, and 5 summarize the involved notation.

3.1. The Procedure to Obtain l_d, h_d, and a_{mc} Assessments

(a.1) Choose a color format and set C_{max}.

(a.2) Load the leaf image.

(a.3) Enter an interval of tolerance of similarity $ST(x)$; $1 \leq x \leq C_{max} - 1$.

(a.4) Select a starting point inside the loaded leaf image.

(a.5) Find the contour of the leaf image through (6), (7), and (8) (these equations identify all adjacent pixels

TABLE 1: Different symbols used in the digital image processing method.

Symbol	Description
l	Leaf length
h	Leaf width
C_{max}	Number of colors in a format of a digital image
$ST(x) = [0, x]$	Interval of Tolerance of Similarity
l_{max}	Maximum observed leaf length
$q = \dfrac{l_{max}}{n}$	Norm of the partition for the interval $[0, l_{max}]$
I_k	Partition interval of the form $[q(k-1), qk)$ for $0 \le k \le n$
$P_0^{l_{max}} = \bigcup_1^n (I_k)$	Collection of n intervals I_k that cover $[0, l_{max}]$
$G_k(l)$	Group of leaves whose lenghts (l) lie in I_k
n_k	Number of leaves in the group $G_k(l)$
$C_G = \bigcup_1^n G_k(l)$	Collection of all groups $G_k(l)$ of leaves

TABLE 2: Symbols for observed, digitally obtained variables and related averages.

Description	Observed data	Digital data	Monte Carlo data
Leaf length (l)	l_o	l_d	—
Leaf width (h)	h_o	h_d	—
Leaf area (a)	a_o	a_d	a_{mc}
Length of the jth leaf in group $G_k(l)$	l_{oj}^k	l_{dj}^k	—
Width of the jth leaf in group $G_k(l)$	h_{oj}^k	h_{dj}^k	—
Area of the jth leaf in group $G_k(l)$	a_{oj}^k	a_{dj}^k	a_{mcj}^k
Average length of the leaves in group $G_k(l)$	$\bar{l}_o^k$	$\bar{l}_d^k$	—
Average width of the leaves in group $G_k(l)$	$\bar{h}_o^k$	$\bar{h}_d^k$	—
Average area of the leaves in group $G_k(l)$	$\bar{a}_o^k$	$\bar{a}_d^k$	$\bar{a}_{mc}^k$

falling within the selected interval of tolerance of similarity $ST(x)$).

(a.6) Obtain l_d, h_d, and a_d by using (3), (4), and (5), respectively.

(a.7) Obtain a_{mc} by using (2).

(a.8) Record $ST(x)$ and the associated l_d, h_d, a_d, and a_{mc} estimations.

(a.9) Repeat steps 2–8 for each one of the available leaf images.

(a.10) Change the $ST(x)$ interval and jump to step (a.3).

Different $ST(x)$ intervals will produce through the above procedure different estimations for l_d, h_d, a_d, and a_{mc}. We now outline a procedure for the selection of the image that produces the most accurate estimations a_d or a_{mc} for the observed leaves area a_o. This requires the identification of

the interval of tolerance of similarity $ST(x)$ that yields the smallest values of the selection index IS_x defined by (9) below.

3.2. The Method for the Selection of an Optimal $ST(x)$ Interval

(b.1) For the entered $ST(x)$ interval, use (E.9) to calculate λ_a (this value gives the proportion of leaves for which a_d produces consistent estimations of a_o).

(b.2) For the entered $ST(x)$ interval, use (E.10) to calculate β_a (this value yields the proportion of leaves for which a_d overestimates observed leaf area a_o).

(b.3) For the entered $ST(x)$ interval, calculate the value of the image selection index IS_x according to

$$IS_x = \frac{\beta_a}{\lambda_a}. \tag{9}$$

(b.4) Record both $ST(x)$ and IS_x.

(b.5) Change the $ST(x)$ interval and repeat steps (b.1) to (b.3) until all the $ST(x)$ intervals generated in Section 3.1 are exhausted.

(b.6) Choose the $ST(x)$ interval that produces the smallest value of IS_x for image processing and leaf area a_d estimations.

The above selection index IS_x criterion can be adapted for Monte Carlo method estimations of leaf area. It becomes

$$IS_{xmc} = \frac{\beta_{amc}}{\lambda_{amc}}, \tag{10}$$

where λ_{amc} and β_{amc} are, respectively, given by (E.11) and (E.12) in Appendix E and are equivalent to λ_a and β_a correspondingly.

4. Results

4.1. Leaf Data Grouping. The present data set was obtained by randomly sampling 5 shoots biweekly from January through December 2009 in a *Zostera marina* field at Punta Banda estuary, a shallow coastal lagoon located near Ensenada, Baja California, Mexico (31° 43–46 N and 116° 37–40 W). For each sampled leaf, a millimeter ruler was used to obtain l_o to the nearest 1/10 mm taken as the distance from the top of the sheath to the leaf tip. Meanwhile, h_o was measured at a point halfway between the top of the sheath and the tip. Observed leaf area estimations a_o were calculated by means of (1).

