

RESEARCH

Open Access



# On the suitability of an allometric proxy for nondestructive estimation of average leaf dry weight in eelgrass shoots I: sensitivity analysis and examination of the influences of data quality, analysis method, and sample size on precision

Héctor Echavarría-Heras<sup>1\*</sup>, Cecilia Leal-Ramírez<sup>1</sup>, Enrique Villa-Diharce<sup>2</sup> and Nohe Cazarez-Castro<sup>3</sup>

## Abstract

**Background:** The effects of current anthropogenic influences on eelgrass (*Zostera marina*) meadows are noticeable. Eelgrass ecological services grant important benefits for mankind. Preservation of eelgrass meadows include several transplantation methods. Evaluation of establishing success relies on the estimation of standing stock and productivity. Average leaf biomass in shoots is a fundamental component of standing stock. Existing methods of leaf biomass measurement are destructive and time consuming. These assessments could alter shoot density in developing transplants. Allometric methods offer convenient indirect assessments of individual leaf biomass. Aggregation of single leaf projections produce surrogates for average leaf biomass in shoots. Involved parameters are time invariant, then derived proxies yield simplified nondestructive approximations. On spite of time invariance local factors induce relative variability of parameter estimates. This influences accuracy of surrogates. And factors like analysis method, sample size and data quality also impact precision. Besides, scaling projections are sensitive to parameter fluctuation. Thus the suitability of the addressed allometric approximations requires clarification.

**Results:** The considered proxies produced accurate indirect assessments of observed values. Only parameter estimates fitted from raw data using nonlinear regression, produced robust approximations. Data quality influenced sensitivity and sample size for an optimal precision.

**Conclusions:** Allometric surrogates of average leaf biomass in eelgrass shoots offer convenient nondestructive assessments. But analysis method and sample size can influence accuracy in a direct manner. Standardized routines for data quality are crucial on granting cost-effectiveness of the method.

**Keywords:** Eelgrass conservation, Nondestructive estimation, Allometric projection suitability, methodological influences, Sensitivity analysis

\* Correspondence: [heheras@icloud.com](mailto:heheras@icloud.com); <http://www.editorialmanager.com/tbmm/default.aspx>

<sup>1</sup>Centro de Investigación Científica y de Estudios Superiores de Ensenada, Carretera Ensenada-Tijuana No. 3918, Zona Playitas, Código Postal 22860, Apdo. Postal 360 Ensenada, B.C., Mexico

Full list of author information is available at the end of the article



## Background

Based on published data for 17 ecosystem services in 16 biomes, Costanza et al. [1] estimated the value of ecosystem services at the level of the whole biosphere. They found a lower bound in the range of US\$16–54 trillion (1012) per year, with an average of US\$33 trillion per year. Marine systems produced near 63% of the annual value. Almost half derived from coastal ecosystems. Approximately 25% of this share related to algal beds and seagrasses. This contribution to human welfare is deemed relevant. Thus maintaining the health of marine ecosystems is subject of scientific concerns. Currently, increasing anthropogenic pressures pose important threats. And global climate change could also threaten future viability of seagrass meadows [2, 3]. For instance, water quality and other local stressors promote unprecedented meadow loss [4–6]. This reduces mitigation of wave action [7] and filtration [8]. Diminishes food and shelter for a myriad of organisms [9–11]. Weakens nutrient cycling [12, 13], erosion abatement and shoreline stabilization [14–16]. Moderates support for detrital food web foundation [17]. And inhibits carbon sequestration [18, 19].

Eelgrass (*Zostera marina* L.) is a dominant along the coasts of both the North Pacific and North Atlantic [20]. This species supports communities known as among the richest and most diverse in sea life [21]. Contribution of organic materials for food webs in shallow environments [22] is noticeable. Indeed, eelgrass produced up to 64% of the whole primary production of an estuarine system [23]. Current deleterious effects of anthropogenic influences on eelgrass prompted special restoration strategies. Among remediation efforts replanting plays an important role [24–27]. Transplant success amounts to reinstatement of ecological functions of natural populations. Evaluation relies on monitoring standing stock and productivity of transplanted plants. Then comparing with assessments of a reference population, which usually settle nearby [28].

Combined biomass of leaves in shoots is an important component of standing stock. Assessments rely on the estimation of the biomass of individual leaves. This requires shoot removal followed by dry weight measurement procedures in the laboratory. Elimination of shoots could infringe damage to natural eelgrass populations [29]. And reduced shoot density makes these effects even more perceptible for transplanted plots. Allometric methods make it possible simplified-indirect estimations of eelgrass productivity and standing stock. Echavarría-Heras et al. [30] considered an allometric representation for eelgrass leaf biomass and related length. Agreeing with Solana et al. [31], the involved parameters are invariant within a given geographical region. Estimates and leaf length measurements grant nondestructive approximations of observed leaf biomass values. This way,

leaf length measurements grant nondestructive approximations for observed leaf biomass values. Leaf growth rates estimation relies on successive measurements of leaf biomass. Then the allometric model in [30] entails nondestructive assessments of eelgrass productivity. But, invariance does not impede local factors to imply variability of parameter estimates. Besides, local influences other factors could explain numerical differences in parameter estimates. There are methodological influences that may render biased parameter estimates. Analysis method, sample size, and data quality can influence scaling results (e.g. [32, 33]). And, since scaling relationships are particularly sensitive to parametric uncertainties, Echavarría-Heras et al. [30] concluded that the actual precision of derived allometric surrogates requires clarification.

Here we deal with allometric surrogates for average leaf biomass in eelgrass shoots. These derive from the model  $w(t) = \beta a(t)^\alpha$  for leaf biomass  $w(t)$  and area  $a(t)$  measured at time  $t$ , and  $\alpha$  and  $\beta$  parameters. Leaf area is more informative of eelgrass leaf biomass than corresponding length. Thus, the present scaling endures a boost in precision of parameter estimates by the model in [30]. This could increase the accuracy of derived surrogates for leaf biomass in shoots. Besides, eelgrass leaf area and length admit an isometric representation [34, 35]. Then, the time invariance found by Solana-Arellano et al. [31] also holds for parameter estimates of the present scaling. This by the way imbeds a nondestructive advantage to the present shoot-biomass substitutes. But, agreeing with Echavarría-Heras et al. [30], we must examine influences on precision of estimates for suitability of projections. Since, such an analysis was not produced before, we took here the try of filling that gap. Achieving the related goals, required the assemblage of an extensive data set. It comprises coupled measurements of eelgrass leaf biomasses and related areas. This is further called “raw data set”. A data cleaning procedure adapted from Echavarría-Heras et al. [30] removed inconsistent leaf biomass replicates from the raw data. Thereby forming what we call a “processed data set”. Differences in reproducibility strength allowed to assess data quality effects in precision. A similar procedure evaluated sample size effects. And a sensitivity analysis evaluated robustness of the projection method. This supports consistent, cost-effective allometric projections of observed values from raw data. But, this depends on nonlinear regression as an analysis method. Besides, sample size must be optimal. Data quality as expected improved reproducibility strength of the allometric projection method. But, this factor was more relevant in optimizing sample size. A detailed explanation of used procedures appears in the methods section. The results section is not only devoted to the presentation of our findings. It also examines the relative strengths of factors influencing the precision of proxies. A Discussion section emphasizes on the gains

and the limitations of the method. [Appendix 1](#) deals with the model selection problem. [Appendix 2](#) is about data processing methodology. [Appendix 3](#) presents the procedure for sensitivity assessment.

## Methods

The symbol  $w_m(t)$  will stand for average leaf dry weight of shoots collected at sampling date  $t$ . An the average of these values over all sampling dates symbolized through  $\langle w_m(t) \rangle$ . Formal representations of these variables appear in [Appendix 3](#) (cf. Eq. (15) and Eq. (16)). The symbols  $w(t)$  and  $a(t)$  will one to one stand for biomass and area of an individual leaf collected at a sampling time  $t$ . We assume that these variables are linked through the scaling relationship

$$w(t) = \beta a(t)^\alpha. \quad (1)$$

The present raw data come from a coastal lagoon located in San Quintin Bay, México [30]. This comprises 10,412 leaves and measured lengths [mm], widths [mm] and dry weights (g). The product of length times width provided estimations of leaf area [mm<sup>2</sup>] [36]. In what follows the symbol  $n_{ra}$  stands for number of leaves in raw data. Processed data results by applying direct and statistical data cleaning techniques. The direct hinges on the consistency of allometric models for eelgrass leaf biomass. Leaf length or area are allometric descriptors of eelgrass leaf biomass [30, 31, 34]. A model selection exploration corroborated a power function like trend assumption for the present data. Details appear in [Appendix 1](#). Severe deviations, from the mean response curve, are inconsistent and must be removed. This took care of sets containing less than ten leaf dry weight replicates. The statistical procedure worked on sets with a larger number of replicates. It centers on properties of the median of a group of data. This is immune to sample size and also a robust estimator of scale. The adapted Median Absolute Deviation (MAD) data cleaning procedure [37] appears in [Appendix 2](#). Processing data resulted in a number of  $n_{qa} = 6094$  pairs of leaf dry weights and areas.

Parameter estimates  $\hat{\alpha}$  and  $\hat{\beta}$  and leaf area values yield allometric proxies  $w_m(\hat{\alpha}, \hat{\beta}, t)$  for  $w_m(t)$ . (cf. Eq. (19)). The symbol  $\langle w_m(\hat{\alpha}, \hat{\beta}, t) \rangle$  (cf. Eq. (20)) stands for the pertinent average over sampling dates. We use Lin's Concordance Correlation Coefficient (CCC) [38] as an evaluation of reproducibility. This meant as the extent to which two connected variables fall on a line through the origin and with a slope of one. We represent this statistic by means of the symbol  $\hat{\rho}$ . Agreement defined as poor whenever  $\hat{\rho} < 0.90$ , moderate for  $0.90 \leq \hat{\rho} < 0.95$ , good for  $0.95 \leq \hat{\rho} \leq 0.99$  or excellent for  $\hat{\rho} > 0.99$  [39]. Values of  $\hat{\rho}$  gave an evaluation of the strength of the  $w_m(\hat{\alpha}, \hat{\beta}, t)$  devise to reproduce observed values.