We obtained $l_{max} = 460$ mm. For data grouping we selected $n = 46$, so we acquired $q = 10$ mm and for the interval $[0, l_{max}]$ we formed a partition P_0^{460} of disjoint intervals I_k of the form $I_k = \{l \mid q(k-1) \le l < qk\}$, with $1 \le k \le 46$. Hence, as described in the appendices for each value of the index k, we formed a group $G_k(l)$ containing leaves with sizes varying in the interval I_k (Table 6). Longer and older leaves displayed darker tonalities than younger and shorter ones. Moreover, leaves with lengths varying on a given partition interval I_k displayed a similar color distribution. For some of the partition intervals there was at most one leaf with

TABLE 3: Approximation errors.

Symbols	Formal expression	Description
e_{lj}^k	$l_{oj}^k - l_{dj}^k$	Difference of observed and image obtained leaf lengths in group $G_k(l)$.
e_{hj}^k	$h_{oj}^k - h_{dj}^k$	Difference of observed and image obtained leaf widths in group $G_k(l)$.
e_{aj}^k	$a_{oj}^k - a_{dj}^k$	Difference of observed and image obtained leaf areas in group $G_k(l)$.
e_{mcj}^k	$a_{oj}^k - a_{mcj}^k$	Difference of observed and Monte Carlo estimated leaf areas in group $G_k(l)$.

TABLE 4: Estimation errors for observed and image obtained variables averages and standard deviations.

Symbol	Formal expression	Description
δ_l^k	$\dfrac{\sum_1^{n_k} e_{lj}^k}{n_k}$	Leaf length average deviation in the group $G_k(l)$.
δ_h^k	$\dfrac{\sum_1^{n_k} e_{hj}^k}{n_k}$	Leaf width average deviation in the group $G_k(l)$.
δ_a^k	$\dfrac{\sum_1^{n_k} e_{aj}^k}{n_k}$	Leaf area average deviation in the group $G_k(l)$.
$\overline{\delta}_l$	$\dfrac{\sum_1^{n} \delta_l^k}{n}$	Leaf length average deviation in C_G.
$\overline{\delta}_h$	$\dfrac{\sum_1^{n} \delta_h^k}{n}$	Leaf width average deviation in C_G.
$\sigma_{\delta l}$	$\left(\dfrac{\sum_1^{n_k} \left(e_{lj}^k - \delta_l^k \right)^2}{(n_k - 1)} \right)^{1/2}$	Standard deviation of e_{lj}^k.
$\sigma_{\delta h}$	$\left(\dfrac{\sum_1^{n_k} \left(e_{hj}^k - \delta_h^k \right)^2}{(n_k - 1)} \right)^{1/2}$	Standard deviation of e_{hj}^k.

length placed in the linked variation range. Therefore, these groups are not taken into account because they do not provide information for the statistical analysis (see bold in Table 6).

4.2. Image Selection Procedure. For each one of the leaves in the collection C_G of groups $G_k(l)$, we applied the procedure described in the pseudo-code (a) aimed to detect the points on the associated peripheral contour and to get the concomitant l_d, h_d, a_d, and a_{mc} estimations. For that purpose we used a variety of equivalence of tones, which permitted an unambiguous framing of the extent of the leaf. A RGB 256 color format was used for all leaves images. Hence, we set $C_{\max} = 256$. Therefore, different tolerances of similarity ranges, $ST(x) = [0, x]$ with $1 \leq x \leq 255$, were used. Moreover, the procedure was automatically applied up to 256 times on each individual leaf image. For every tolerance of similarity interval $ST(x)$ we selected a starting point inside a chosen leaf image and we identified all adjacent pixels falling within the named similarity range. This identified the peripheral contour of the leaf image so the linked l_{dj}^k, h_{dj}^k, and a_{dj}^k assessments as well as leaf area estimations a_{mcj}^k were acquired (see B.3 in Appendix B).

For each group $G_k(l)$ of leaves determined by the partition P_0^{460} we calculated deviation values δ_h^k and δ_l^k and their averages $\overline{\delta}_h$ and $\overline{\delta}_l$ taken over C_G. We also calculated the associated standard deviation values $\sigma_{\delta h}$ and $\sigma_{\delta l}$ (see Appendix D) and the proportions, λ_h, λ_l, θ_h, θ_l, λ_a, β_a, λ_{amc}, and β_{amc} (see Appendix E); calculated values are presented in Tables 7 and 8. Values of the image selection index IS_x were obtained and compared. For easy of presentation we focus on the results obtained for ST(68), ST(128), and ST(192) which include the smallest obtained value for IS_x (Table 8). Figures 1, 2, and 3 display comparisons of the averages $\overline{l}_o^k$ and $\overline{h}_o^k$ of observed leaf lengths and widths in groups $G_k(l)$ versus estimations $\overline{l}_d^k$ and $\overline{h}_d^k$ from images acquired using these tolerance of similarity intervals.

For $1 \leq x \leq 67$, λ_l and λ_h values were greater than those calculated for ST(68). Moreover, λ_l and λ_h values obtained using $1 \leq x \leq 127$ were greater than those obtained using ST(128) but smaller than those obtained for $1 \leq x \leq 67$. Nevertheless, λ_l values produced for $1 \leq x \leq 196$ were smaller than those linked to $1 \leq x \leq 128$, but generally λ_h values increased implying a greater concentration of a particular tonality within the range of colors forming the color spectrum of the image. That is, $ST(x)$ broadens or reduces the collection of colors which can be taken into account for pixel identification within the extent of the image. Whenever x stands for a greater amount of colors relative to the assortment defining the image, the pixel selection procedure will lead to subjective identification. This influences $\overline{\delta}_h$ and $\overline{\delta}_l$ values in a direct way. Indeed for ST(192), $\overline{\delta}_h$ was negative (Table 7) which means that most h_{dj}^k values were greater than h_{oj}^k values. Nevertheless, this outcome is limited by the maximum value that x can attain and also by image size. For $1 \leq x \leq 255$, the value of λ_l remained unchanged.

By using the proportion values λ_a, β_a, λ_{amc} and β_{amc}, we assessed that for ST(192) leaf area was overestimated by the method. And although for ST(128) the method slightly overestimated leaf area, for $ST(x)$ intervals beyond ST(128) leaf area overestimation always increased. Moreover, in Table 9 we provide root-mean-square deviation (RMSD) values for comparisons of directly versus image obtained estimations when using $ST(x)$ intervals. We can assess that a ST(128) interval produced the highest RMSD values for the comparison of directly versus image obtained estimations of leaf width, length or area. This is consistent with the smallest value obtained for the IS_x selection index. Therefore, we suggest that a similarity ST(128) interval must be set to process the present *Zostera marina* leaf image set through the method proposed.

TABLE 5: Auxiliary statistics λ_l, λ_h, θ_l, θ_h, λ_a, β_a, λ_{amc} and β_{amc} used to obtain the set of leaves with estimation errors in range for a reliable estimation.

Symbol	Description	Reference equation
λ_l	Proportion of leaves in C_G for which δ_l^k satisfies: $\overline{\delta}_l - \sigma_{\delta l} \le \delta_l^k \le \overline{\delta}_l + \sigma_{\delta l}$	(E.5)
λ_h	Proportion of leaves in C_G for which δ_h^k satisfies: $\overline{\delta}_h - \sigma_{\delta h} \le \delta_h^k \le \overline{\delta}_h + \sigma_{\delta h}$	(E.6)
θ_l	Proportion of leaves in C_G for which δ_l^k do not satisfies: $\overline{\delta}_l - \sigma_{\delta l} \le \delta_l^k \le \overline{\delta}_l + \sigma_{\delta l}$	(E.7)
θ_h	Proportion of leaves in C_G for which δ_h^k do not satisfies: $\overline{\delta}_h - \sigma_{\delta h} \le \delta_h^k \le \overline{\delta}_h + \sigma_{\delta h}$	(E.8)
λ_a	Proportion of leaves in C_G for which δ_l^k and δ_h^k satisfies: $\overline{\delta}_l - \sigma_{\delta l} \le \delta_l^k \le \overline{\delta}_l + \sigma_{\delta l}$ and $\overline{\delta}_h - \sigma_{\delta h} \le \delta_h^k \le \overline{\delta}_h + \sigma_{\delta h}$, and $e_{aj}^k \ge 0$.	(E.9)
β_a	Proportion of leaves in C_G for which δ_l^k and δ_h^k do not satisfies: $\overline{\delta}_l - \sigma_{\delta l} \le \delta_l^k \le \overline{\delta}_l + \sigma_{\delta l}$ and $\overline{\delta}_h - \sigma_{\delta h} \le \delta_h^k \le \overline{\delta}_h + \sigma_{\delta h}$	(E.10)
λ_{amc}	Proportion λ_{amc} equivalent to λ_a respectively but linked to leaf area estimation by Monte Carlo method (cf. (2)).	(E.11)
β_{amc}	Proportion β_{amc} equivalent to β_a but linked to leaf area estimation by Monte Carlo method (cf. (2)).	(E.12)

TABLE 6: Numbers n_k of whole leaves classified in groups $G_k(l)$ formed by leaf sizes varying in corresponding length intervals I_k.

k	I_k	n_k	k	I_k	n_k
1	[0, 10)	10	24	[230, 240)	24
2	[10, 20)	43	25	[240, 250)	23
3	[20, 30)	38	26	[250, 260)	15
4	[30, 40)	38	27	[260, 270)	21
5	[40, 50)	32	28	[270, 280)	16
6	[50, 60)	37	29	[280, 290)	12
7	[60, 70)	43	30	[290, 300)	10
8	[70, 80)	32	31	[300, 310)	9
9	[80, 90)	34	32	[310, 320)	9
10	[90, 100)	38	33	[320, 330)	4
11	[100, 110)	28	34	[330, 340)	7
12	[110, 120)	40	35	[340, 350)	3
13	[120, 130)	28	36	[350, 360)	4
14	[130, 140)	29	37	[360, 370)	3
15	[140, 150)	19	38	[370, 380)	3
16	[150, 160)	27	**39**	**[380, 390)**	**1**
17	[160, 170)	19	**40**	**[390, 400)**	**1**
18	[170, 180)	14	**41**	**[400, 410)**	**1**
19	[180, 190)	17	**42**	**[410, 420)**	**1**
20	[190, 200)	21	**43**	**[420, 430)**	**1**
21	[200, 210)	19	**44**	**[430, 440)**	**0**
22	[210, 220)	20	**45**	**[440, 450)**	**0**
23	[220, 230)	14	**46**	**[450, 460)**	**1**