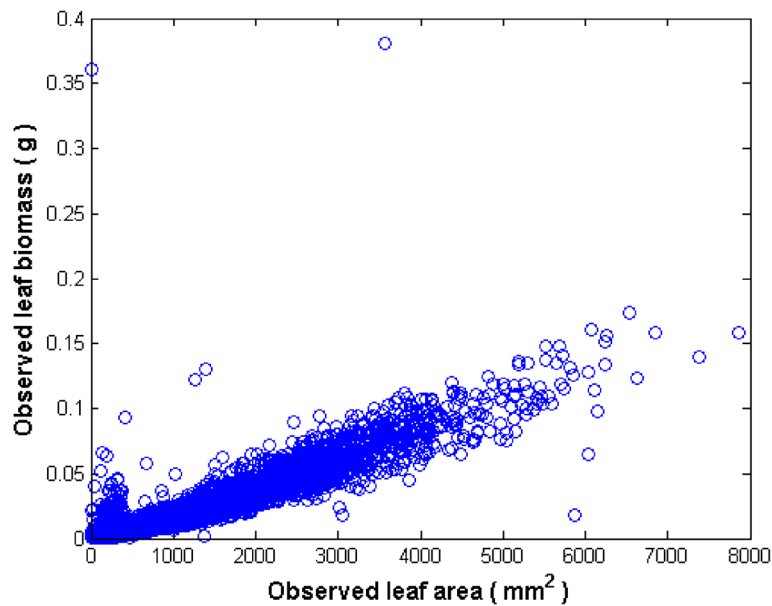
In getting parameter estimates  $\hat{\alpha}$  and  $\hat{\beta}$  we relied on two procedures. The traditional analysis method of allometry and nonlinear regression. Assessing analysis method effects on reproducibility strength of  $w_m(\hat{\alpha}, \hat{\beta}, t)$  depended on testing differences in  $\hat{\rho}$ . The traditional approach involves a linear regression equation (cf. Eq. (4)). This obtained through logarithmic transformation of response and descriptor in Eq. (1). The nonlinear regression analysis method relied on maximum likelihood [40, 41]. This approach fitted the model of Eq. (1) in a direct way in the original arithmetical scale. The nonlinear fit allowed the consideration of homoscedasticity or heteroscedasticity (cf. Eqs. (5) and (6)). All the required fittings for both raw and processed data depended on the use of the R software.

We also fitted the model of Eq. (1) to samples of different sizes taken out from primary and processed data sets. Each sample of size  $k$ ; with  $100 \leq k \leq n_{ra}$  produced estimates  $\hat{\alpha}(k)$  for  $\alpha$  and  $\hat{\beta}(k)$  for  $\beta$ , and resulting  $w_m(\hat{\alpha}(k), \hat{\beta}(k), t)$  projections. The symbol  $\hat{\rho}(k)$  denotes the value of  $\hat{\rho}$  for a sample of size  $k$ . Differences in  $\hat{\rho}(k)$  allow exploring sample size influences in reproducibility.

Deviations  $\Delta\alpha_q$  and  $\Delta\beta_r$  convey fluctuating values  $\alpha_q = \alpha + \Delta\alpha_q$  and  $\beta_r = \beta + \Delta\beta_r$  for the parameters  $\alpha$  and  $\beta$  one to one. The modulus of the vector of parametric changes  $(\Delta\alpha_q, \Delta\beta_r)$  defines a tolerance range  $\theta(q, r)$ . And the value of  $\theta(q, r)$  determined by the standard errors of parameter estimates denoted by mean of  $\theta_{ste}$ . A fixed value of  $\theta(q, r)$  leads to four possible characterizations of the pair  $(\Delta\alpha_q, \Delta\beta_r)$ . Each one associates to a trajectory  $w_m(\alpha_q, \beta_n, t)$  shifting from a reference one  $w_m(\alpha, \beta, t)$ . The symbol  $\delta w_{m\theta}(\alpha_q, \beta_n, t)$  (cf. Eq. (42)) denotes deviations between reference and average of shifting trajectories at sampling dates. And the average of  $\delta w_{m\theta}(\alpha_q, \beta_n, t)$  values taken over all sampling dates denoted through  $\langle \delta w_{m\theta}(\alpha_q, \beta_n, t) \rangle$  (cf. Eq. (43)). The absolute value of the ratio of  $\langle \delta w_{m\theta}(\alpha_q, \beta_n, t) \rangle$  to  $\langle w_m(\alpha, \beta, t) \rangle$  defines a relative deviation index  $\vartheta(\theta)$ . It measures sensitivity of  $\langle w_m(\alpha, \beta, t) \rangle$  to fluctuations of tolerance  $\theta(q, r)$  on  $\alpha$  and  $\beta$ . [Appendix 3](#) presents detailed formulae.

## Results

Figure 1 shows the variation of leaf dry weight and area observed in the raw data. Smallest and largest leaf areas were 2 mm<sup>2</sup> and 7868 mm<sup>2</sup> respectively. Associated dry weights were  $1 \times 10^{-5}$  g and 0.1588 g one to one. The time average of mean leaf dry weight in shoots was  $\langle w_m(t) \rangle = 0.01461$ g (cf. Eq. (16)). Each leaf area measurement associate to several replicates of leaf biomass. Number of replicates increased from a single association up to a largest value of 84. Dispersion masks a power function like trend. Contents of [Appendix 1](#) corroborate this at formal level. And exploration of dispersion reveals severe deviations from the inherent power function-like trend. Inconsistencies are more visible for leaves with areas under 350 mm<sup>2</sup> and also for

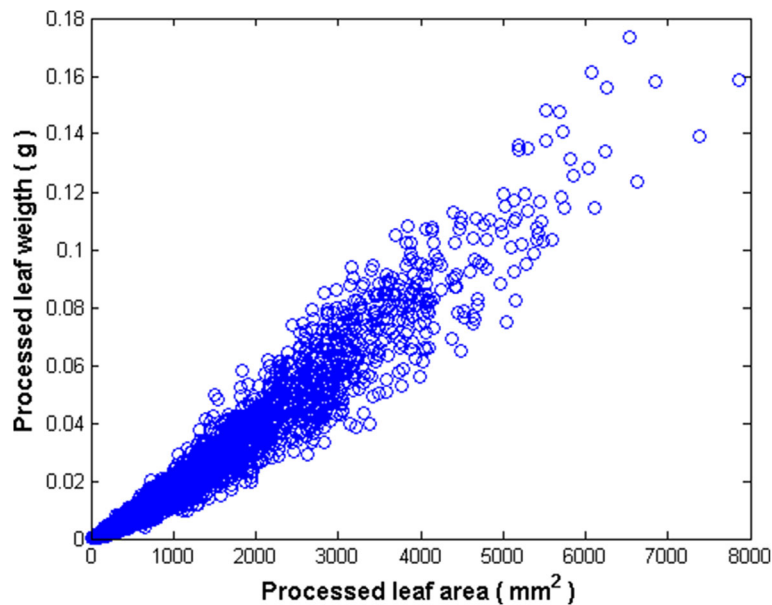


**Fig. 1** Distribution of eelgrass leaf dry weight and linked area values in raw data. Dispersion shows a masked power function-like trend. Deviations from this trend are more manifest for areas under 350 mm<sup>2</sup> and also for those bigger than 5000 mm<sup>2</sup>. This suggests data quality effects on accuracy of allometric projections

those over 5000 mm<sup>2</sup>. This hints about the relevance of data quality.

Figure 2 exhibits the spreading of leaf dry weight after data quality control. About 40% of the replicates in the raw data were eliminated. A power function like trend appears more depicted than that showing on Fig. 1. But

dispersion still shows significant deviations around the expected power function like trend. This suggests lack of standardized routines for data gathering. In this work the length times width proxy [36] approximated leaf area. Errors in estimation of area of older damaged leaves could explain uneven replicates. Faulty



**Fig. 2** Plot of processed data. Distribution of eelgrass leaf dry weight and area values remaining after data quality control procedures. Although about 40% of the replicates in the raw data were found inconsistent and eliminated, this plot still shows significant residual variability around an expected power function like trend

equipment, or incorrect recording could explicate inconsistencies for small leaves.

Table 1 gives estimates  $\hat{\alpha}$  and  $\hat{\beta}$  for the parameters  $\alpha$  and  $\beta$  and corresponding standard errors. Assuming heteroscedasticity in the model of Eqs. (5) and (6) did not affect estimates. Thus, presentation of results of nonlinear regression refers to the homoscedastic case of the model of Eqs. (5) and (6). Figure 3 displays mean response curves fitted using raw data. Figure 4 shows those associated to quality controlled data. Results for the log-linear transformation method included correction for bias of re-transformation [42]. The smearing estimate of bias of Duan [43] provided the form of the correction factor.

For raw data the log-linear transformation method produced  $\hat{\rho} = 0.8910$ , entailing poor reproducibility. This explains a biased distribution of replicates around the mean response curve (Fig. 3a). Meanwhile, estimates acquired by nonlinear regression from raw data conveyed adequate reproducibility ( $\hat{\rho} = 0.9307$ ). This explains a displayed fair distribution of projected leaf biomass values (Fig. 3b).

Estimates via log-linear transformation for processed data seemed enhance reproducibility ( $\hat{\rho} = 0.9777 = 0.9455$ ). But, Fig. 4a, reveals a bulk of inconsistent replicates for leaves areas under 5000 mm<sup>2</sup>. Notice that this subset of replicates distributes almost evenly around the mean response curve. Yet replicate spread for areas beyond 5000 mm<sup>2</sup> shows significant bias (Fig. 4a). Meanwhile, nonlinear regression and processed data associate to  $\hat{\rho} = 0.9777$ . This corresponded to good reproducibility strength. Indeed, spread of replicates around the mean response is fair (Fig. 4b).

Estimates acquired from raw data via the traditional analysis method of allometry returned a value of  $\hat{\rho} = 0.9285$  for  $w_m(\hat{\alpha}, \hat{\beta}, t)$  projections (Table 2). This seems to correspond to moderate reproducibility. Yet, corresponding  $rms = 0.01265$  was largest among analysis method- data set combinations (Table 2). Figure 3a shows a relative wider bias in spread around the mean response curve for larger leaves. This explains resulting

inconsistencies in reproducibility of  $w_m(\hat{\alpha}, \hat{\beta}, t)$  projections shown in Fig. 5a. Display reveals biased  $w_m(\hat{\alpha}, \hat{\beta}, t)$  projections for near 50% of sampling dates. This, led to a smallest value of 0.8436 for the ratio of projected to observed averages.