FIGURE 1: (a) Comparison of observed $\overline{h}_o^k$ and image estimated $\overline{h}_d^k$ width averages taken over groups $G_k(l)$. (b) Comparison of observed $\overline{l}_o^k$ and image estimated $\overline{l}_d^k$ length averages taken over groups $G_k(l)$ (see Table 6). The values obtained from digitized leaves were estimated by using ST(68).

5. Discussion

Deleterious effects derived from anthropogenic influences are currently increasing worldwide, threatening the health of many eelgrass meadows [19]. Remediation efforts have included transplant projects 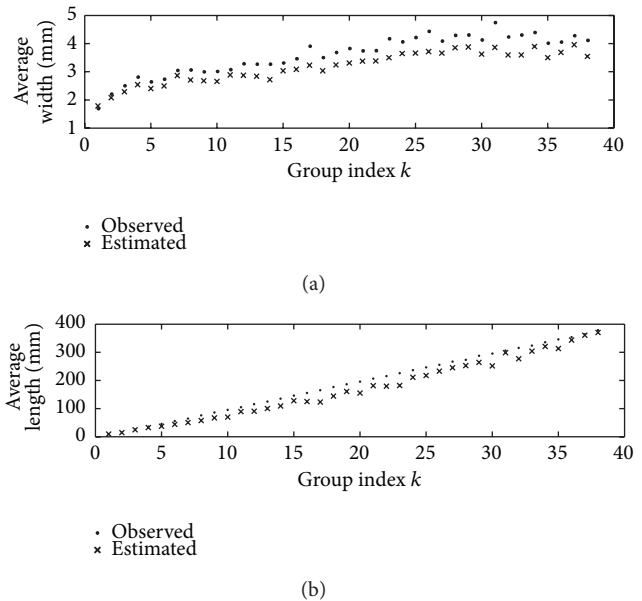[20] and the valuation of their status, requires the estimations of key variables such as standing stock or productivity. Although shoot removal for the measurement of these variables does not damage natural seagrass populations, these procedures could produce undesirable effects on transplants. Therefore, when standing stock or productivity assessment are performed over the early stages of an eelgrass restoration experiment,

TABLE 7: Direct comparison statistics for different ST(x) range values.

ST(x)	$\overline{\delta}_h$	$\sigma_{\delta h}$	$\overline{\delta}_l$	$\sigma_{\delta l}$	θ_l	θ_h	λ_l	λ_h
ST(68)	0.4493	0.2721	24.0157	23.6548	0.0161	0.1038	0.9839	0.8962
ST(128)	0.2599	0.2576	5.0342	13.7282	0.0049	0.0445	0.9951	0.9555
ST(192)	−0.1291	0.2496	3.8965	12.9700	0.0049	0.1669	0.9951	0.8331

TABLE 8: Proportions of overestimation and underestimation of leaf area and selection index values for a given ST(x) range.

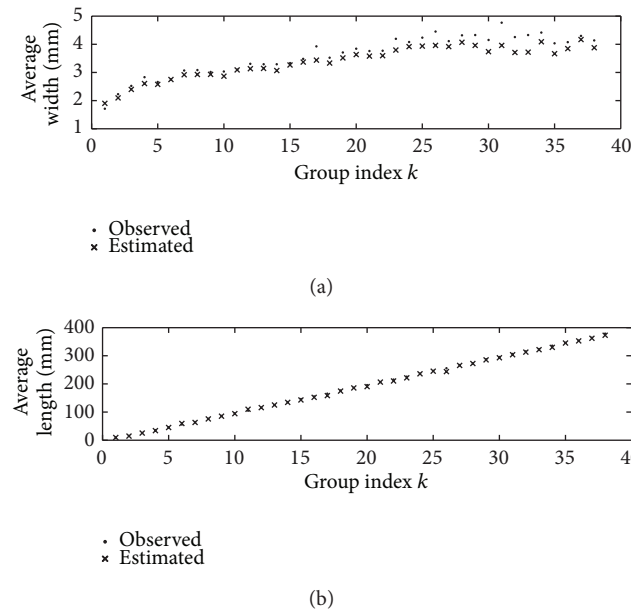| ST(x)	λ_a	β_a	IS$_x$	λ_{amc}	β_{amc}	IS$_{xmc}$
ST(68)	0.6820	0.3180	0.4662	0.6666	0.3344	0.5016
ST(128)	0.7005	0.2995	0.4275	0.7197	0.2803	0.3894
ST(192)	0.4982	0.5018	1.0072	0.4917	0.5083	1.0337

(a)

(b)

FIGURE 2: (a) Comparison of observed $\overline{h}_o^k$ and image estimated $\overline{h}_d^k$ width averages taken over groups $G_k(l)$. (b) Comparison of observed $\overline{l}_o^k$ and image estimated $\overline{l}_d^k$ length averages taken over groups $G_k(l)$ (see Table 6). The values obtained from digitized leaves were estimated by using ST(68). The values obtained from digitized leaves were estimated by using ST(128).

(a)

(b)

FIGURE 3: (a) Comparison of observed $\overline{h}_o^k$ and image estimated $\overline{h}_d^k$ width averages taken over groups $G_k(l)$. (b) Comparison of observed $\overline{l}_o^k$ and image estimated $\overline{l}_d^k$ length averages taken over groups $G_k(l)$ (see Table 6). The values obtained from digitized leaves were estimated by using ST(68). The values obtained from digitized leaves were estimated by using ST(192).

data gathering approaches that avoid disruptive interference are essential. Allometric methods can provide convenient proxies, which reduce leaf biomass and growth assessments to simple blade length or area measurements. What is more, if the estimation of these leaf attributes can be done without removing the blades allometric approaches could furnish truly non-destructive assessments. Modern electronic scanning technologies could be used under water in order to produce reliable images of marine plants leaves, which guarantee non-destructive sampling of leaves length, width or area. However, insitu scanning of eelgrass leaves could add extraneous information mainly due to the inherent humidity content or to materials attached to blades like mud related particles. Hence, for estimating leaf attributes such
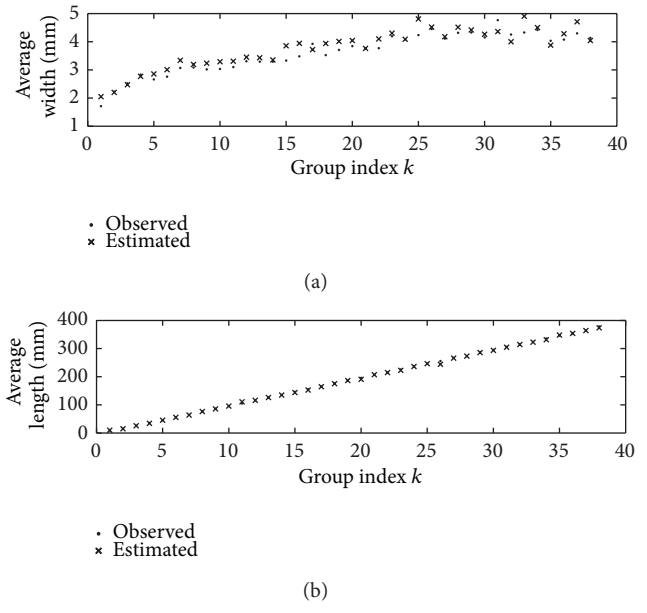
as length or area, we must take into account that image noise effects could render biased assessments. Moreover, if we strive to use (6), the digital image counterpart of (1), we must take into account that a *Zostera marina* leaf does not show a perfect rectangular shape. We should also notice that since width attains small values, noise produced by humidity could increase the width of the digital leaf in a noticeable way which could certainly heighten uncertainty in leaf area estimations produced through (5). What is more, even when leaf area is estimated from images using Monte Carlo methods, noise effects could produce ambiguity in peripheral contour identification rendering biased results.

In the other hand, we need to be aware that since the power functions involved in allometric approaches are highly

TABLE 9: RMSD calculated by using observed versus image calculated variables.

ST(x)	RMSD(h_o, h_d)	RMSD(l_o, l_d)	RMSD(a_o, a_d)	RMSD(a_o, a_{mc})
ST(68)	0.4590	26.4500	360.8746	151.2869
ST(128)	0.4016	12.9587	99.1725	90.6759
ST(192)	0.7303	10.8674	155.3371	160.6715

sensitive to parameter uncertainty we require consistent estimations of blade length or area [10]. Therefore, when we estimate eelgrass leaf area, using digital imagery in combination with allometric proxies, we must be aware that image noise could certainly reduce the accuracy of estimates. This makes it necessary to rely on efficient image selection methodologies for uncertainty reduction. Our results show that the present methods produce reliable results. This conclusion is mainly substantiated by the obtained values for the RMSD. We used these statistics to determine consistency between directly obtained measurements and image estimated assessments. Table 9 shows that the highest RMSD values corresponded with the smallest values obtained for the IS_x selection index. Moreover, the ST(x) interval selected using the IS_x criteria produced also the highest RMSD values for the comparison of observed values and those obtained by means of Monte Carlo method. This justifies our claim that the proposed procedure abridged by the selection index IS_x can be expected to produce consistent estimations of the leaf attributes necessary for allometric estimation of relevant variables required to assess the status of an eelgrass population. Moreover, the presented procedure could be straightforwardly applied to other eelgrass populations or seagrass species that exhibit similar leaf architectures making leaf length times width a reliable proxy for the pertinent area.

Appendices

We now explain how to identify the interval of tolerance of similarity ST(x) that yields accurate estimations for the observed leaf area values. The task, is achieved through statistical methodologies, which requires the completion of the following steps.