Instead, nonlinear regression and raw data produced a value of  $\hat{\rho} = 0.9915$ . And root mean squared deviation attained a value of  $rms = 0.00460$  (Table 2). This suggest a remarkable reproducibility strength for  $w_m(\hat{\alpha}, \hat{\beta}, t)$  projections (Table 2). Correspondence between projected and observed values, shown in Fig. 5b. corroborates high agreement. Moreover, the ratio of projected to observed leaf dry weight averages attained an outstanding value of 0.9773 (Table 2).

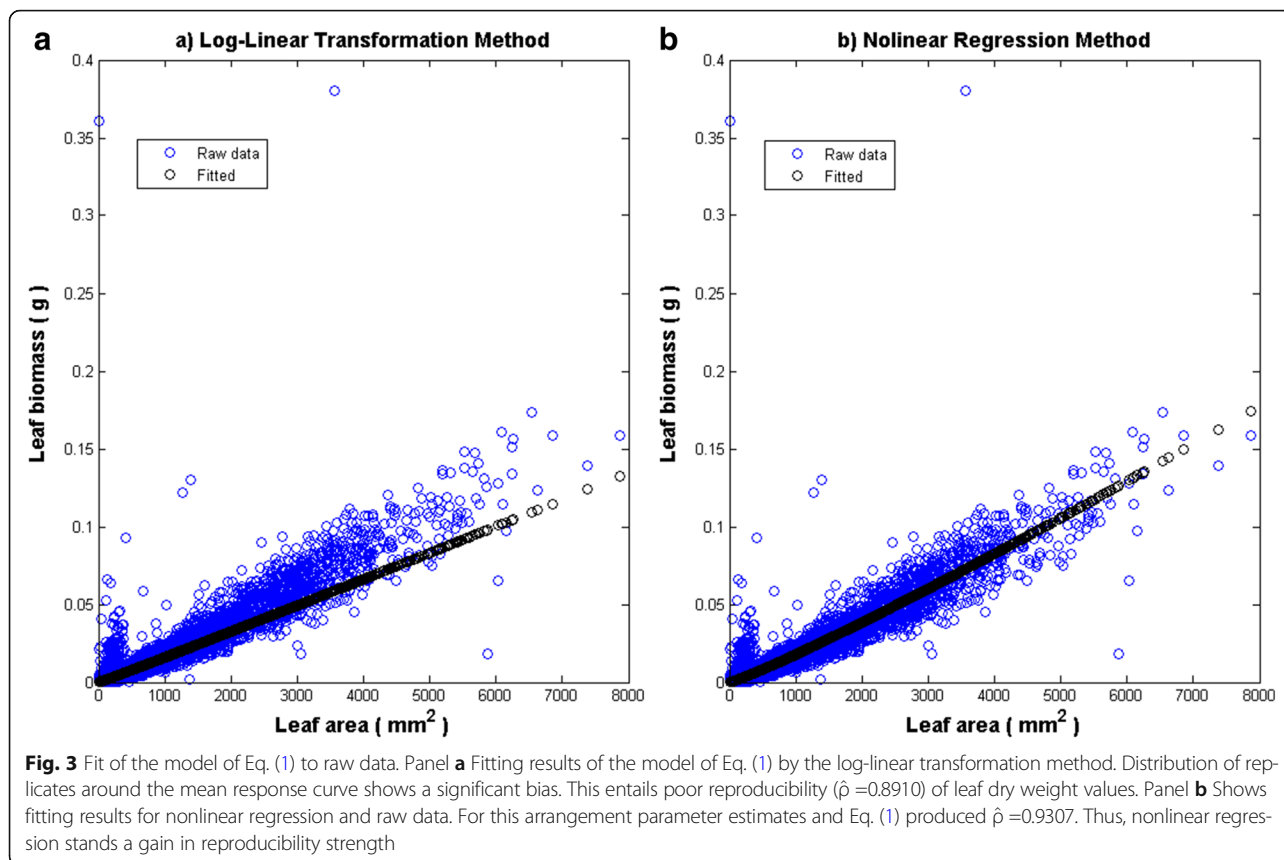
Processed data and log-linear transformation analysis produced  $\hat{\rho} = 0.9489$  for  $w_m(\alpha, \beta, t)$  projections. This figure is bigger than corresponding to raw data for this method. Nevertheless, lines in Fig. 6a show that this result does not correspond to a real gain in reproducibility. Besides data quality could not significantly reduce calculated root mean squared deviation (Table 2). Also, a value of 0.8588 for a ratio of projected to observed leaf dry weight averages is still low for suitable agreement (Table 2). Thus, regardless data quality, log-linear analysis failed to produce consistent  $w_m(\alpha, \beta, t)$  projections of  $w_m(t)$  averages.

In turn,  $w_m(\alpha, \beta, t)$  projections made by nonlinear regression and processed data yield the highest value of  $\hat{\rho} = 0.9976$ . (Table 2). And also the smallest root mean squared deviation among analysis method-data set combinations (Table 2). As shown by Fig. 6b this corresponds to a fairly good reproducibility strength. Additionally, data quality and nonlinear regression led to an outstanding value of 0.9975 for the ratio of projected  $\langle w_m(\alpha, \beta, t) \rangle$  to observed  $\langle w_m(t) \rangle$  averages.

Results exhibit that log-linear transformations failed to produce consistent projections for observed  $w_m(t)$  averages. In contraposition, nonlinear regression entailed

**Table 1** Parameter estimates  $\hat{\alpha}$  and  $\hat{\beta}$  associated standard errors ( $ste(\hat{\alpha}), ste(\hat{\beta})$ ) found by fitting the model of Eq. (1). Nonlinear regression estimates associate to the homoscedastic case of the model of Eqs. (5) and (6) (see Appendix 1). Values of  $\hat{\rho}$  give an evaluation of reproducibility strength of the proxy of Eq. (1)

Analysis method Data	$\hat{\beta}$	$ste(\hat{\beta})$	$\hat{\alpha}$	$ste(\hat{\alpha})$	$\hat{\rho}$
Log-linear Transformation Raw	$1.3674 \times 10^{-5}$	$2.9355 \times 10^{-7}$	1.023	$3.662 \times 10^{-3}$	0.8910
Nonlinear Regression Raw	$8.718 \times 10^{-6}$	$3.530 \times 10^{-7}$	1.104	$5.101 \times 10^{-3}$	0.9307
Log-linear Transformation Processed	$1.142 \times 10^{-5}$	$2.0831 \times 10^{-7}$	1.046	$3.035 \times 10^{-3}$	0.9455
Nonlinear Regression Processed	$6.956 \times 10^{-6}$	$2.200 \times 10^{-7}$	1.132	$3.954 \times 10^{-3}$	0.9777



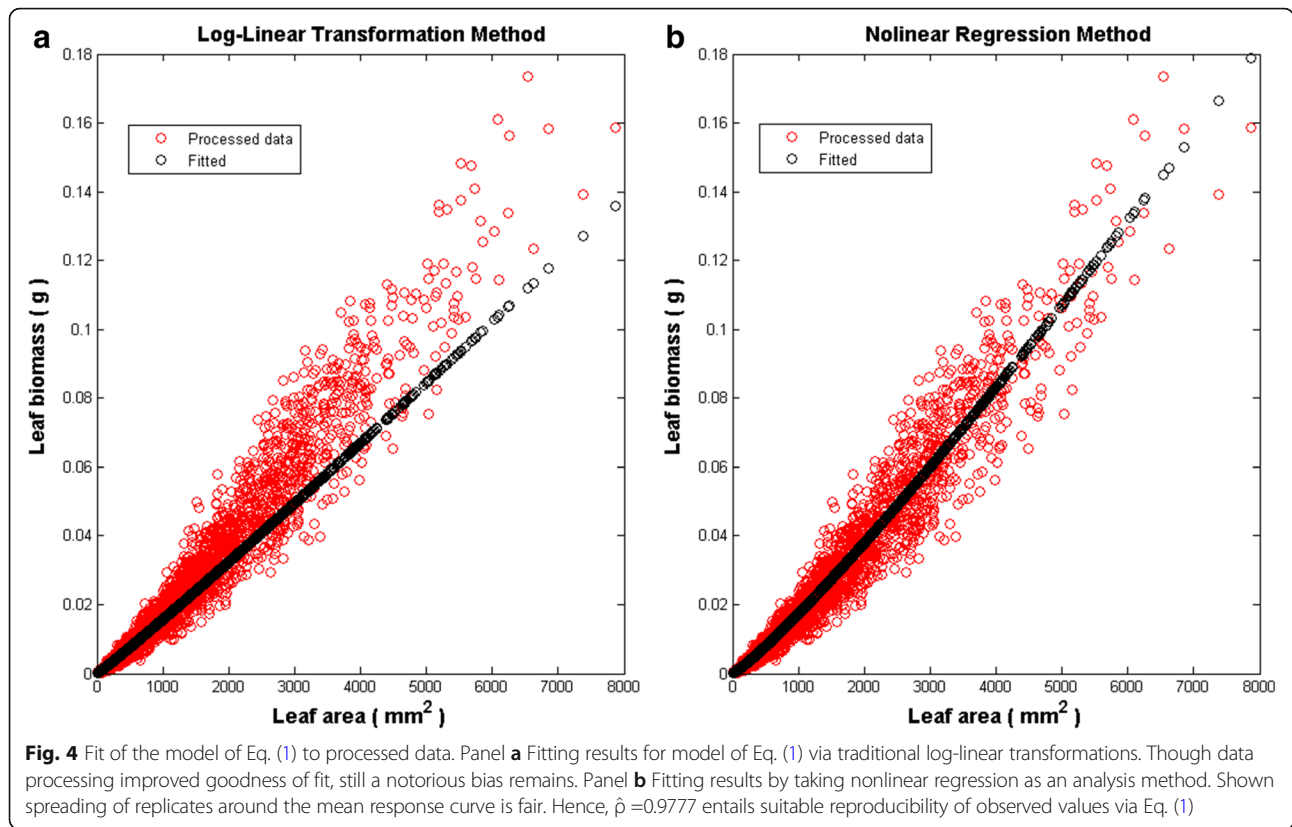
parameter estimates of noteworthy reliability. Thus studying sample size effects on reproducibility addressed only this analysis method. Figure 7 exhibits variation of CCC values depending on sample size  $k$ . This is expressed as a percentage of the extent of data set ( $n_{ra} = 10412$  for raw) or ( $n_{qa} = 6094$  for processed). For raw data, a sample sized  $k = 0.20n_{ra}$  endures reasonable reproducibility ( $\hat{\rho}(k) = 0.9889$ ). But samples beyond this threshold would not induce a significant change in reproducibility. Meanwhile, for the quality controlled data, the sample size threshold for excellent reproducibility was  $k = 0.10n_{rq}$ . This leads to a high reproducibility strength of  $\hat{\rho}(k) = 0.9929$ . Thus, a sample 10% the size of processed data set produced excellent reproducibility. Again sample sizes beyond this threshold failed to increase  $\hat{\rho}(k)$  values.

In Appendix 3 we consider variations  $\alpha_q = \alpha + \Delta\alpha_q$  and  $\beta_r = \beta + \Delta\beta_r$ . We found that shifting trajectories,  $w_m(\alpha_q, \beta_r, t)$  overestimated reference projections  $w_m(\alpha, \beta, t)$  whenever  $\Delta\alpha_q > 0$  and  $\Delta\beta_r > 0$ . In correspondence underestimation of  $w_m(\alpha, \beta, t)$  occurs for  $-\alpha < \Delta\alpha_q < 0$  and  $-\beta < \Delta\beta_r < 0$ . Fig. 8 explains that for  $\Delta\alpha_q \cdot \Delta\beta_r > 0$ , shifting trajectories overestimate (see red lines in panel a) or underestimate the reference one (see blue lines in panel a)). We can also make certain that relatively smaller deviations between  $w_m(\alpha_q, \beta_r, t)$  and  $w_m(\alpha, \beta, t)$ , values

occur for the case,  $\Delta\alpha_q \cdot \Delta\beta_r < 0$ , (see red and blue lines in panel b)).

The simulation code of Eqs. (39) through (44) explored the sensitivity of the  $w_m(\alpha, \beta, t)$  projection method, to numerical variation of parameters  $\alpha$  and  $\beta$ . Available parameter estimates yield reference values for  $\alpha$  and  $\beta$  (Table 2). Again, since nonlinear regression associates to highest reproducibility strength, for easier presentation, we only explain results using this analysis method.

The variation of the absolute deviation index for raw data is shown in Fig. 9. A range  $0 \leq \theta(q, r) \leq \theta_{ste}$  places the relative  $\vartheta(\theta)$  deviation index within the domain  $0 \leq \vartheta(\theta) \leq 0.0205$  (Table 3). Therefore, for a bound of  $\theta(q, r)$  set by the standard errors of estimates largest absolute deviation between  $w_m(\alpha_q, \beta_r, t)$  and  $w_m(t)$  amounts to about 2% of  $\langle w_m(t) \rangle$ . Moreover, a range  $0 \leq \theta(q, r) \leq 2\theta_{ste}$  produces  $0 \leq \vartheta(\theta) \leq 0.031$ . This leads to an equivalent 3% of  $\langle w_m(t) \rangle$ . Figure 10 displays the dynamics of  $\vartheta(\theta)$  depending on  $\theta(q, r)$  for processed data. We have that  $\theta(q, r)$  varying in a range of  $0 \leq \theta(q, r) \leq \theta_{ste}$  implies  $0 \leq \vartheta(\theta) \leq 0.003$  (Table 3). This time largest absolute deviation was only 0.03% of  $\langle w_m(t) \rangle$ . Comparing with results for raw data, we ascertained remarkable gain in precision of  $w_m(\alpha, \beta, t)$  projections. This exploration highlights on importance of data quality control as a



procedure leading the consistent  $w_m(\alpha, \beta, t)$  projections. But, results show that the projection method is robust even when parameter estimates are obtained from raw data.

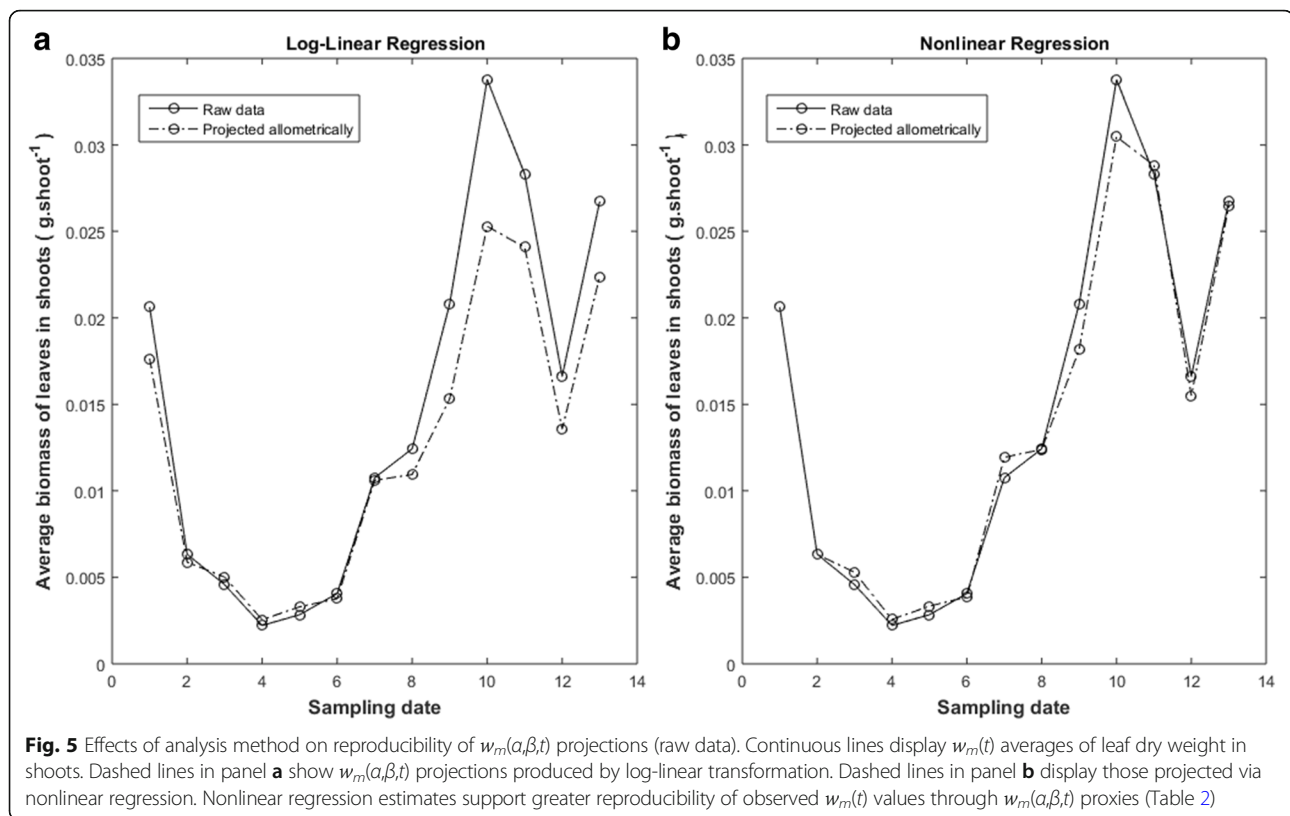
**Discussion**

Results of Solana-Arellano et al. [31] explain invariance of the allometric parameters  $\alpha$  and  $\beta$  in Eq. (1). This suggest  $w_m(\alpha, \beta, t)$  proxies as possible nondestructive estimations of the average leaf dry weight in eelgrass shoots. These assessments are essential for monitoring the efficiency of transplanted eelgrass plots, fundamental in remediation aims. The present examination shows that the  $w_m(\alpha, \beta, t)$  proxies could in fact offer reliable and cost-effective assessments. This on condition that practitioners take in to account our guidelines. For instance, our results typify the extent on which accuracy of estimates of the parameters  $\alpha$  and  $\beta$  influences the predictive power of the  $w_m(\alpha, \beta, t)$  projections.

And, our findings clarify that there are methodological factors affecting the accuracy of estimates. Related influences could spread significant bias in approximations supported by the  $w_m(\alpha, \beta, t)$  device. Indeed, analysis method turned into a main factor inducing bias in parameter estimates of the model of Eq. (1). Moreover, only parameter estimates acquired by nonlinear regression yield consistency of the model of Eq. (1) (Table 1 and lines in Fig. 3b and Fig. 4b). And, only these estimates upheld conclusive predictive power of the  $w_m(\alpha, \beta, t)$  proxies (Table 2, as well as, lines in Fig. 5b and Fig. 6b). Our results also show that data quality could not improve the performance of  $w_m(\alpha, \beta, t)$  projections acquired via log-linear transformations. Without doubt, parameter estimates acquired from processed data by this method still led to significant bias in  $w_m(\alpha, \beta, t)$  projections (Fig. 6a). Meanwhile, data processing improved reproducibility of projections built for raw data using nonlinear regression (Table 2 and lines in Fig. 5b and Fig. 6b).