A. Grouping Leaf Data: l_o and h_o

(A.1) Identify the maximum observed leaf length (l_{max}).

(A.2) Chose a positive integer n and define a partition of the interval $[0, l_{max}]$ with norm $q = (l_{max}/n)$.

(A.3) Form the collection $\bigcup_1^n(I_k)$ of n disjoint intervals of the form $I_k = [q(k-1), qk)$ with $1 \leq k \leq n$. This collection is denoted thought $P_0^{l_{max}}$.

(A.4) For each value of the index k identify the group $G_k(l)$ of leaves whose lengths are contained in I_k. Notice that $G_k(l)$ holds the leaves whose size differences are bounded by q.

(A.5) For each value of the index k obtain and record n_k standing for the number of leaves in the group $G_k(l)$.

(A.6) For each value of the index k introduce an index j such that $1 \leq j \leq n_k$ and label as l_{oj}^k, h_{oj}^k and a_{oj}^k respectively, the straight length, width and area of the jth leaf in $G_k(l)$. The character a_{oj}^k denotes the associated estimations of leaf area obtained by means of (1).

(A.7) Form and record the collection $\bigcup_1^n G_k(l)$ of all groups of leaves $G_k(l)$. This collection is denoted by means of C_G.

(A.8) Obtain the average length $\bar{l}_o^k$ for each group of leaves $G_k(l)$. That is, calculate and record

$$\bar{l}_o^k = \frac{1}{n_k}\sum_1^{n_k} l_{oj}^k. \tag{A.1}$$

(A.9) Obtain the average width $\bar{h}_o^k$ for each group of leaves $G_k(l)$. That is, calculate and record

$$\bar{h}_o^k = \frac{1}{n_k}\sum_1^{n_k} h_{oj}^k. \tag{A.2}$$

(A.10) Obtain the average area $\bar{a}_o^k$ for each group of leaves $G_k(l)$. That is, calculate and record,

$$\bar{a}_o^k = \frac{1}{n_k}\sum_1^{n_k} a_{oj}^k. \tag{A.3}$$

B. Obtaining Length, Width and Area from the Image of each Leaf

(B.1) For processing all digital images, we chose a specified color format with a number C_{max} of colors.

(B.2) For processing the digital images of all collected leaves, we choose different intervals of tolerance of similarity ST(x) = $[0, x]$, with the upper bound x satisfying $0 \leq x \leq C_{max} - 1$.

(B.3) For a picked ST(x) interval, for $1 \leq j \leq n_k$ use the algorithm (a) described in the method section to obtain l_{dj}^k and h_{dj}^k, which respectively denote the length and width of the image of the jth leaf in $G_k(l)$. Also, obtain a_{dj}^k and a_{mcj}^k, which respectively stand for leaf area obtained from the image and calculated by means of (1) and (2) respectively. Record these values.

(B.4) interval obtain and record the concomitant averages $\bar{l}_d^k$, $\bar{h}_d^k$, $\bar{a}_d^k$ and $\bar{a}_{cm}^k$ (cf. (A.1) through (A.3)).

C. Obtain Estimation Errors between the Observed and Image Obtained Values in Steps A and B

(C.1) For the picked $ST(x)$ interval and for $1 \leq j \leq n_k$, calculate the leaf length approximation errors through

$$e_{lj}^k = l_{oj}^k - l_{dj}^k. \tag{C.1}$$

(C.2) For the picked $ST(x)$ interval and for $1 \leq j \leq n_k$, calculate the individual leaf width approximation errors through

$$e_{hj}^k = h_{oj}^k - h_{dj}^k. \tag{C.2}$$

(C.3) For the picked $ST(x)$ interval and for $1 \leq j \leq n_k$, calculate the leaf area approximation errors through

$$e_{aj}^k = a_{oj}^k - a_{dj}^k. \tag{C.3}$$

(C.4) For the picked $ST(x)$ interval and for $1 \leq j \leq n_k$, calculate the leaf area approximation errors linked to the Monte Carlo method through

$$e_{mcj}^k = a_{oj}^k - a_{mcj}^k. \tag{C.4}$$

D. Obtain the Average Deviations Produced by the Individual Estimation Errors

(D.1) For the picked $ST(x)$ interval, obtain the average leaf length deviations δ_l^k, which are calculated by averaging the e_{lj}^k values. That is,

$$\delta_l^k = \frac{1}{n_k} \sum_1^{n_k} e_{lj}^k. \tag{D.1}$$

We notice that $\delta_l^k = (\bar{l}_o^k - \bar{l}_d^k)$ and also that negative values of δ_l^k imply that in a lot $G_k(l)$ most image assessments l_{dj}^k overestimate observed l_{oj}^k values.

(D.2) For the picked $ST(x)$ interval, obtain the average leaf width deviations δ_h^k, which are calculated by averaging the e_{hj}^k values. That is,

$$\delta_h^k = \frac{1}{n_k} \sum_1^{n_k} e_{hj}^k. \tag{D.2}$$

notice that since $\delta_h^k = (\bar{h}_o^k - \bar{h}_d^k)$, negative values of δ_h^k imply that in a group $G_k(l)$ most image assessments h_{dj}^k overestimate observed h_{oj}^k values.