**Table 2** Reproducibility results for  $w_m(\alpha, \beta, t)$ . Entries include, Lin’s concordance correlation coefficients ( $\hat{\rho}$ ), root mean square deviations (*rms*) and ratios of  $\langle w_m(\alpha, \beta, t) \rangle$  to  $\langle w_m(t) \rangle$  averages

Analysis method Data	$\langle w_m(\alpha, \beta, t) \rangle / \langle w_m(t) \rangle$	$\hat{\rho}$	<i>rms</i>
Log-linear Transformation Raw	0.8436	0.9285	0.01265
Nonlinear Regression Raw	0.9773	0.9915	0.00460
Log-linear Transformation Processed	0.8588	0.9489	0.01264
Nonlinear Regression Processed	0.9975	0.9976	0.00293



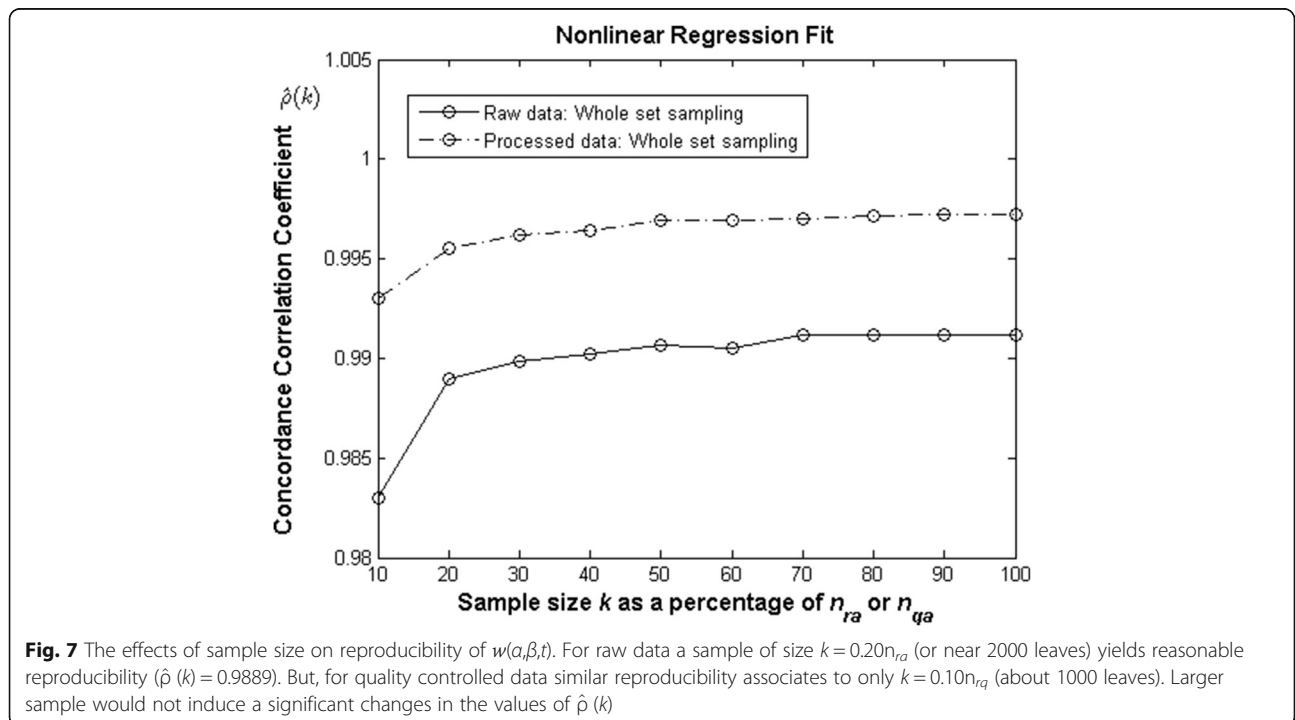
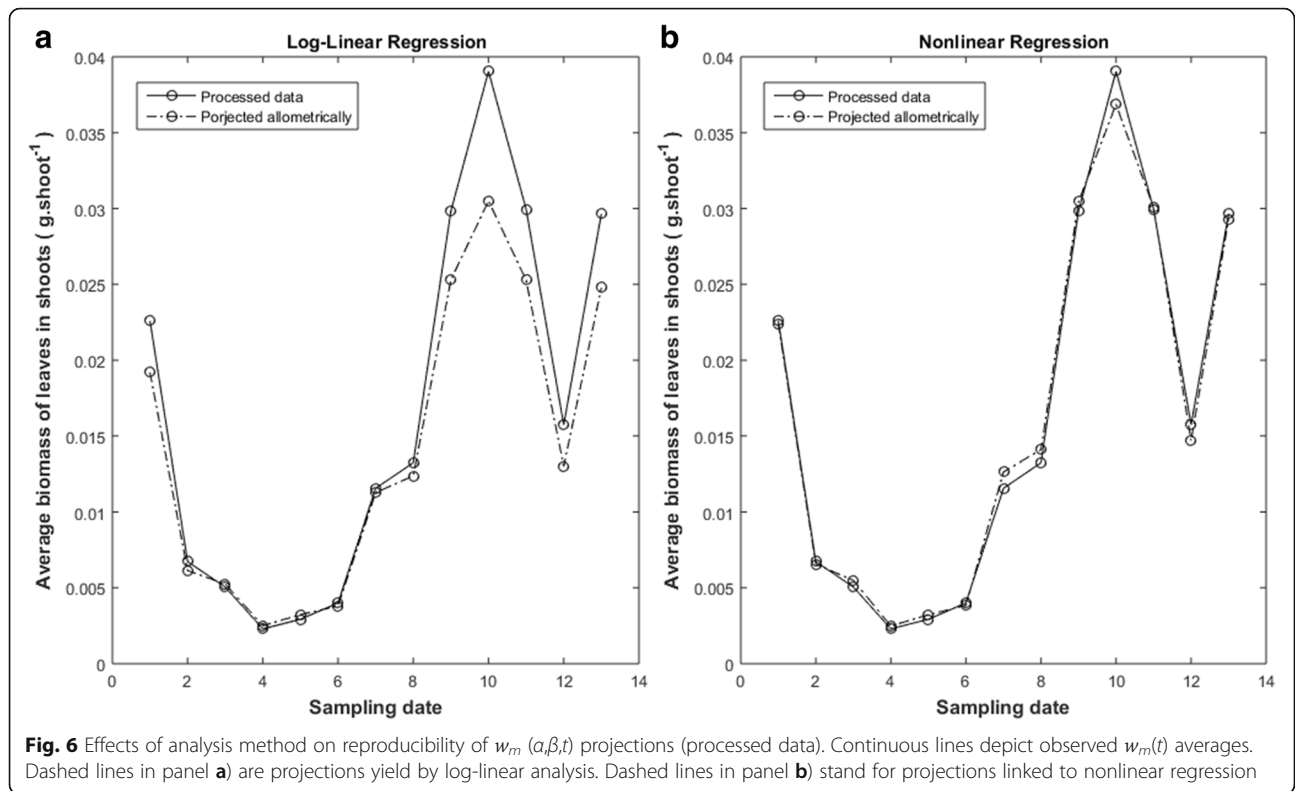
Besides, relevance of data quality was also evident for optimizing sample size. Indeed, while for raw data, a sample of approximately 2000 leaves shows reasonable reproducibility, for the quality controlled data this threshold drops to near 1000. However, samples sized beyond these thresholds would not induce a noteworthy gain in reproducibility. This result on its own, ties to efficiency of the  $w_m(\alpha, \beta, t)$  projection method. Undoubtedly routines for leaf dry weight assessment are tedious and time consuming. So, reducing size of data set for parameter estimation increases cost-effectiveness of the  $w_m(\alpha, \beta, t)$  projection method.

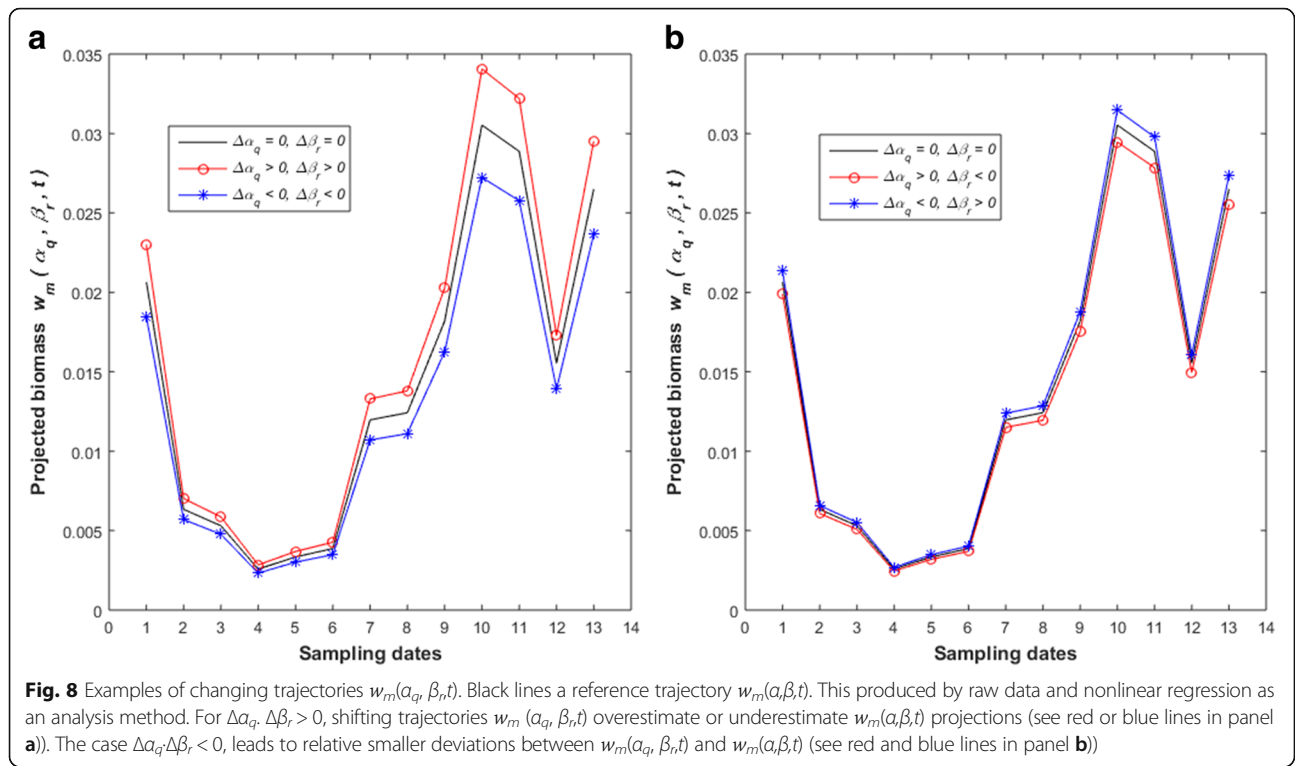
Nonlinear regression estimation also showed advantages in sensitivity over the log-linear analysis counterpart. Estimates from raw data led to a largest absolute deviation between  $w_m(\alpha, \beta, t)$  and  $w_m(t)$  values amounting only 3% of the average of  $w_m(t)$  over sampling dates. And, for processed data, the fluctuation range for equivalent sensitivity widened to 2.8 times the range for standard errors of estimates. But, on spite of data quality relevance, sensitivity results for raw data reveal that the  $w_m(\alpha, \beta, t)$  projection method is robust relative to expected fluctuations in parameter estimates.

Our results show that both the accuracy and cost-effectiveness of projections can be enhanced by the addition of data quality control procedures. However, including data processing can become a weakness for the  $w_m(\alpha, \beta, t)$  projection method. Indeed, data cleaning

procedures convey niceties that relate in a fundamental way to detection and rejection of inconsistent replicates. Also, compromising about which particular rejection edge should work, is hard to determine. Thus, the use of any data processing will endure a doubt, that the examiner selects an arrangement producing the most probable results [37]. In that order of ideas, when attempting to enhance the reproducibility power of  $w_m(\alpha, \beta, t)$  projections it is desirable to avoid depending in any form of data processing. For that aim, prior to data assembly, we must bear in mind standardized routines yielding accurate measurements for  $w(t)$  and  $a(t)$ . This will favor direct identification of the model of Eq. (1) in a consistent way. It is of a fundamental importance to be aware, that errors in leaf dry weight or area assessment differentiate in terms of leaf size. Certainly, leaves produced anew normally present a complete and undamaged span. But, they normally yield very reduced dry weight values. Therefore, we can expect estimation errors imputable to the precision of the analytical scale for individual leaf dry weight assessments. To this, we may add errors in the reading and/or recording of the scale output. These issues could explain a perceptible accumulation of inconsistent replicates for leaves with areas between 2 mm<sup>2</sup> and 350 mm<sup>2</sup> (Fig. 1). And, even after data cleaning procedures, leaf dry weight replicate spread for leaves bigger than 2000 mm<sup>2</sup> shows significant residual variability (Fig. 2). Likewise, as far as, bigger and



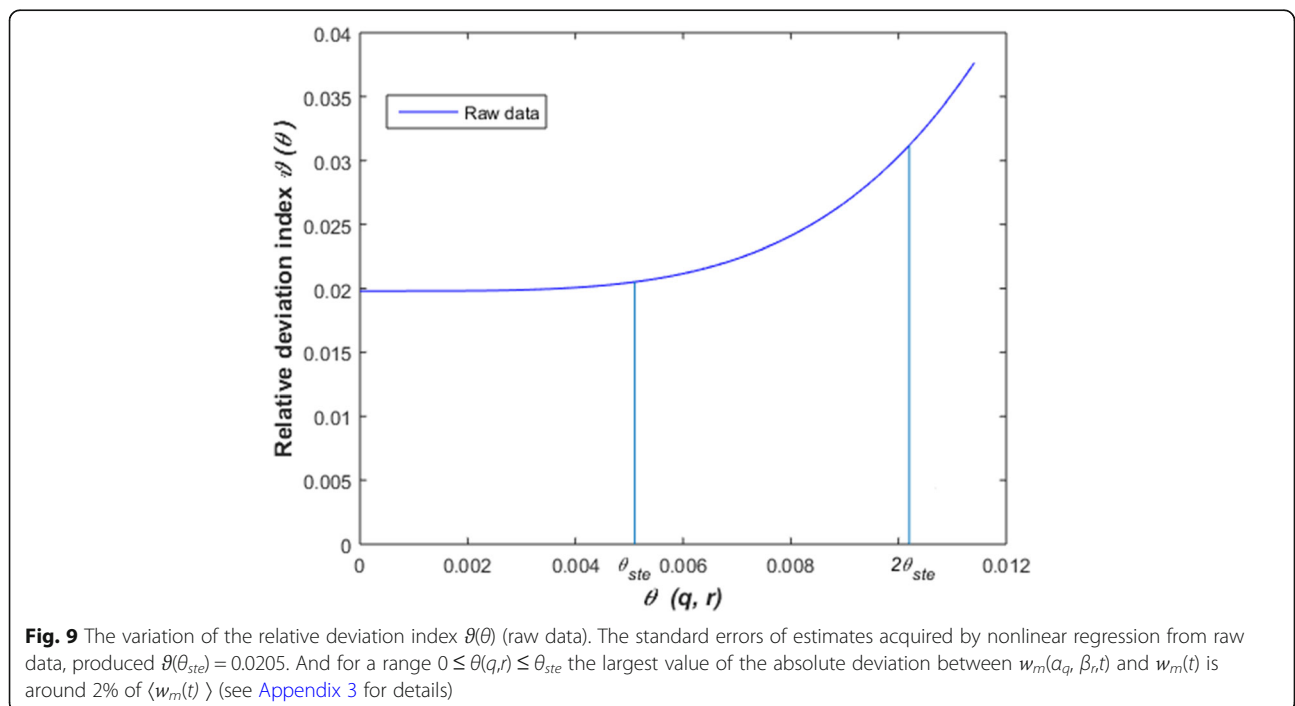




older leaves is concerned, there are issues on dry weight estimation errors. These seem to relate to damage caused by exposure to environmental factors. The fact that we estimated leaf area by means of the product of related length and width could explain these effects. For complete undamaged eelgrass leaves, the use of a leaf times width

proxy grants simplified and accurate estimations of leaf area [36].

But, this approach could deliver inaccurate estimations for long and damaged leaves. Actually, bigger leaves remain exposed during significant periods of time to environmental influences such as drag forces or herbivory.



**Table 3** Sensitivity of the  $w_m(\alpha, \beta, t)$  projections to changes in estimates of the parameters  $\alpha$  and  $\beta$ . Included are calculated  $\theta_{ste}$  values. This gives  $\theta(q, r)$  as determined by the standard errors of estimates. We also present corresponding values of the relative deviation index  $\vartheta(\theta_{ste})$

Analysis method Data	$\theta_{ste}$	$\vartheta(\theta_{ste})$ .
Log-Linear transformation Raw	$3.662 \times 10^{-3}$	0.1598
Nonlinear regression Raw	$5.101 \times 10^{-3}$	0.0205
Log-Linear transformation Processed	$3.035 \times 10^{-3}$	0.1419
Nonlinear regression Processed	$3.954 \times 10^{-3}$	0.003

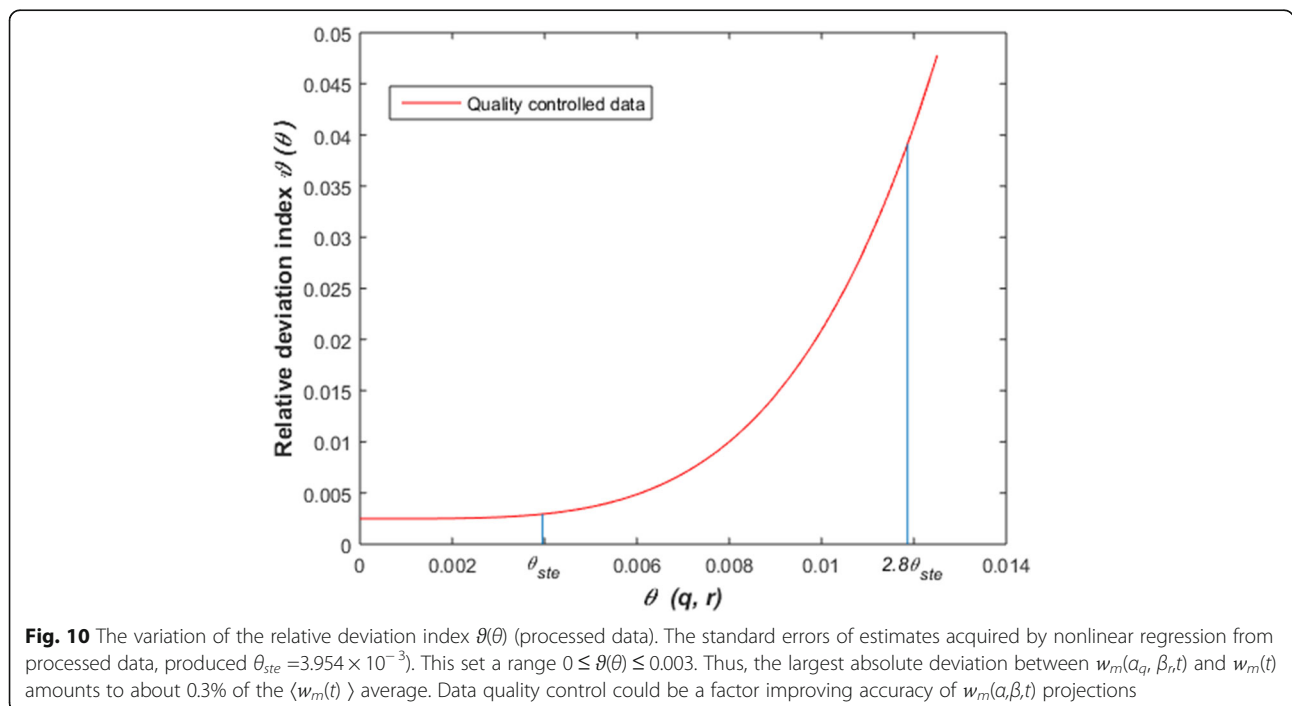
This could remove large amounts of leaf tissue while leaving length unaffected. Thus, causing overestimation of true leaf area when using a width and length product estimation. At the same time, lost portions of a leaf will produce a smaller dry weight than expected for an overestimated area. These effects will bring dry weight replicates that deviate from the power function-like trend associated to the model of Eq. (1). Estimation bias for the dry weights of smaller and longer leaves could explain the anomalous proliferation of inconsistencies (around 41 %) found while applying the proposed data cleaning procedure to the present raw data. These effects will propagate uncertainty of parameter estimates of the model of Eq. (1), influencing accuracy of the  $w_m(\alpha, \beta, t)$  projections. Hence, for the sake of consistent reproducibility of observed values via  $w_m(\alpha, \beta, t)$  projections, we need to be aware of these effects. And as elaborated, a good starting point for granting consistency, is appropriate gathering of primary data for

the identification of the model of Eq. (1). This will make subsequent data cleaning procedures unnecessary.

**Conclusion**

This research show that precise estimates of allometric parameters in Eq. (1) grant accurate non-destructive projections of the average leaf dry weight in eelgrass shoots. These projections offer efficient appraisal of eelgrass restoration projects, thereby contributing to the conservation of this important seagrass species. Our findings support views in Hui and Jackson [32], Packard and Birchard [33] and Packard et al. [44], on the relevance of analysis method in scaling studies. Indeed, we found that for assuring suitability of the  $w_m(\alpha, \beta, t)$  proxies, the use of nonlinear regression is crucial. On spite of claims that the use of the traditional log-linear analysis method is a must in allometric examination [45], exploration of spread of present crude data reveals curvature [46]. This explaining failure of the traditional analysis method to produce unbiased results for the present data. Besides proxies supported by nonlinear regression and raw data, are robust.