(D.3) For the picked $ST(x)$ interval, calculate the average leaf area deviations δ_a^k by averaging the e_{aj}^k values. That is,

$$\delta_a^k = \frac{1}{n_k} \sum_1^{n_k} e_{aj}^k. \tag{D.3}$$

again since $\delta_a^k = \bar{a}_o^k - \bar{a}_d^k$ negative values of δ_a^k imply that in a group $G_k(l)$ most image assessments a_{dj}^k overestimate observed a_{oj}^k values.

(D.4) For the picked $ST(x)$ interval, calculate $\bar{\delta}_l$, the average value of deviations δ_l^k taken over C_G. Calculate also the associated standard deviation $\sigma_{\delta l}$.

(D.5) For the picked $ST(x)$ interval, calculate $\bar{\delta}_h$, the average value of deviation δ_h^k taken over C_G. Calculate also the associated standard deviation $\sigma_{\delta h}$.

E. Criteria for Selecting the $ST(x)$ Interval That Produces the Highest Correspondence Level between Image Obtained Measurements and Those Obtained Directly from Collected Leaves

(E.1) For a given range of similarity values $ST(x) = [0, x]$, identify the leaves satisfying the conditions

$$\bar{\delta}_h \geq 0, \tag{E.1}$$

$$\bar{\delta}_l \geq 0, \tag{E.2}$$

$$\bar{\delta}_l - \sigma_{\delta l} \leq \delta_l^k \leq \bar{\delta}_l + \sigma_{\delta l}, \tag{E.3}$$

$$\bar{\delta}_h - \sigma_{\delta h} \leq \delta_h^k \leq \bar{\delta}_h + \sigma_{\delta h}. \tag{E.4}$$

(E.2) Calculate the proportion λ_l of leaves in C_G that comply with the condition (E.3) through

$$\lambda_l = \sum_{k=1}^{n} \sum_{j=1}^{n_k} \left[l_{dj}^k \mid \text{leaves in } G_k \text{ that comply with} \right.$$
$$\left. \text{condition (E.3)} \right] \times \left(\sum_{k=1}^{n} \sum_{j=1}^{n_k} l_{oj}^k \right)^{-1}. \tag{E.5}$$

(E.3) Calculate the proportion λ_h of leaves in C_G that comply with the condition (E.4) through

$$\lambda_h = \sum_{k=1}^{n} \sum_{j=1}^{n_k} \left[h_{dj}^k \mid \text{leaves in } G_k \text{ that comply with} \right.$$
$$\left. \text{condition (E.4)} \right] \times \left(\sum_{k=1}^{n} \sum_{j=1}^{n_k} h_{oj}^k \right)^{-1}. \tag{E.6}$$

(E.4) Calculate the proportion θ_l of leaves in C_G that do not comply with the condition (E.3) through,

$$\theta_l = 1 - \lambda_l. \tag{E.7}$$

(E.5) Calculate the proportion θ_h of leaves in C_G that do not comply with the condition (E.4) through,

$$\theta_h = 1 - \lambda_h. \tag{E.8}$$

(E.6) Obtain the concomitant proportions of leaves in C_G that provide consistent leaf area estimations by the proxy of (1)

$$\lambda_a = \sum_{k=1}^{n} \sum_{j=1}^{n_k} \left[a_{dj}^k \mid \text{leaves in } G_k \text{ that comply with}\right.$$

$$\left. \text{condition } (E.3), (E.4) \text{ and } e_{aj}^k \geq 0 \right] \quad (E.9)$$

$$\times \left(\sum_{k=1}^{n} \sum_{j=1}^{n_k} a_{oj}^k \right)^{-1}.$$

(E.7) Calculate the proportion of leaves in C_G for which image estimated blade length and width measurements overestimate leaf area calculated through (1)

$$\beta_a = 1 - \lambda_a. \quad (E.10)$$

(E.8) The proportions λ_{amc} and β_{amc} equivalent to λ_a and β_a respectively but linked to leaf area estimation by Monte Carlo method (cf. (2)) that is,

$$\lambda_{amc} = \sum_{k=1}^{n} \sum_{j=1}^{n_k} \left[a_{mcj}^k \mid \text{leaves in } G_k \text{ that comply with}\right.$$

$$\left. \text{condition } (E.3), (E.4) \text{ and } e_{mcj}^k \geq 0 \right]$$

$$\times \left(\sum_{k=1}^{n} \sum_{j=1}^{n_k} a_{oj}^k \right)^{-1},$$

$$(E.11)$$

$$\beta_{amc} = 1 - \lambda_{amc}. \quad (E.12)$$

Conditions (E.1) and (E.2) grant bounded estimation errors, for h and l respectively. Moreover, the groups of leaves that also satisfy conditions (E.3) and (E.4) can be identified as those groups for which image l_d and h_d estimations are closer to directly obtained l_o and h_o measurements. Therefore, groups in C_G, which do not comply with conditions (E.1)–(E.4), denote the set of leaves with estimation errors out of range for a reliable estimation.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

References

[1] H. Echavarria-Heras, E. Solana-Arellano, C. Leal-Ramirez, and E. Franco-Vizcaíno, "The length-times-width proxy for leaf area of eelgrass: criteria for evaluating the representativeness of leaf-width measurements," *Aquatic Conservation: Marine and Freshwater Ecosystems*, vol. 21, no. 7, pp. 604–613, 2011.