Data cleaning could only marginally enhance the accuracy of proxies produced by nonlinear regression and raw data. But results underline a relevant influence of data quality in setting optimal sample size for a suitable precision of parameter estimates. Nevertheless, the use of data cleaning procedures leads to intricacies. They in a fundamental way relate to choosing thresholds for rejection of detected inconsistencies, which are often regarded as subjective. Thus, instead of using later data cleaning, data gathering should seek for suitability. Special care must be taken when



**Table 4** Residual statistics for the fitting of model of Eq. (4)

Minimum	1Q	Median	3Q	Maximum
-4.7535	-0.2642	0.0042	0.2151	8.3509

processing bigger and older leaves. These are often damaged or even trimmed so that their dry weights do not conform to a true weight to area relationship. Irregularities in raw data may also associate to biased estimation of leaf length or width. Moreover, in a lesser way faulty equipment for leaf dry weight assessment, rounding off, or even incorrect data recording could as well contribute.

Taking advantage of a time invariance of the parameters in Eq. (1) the  $w_m(\alpha, \beta, t)$  device could offer to the eelgrass conservation practitioner highly consistent and truly nondestructive assessments of the average value of leaf dry weight in shoots. But the explained guidelines on analysis method, sample size and data appropriateness are mandatory for cost-effectiveness. Moreover, the present results suggest that the use of the  $w_m(\alpha, \beta, t)$  method could be extended to other seagrasses species with similar leaf architecture to eelgrass.

**Appendix 1**

**Model selection**

This Appendix covers the selection problem of Eq. (1) as a model for the variation of eelgrass leaf dry weight in terms of related area. On first sight the spread in Fig. 1 may suggest that instead of clinging to the power function model of Eq. (1) we should explore fitting an alternative linear model. There are two possible ways to achieve a linear model fit for the present data. One by fitting an isometric model. This follows by setting  $\alpha = 1$  in Eq. (1) to yield a linear model with  $w$ -intercept through origin, that is,

$$w(t) = \beta a(t) \tag{2}$$

A second choice is considering a standard linear equation with intercept different from zero, that is,

$$w(t) = \beta a(t) + \delta, \tag{3}$$

where  $\beta$  is the slope and  $\delta$  the intercept. But, by looking at the spread in Fig. 1, we realize that this model turns out to be inconsistent. This because for a vanishing leaf area we expect a vanishing dry weight. In any event, we fit the model of Eq. (3) as a device to analyze the suitability of the model of Eq. (2).

**Table 5** Fitting results for the model of Eq. (4)

Parameters	Estimate	Std. Error	t value	Pr(> t )	Confidence Interval (95%)
$\alpha$	1.022775	0.003662	279.3	<2e-16	(1.015597, 1.029953)
$\ln\beta$	-11.202199	0.021515	-520.7	<2e-16	(- 11.24437, - 11.16003)

**Fitting of the model of Eq. (1) by means the traditional analysis method of allometry**

For this fit we used raw data. Applying a logarithmic transformation on both sides of the Eq. (1), since we have available data pairs  $(w_i, a_i)$ , we get

$$\ln(w_i) = \ln(\beta) + \alpha \ln(a_i) + \varepsilon_i, i = 1, 2, \dots, n, \tag{4}$$

where the additive errors  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$  are independent and identically distributed, with distribution  $N(0, \sigma^2)$ . Fitting the model of Eq. (4) to data produced entries in Tables 4, 5 and 6.

Residual and normal probability plots for the fit of the model of Eq. (4) appear in Fig. 11. We notice that a value  $\alpha = 1.0$  is not included in the confidence interval for the parameter  $\alpha$ . This impairs Eq. (2) as a consistent model for the present raw data.

**Fitting of the model of Eq. (1) via non-linear regression**

For the aim of fitting the model of Eq. (1) to raw data using nonlinear regression, we use the equation

$$w_i = \beta a_i^\alpha + \varepsilon_i \tag{5}$$

the error term  $\varepsilon_i$  has a normal distribution with mean  $\mu$  zero and standard deviation  $\sigma_i$ . This chosen to contemplate heteroscedasticity or homoscedasticity. Moreover,  $\sigma_i$  depends on the predictor variable  $a_i$  through the model

$$\sigma_i = \gamma_0 + \gamma_1 a_i, \tag{6}$$

where  $\gamma_0$  and  $\gamma_1$  are parameters to be estimated. The case  $\gamma_1$  different from zero associates to heteroscedasticity. Whenever  $\gamma_1$  vanishes allows the consideration of homoscedasticity.

The response variable  $w_i$  resulting from Eq. (5) is normally distributed, with mean  $\mu_i = \beta a_i^\alpha$  and standard deviation  $\sigma_i$  given by Eq. (6). Fitting this regression model to raw data relied on a likelihood approach, i.e., acquiring estimates for the parameters  $\beta, \alpha, \gamma_0$  and  $\gamma_1$ , which yield the largest value of the log likelihood function [40, 41]

$$l(\beta, \alpha, \gamma_0, \gamma_1) = -\frac{n}{2} \log(2\pi) - \sum_1^n \log(\sigma_i) - \frac{1}{2} \sum_1^n \left(\frac{w_i - \mu_i}{\sigma_i}\right)^2, \tag{7}$$

with  $\mu_i = \beta a_i^\alpha$  and  $\sigma_i$  given by Eq. (6).

Estimates  $\hat{\beta}$  and  $\hat{\alpha}$  are very similar for the homoscedastic and heteroscedastic cases of Eq. (5). Notice that a

**Table 6** Fitting tests for the model of Eq. (4)

Test	Value
Residual standard error	0.5723 on 10,410 degrees of freedom
Multiple R-squared	0.8823
Adjusted R-squared	0.8823
F-statistic	7.802e + 04 on 1 and 10,410 DF
p-value	< 2.2e-16

value  $\alpha = 1.0$  is not included in the confidence interval for the parameter  $\alpha$ . Thus, this fit also excludes the model of Eq. (2) (isometric case).

Confidence intervals in Table 7 show a wide overlap region. It follows that parameter estimates do not differ in a significant way. This is also supported by the estimated CCC value ( $\hat{\rho} = 0.9307$ ) that was the same for the heteroscedastic and homoscedastic cases. Likewise, plotting the response curves  $w_i = \beta a_i^\alpha$  on a dispersion diagram, reveals that curves estimated considering heteroscedasticity or homoscedasticity coincide (Fig. 12).

Figure 13 shows the scatter diagram of dry weight residuals against leaf area. It also displays the region bounded by lines for 95% confidence intervals (blue lines). Figure 14 displays raw data and response curve (median) for the heteroscedastic model. It includes curves for region comprising 95% of leaf dry weight values (blue lines).

Thus, predictive strength from the homoscedastic or heteroscedastic regression models are similar. This because assumed distributions of the error term, in both models are symmetric. What makes a difference is the form capturing the uncertainty pattern.

**Fitting of the linear model of Eq. (3)**

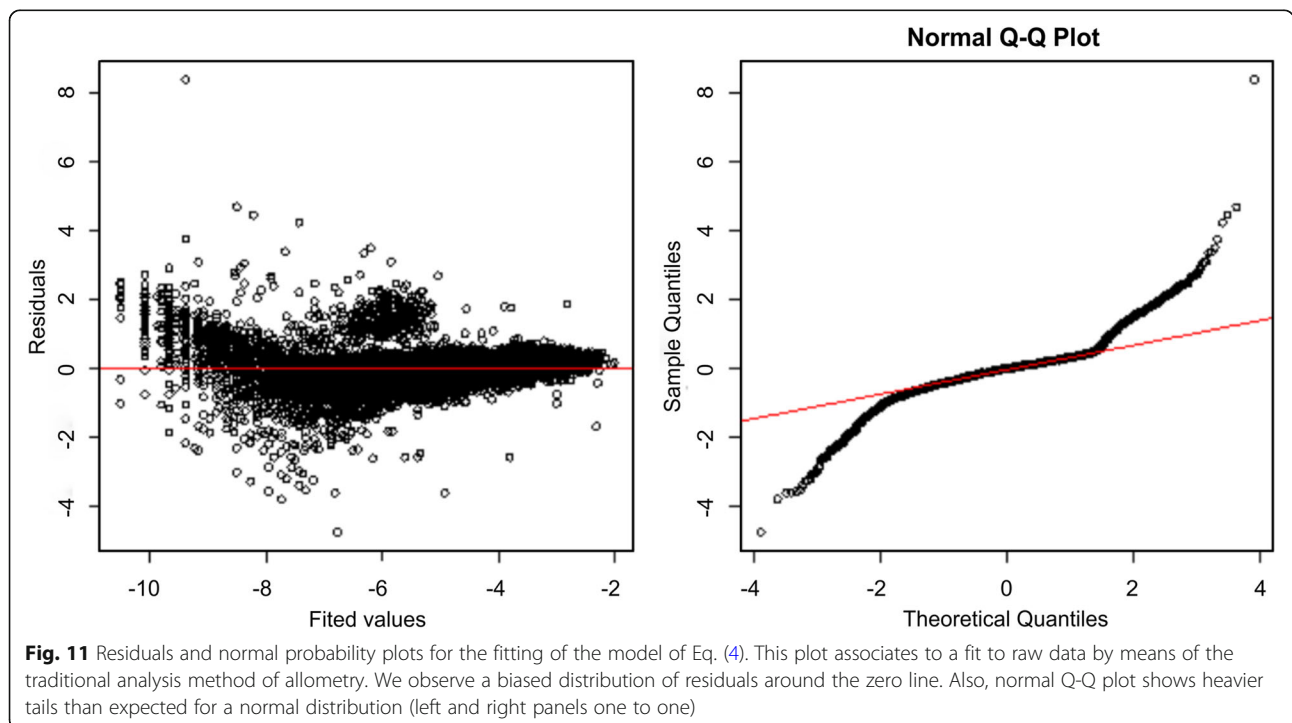
Since, raw data composes of pairs  $(w_i, a_i)$ , in order to test the consistency of the model of Eq. (3), we consider the regression equation

$$w_i = \beta a_i + \delta + \varepsilon_i, \tag{8}$$

where the additive errors  $\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n$  are independent and identically distributed  $N(0, \sigma^2)$ . Fitting regression Eq. (8) to raw data produced the results in Table (8), Table (9) and Table (10).

Residual and normal probability plot for this fit appear in Fig. 15. Looking at the confidence intervals in Table 7, we observe that a value  $\alpha = 1$  is not included. This implies rejection for the model of Eq. (2) (isometric case). Also, the intercept in Eq. (8) is negative, which is inconsistent. This is inconsistent with observations. Thus, we must reject the model of Eq. (3).