[2] T. P. Williams, J. M. Bubb, and J. N. Lester, "Metal accumulation within salt marsh environments," *Marine Pollution Bulletin*, vol. 28, no. 5, pp. 277–289, 1994.

[3] F. T. Short and C. A. Short, "The seagrass filter: purification of coastal water," in *The Estuary as a Filter*, V. S. Kennedy, Ed., pp. 395–413, Academic Press, 1984.

[4] L. G. Ward, W. Michael Kemp, and W. R. Boynton, "The influence of waves and seagrass communities on suspended particulates in an estuarine embayment," *Marine Geology*, vol. 59, no. 1–4, pp. 85–103, 1984.

[5] M. S. Fonseca and J. S. Fisher, "A comparison of canopy friction and sediment movement between four species of seagrass with reference to their ecology and restoration," *Marine Ecology Progress Series*, vol. 29, pp. 15–22, 1986.

[6] F. Tuya, L. Ribeiro-Leite, N. Arto-Cuesta, J. Coca, R. Haroun, and F. Espino, "Decadal changes in the structure of *Cymodocea nodosa* seagrass meadows: natural vs. human influeces," *Estuarine Coastal and Shelf Science*, vol. 137, pp. 41–49, 2014.

[7] W. C. Dennison, "Leaf production," in *Seagrass Research Methods*, C. P. Ronald and C. P. McRoy, Eds., pp. 77–79, 1990.

[8] J. L. Gaeckle and F. T. Short, "A plastochrone method for measuring leaf growth in eelgrass, *Zostera marina* L.," *Bulletin of Marine Science*, vol. 71, no. 3, pp. 1237–1246, 2002.

[9] H. A. Echavarria-Heras, M. E. Solana-Arellano, K. S. Lee, S. Hosokawa, and E. Franco-Vizcaino, "An evaluation of leaf biomass: length ratio as a tool for nondestructive assessment in eelgrass (*Zostera marina* L.)," *The ScientificWorldJournal*, vol. 2012, Article ID 543730, 2012.

[10] H. Echavarría-Heras, E. Solana-Arellano, and E. Franco-Vizcaíno, "An allometric method for the projection of eelgrass leaf biomass production rates," *Mathematical Biosciences*, vol. 223, no. 1, pp. 58–65, 2010.

[11] C. Igathinathane, A. R. Womac, S. Sokhansanj, and L. O. Pordesimo, "Mass and moisture distribution in aboveground components of standing corn plants," *Transactions of the ASABE*, vol. 49, no. 1, pp. 97–106, 2006.

[12] E. . Rico-García, F. Hernández-Hernández, G. M. Soto-Zarazúa, and G. Herrera-Ruiz, "Two new methods for the estimation of leaf area using digital photography," *International Journal of Agriculture and Biology*, vol. 11, pp. 397–400, 2009.

[13] A. Femat-Diaz, D. Vargas-Vazquez, E. Huerta-Manzanilla, E. Rico-Garcia, and G. Herrera-Ruiz, "Scanner image methodology (SIM) to measure dimensions of leaves for agronomical applications," *African Journal of Biotechnology*, vol. 10, no. 10, pp. 1840–1847, 2011.

[14] S. B. Patil and S. K. Bodh, "Image processing method to measure sugarcane leaf area," *International Journal of Engineering Science and Technology*, vol. 3, no. 8, pp. 6394–6400, 2011.

[15] C. Lü, H. Ren, Y. Zhang, and Y. Shen, "Leaf area measurement based on image processing," in *Proceedings of the International Conference on Measuring Technology and Mechatronics Automation (ICMTMA '10)*, vol. 2, pp. 580–582, IEEE Computer Society, March 2010.

[16] A. Ramfos, A. Gazis, and G. Katselis, "Development and evaluation of an automated digital image analysis software for obtaining seagrass leaf metrics," *Botanica Marina*, vol. 55, no. 6, pp. 601–610, 2012.

[17] K. J. . Juneau and C. S. Tarasoff, "Leaf area and water content changes after permanent and temporary storage," *PLoS ONE*, vol. 7, no. 8, Article ID e42604, 2012.

[18] J. M. Engle and K. A. Miller, "Distribution and morphology of eelgrass (*Zostera marina* L.) at the California channel islands," in *Proceedings of the 6th California Islands Symposium*, D. K. Garcelon and C. A. Scwemm, Eds., National Park Service Technical Publication CHIS-05-01, pp. 405–414, 2005.

[19] R. J. Orth, T. J. B. Carruthers, W. C. Dennison et al., "A global crisis for seagrass ecosystems," *BioScience*, vol. 56, no. 12, pp. 987–996, 2006.

[20] M. Björk, F. Short, E. McLeod, and S. Beer, "Managing seagrasses for resilience to climate Change," IUCN Resilience Science Group Working Paper Series no 3, 2008.