In summary, fitting results for the model of Eq. (1) excludes the value  $\alpha = 1$ . Now, for  $\alpha = 1$ , the model of Eq. (1) takes the form of a straight line through the origin (isometric model of Eq. (2)). In the other hand, for the fit of the linear model of Eq. (3) the intercept is negative.



**Table 7** Maximum likelihood estimates for parameters  $\beta$  and  $\alpha$  in Eq. (5). This fit considers heteroscedasticity ( $\gamma_1 \neq 0$ ) or homoscedasticity ( $\gamma_1 = 0$ )

Parameters	Estimate	Std. Error	Confidence Interval (95%)
Homoscedastic model			
$\alpha$	1.104	$5.101 \times 10^{-3}$	(1.094, 1.114)
$\beta$	$8.718 \times 10^{-6}$	$3.530 \times 10^{-7}$	( $8.052 \times 10^{-6}$ , $9.433 \times 10^{-6}$ )
Heteroscedastic model			
$\alpha$	1.107	$6.116 \times 10^{-3}$	(1.095, 1.119)
$\beta$	$8.477 \times 10^{-6}$	$3.950 \times 10^{-7}$	( $7.703 \times 10^{-6}$ , $9.251 \times 10^{-6}$ )

This means that this model could predict negative values of leaf dry weight. Thus, the model of Eq. (1) is the only one supporting a good selection criterion.

**Appendix 2**  
**Data processing**

This Appendix explains the Median Absolute Deviation (MAD) criterion for data cleaning. A first step is detecting the group of leaf dry weight replicates that associate to a given leaf area. The symbol  $G(a)$  stands for this group, and the number of replicates it contains denoted by means of the symbol  $n(G)$ . Formally,  $G(a) = \{w_i(a) \mid 1 \leq i \leq n(G)\}$ . The median of  $G(a)$  is denoted by means of  $MED(G(a))$ , that is,

$$MED(G(a)) = MED\{w_1(a), \dots, w_{n(G)}(a)\}. \tag{9}$$

And, for  $1 \leq i \leq n(G)$  the absolute deviation of replicate  $w_i(a)$  from the median  $MED(G(a))$  represented by means of  $\delta_i(a)$ , that is,

$$\delta_i(a) = |w_i(a) - MED(G(a))|. \tag{10}$$

Similarly, the median of the set of absolute deviations is signified by the symbol  $MED(\delta G(a))$ , that is,

$$MED(\delta G(a)) = MED\{\delta_1(a), \dots, \delta_{n(G)}(a)\}. \tag{11}$$

We use the character  $MAD(G(a))$  to denote median absolute deviation of a group  $G(a)$ . After Huber [47] and recalling that eelgrass leaf dry weight values are log-normally distributed, this is calculated through,

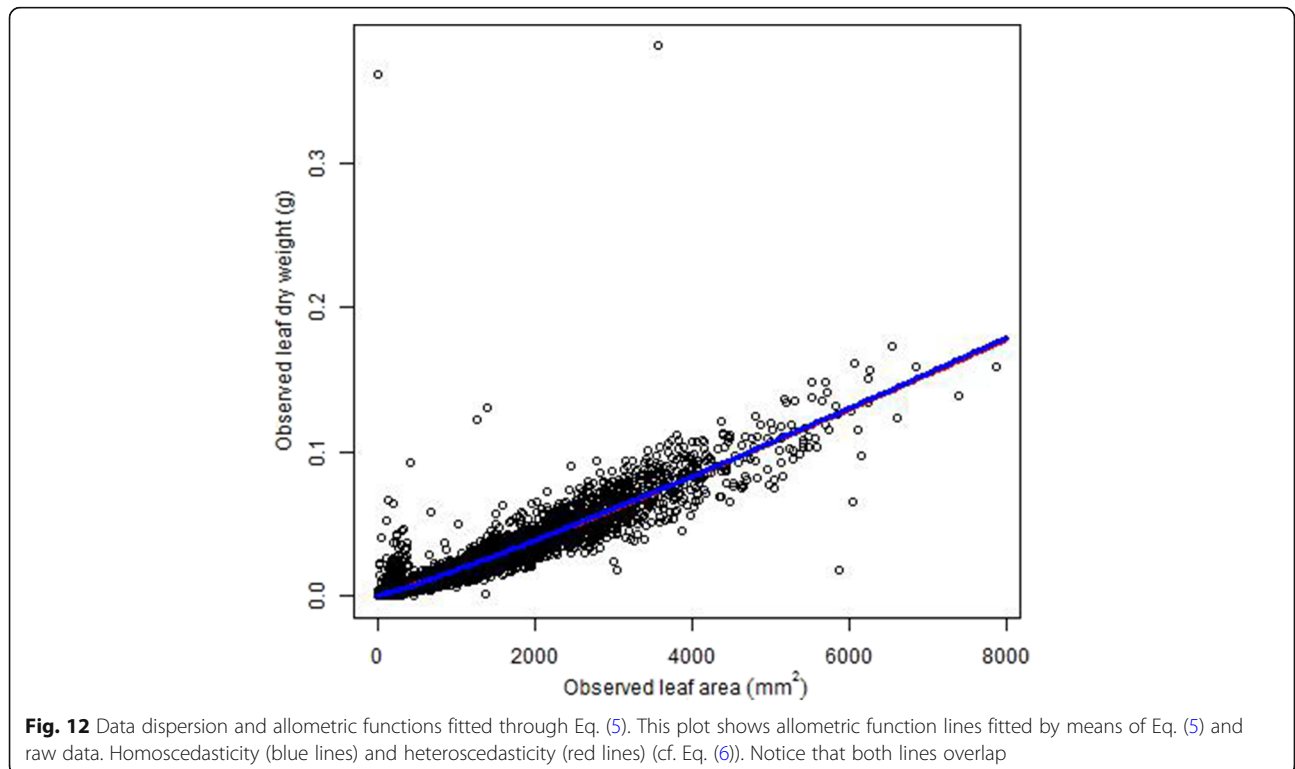
$$MAD(G(a)) = bMED\{\delta_1(a), \dots, \delta_{n(G)}(a)\}, \tag{12}$$

where  $b = 1/Q(0.75)$ , being  $Q(0.75)$  the 0.75 quantile of the lognormal distribution.

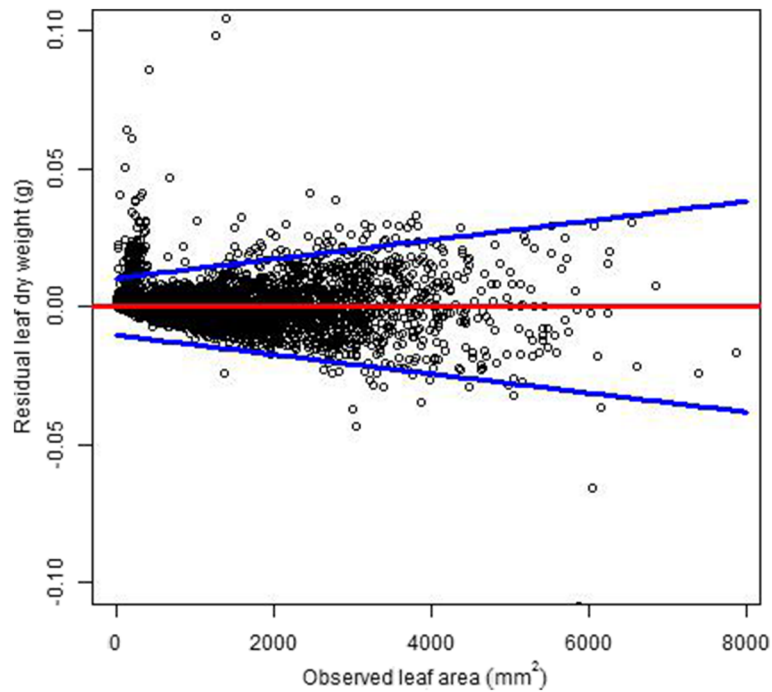
The removal of inconsistent replicates in a group  $G(a)$  follows the decision criterion

$$MED(G(a)) - T \cdot MAD(G(a)) < w_i(a) < MED(G(a)) + T \cdot MAD(G(a)), \tag{13}$$

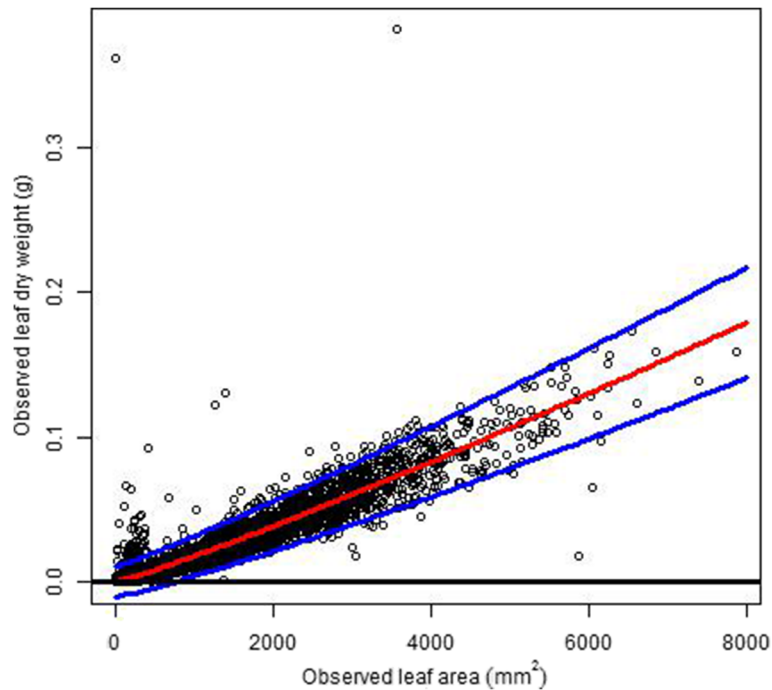
where  $T$  is the rejection threshold. Following Miller [48], we set  $T = 3$ . In what follows, the symbol  $n_{qa}$  stands for the size of the resulting quality controlled data set.



**Fig. 12** Data dispersion and allometric functions fitted through Eq. (5). This plot shows allometric function lines fitted by means of Eq. (5) and raw data. Homoscedasticity (blue lines) and heteroscedasticity (red lines) (cf. Eq. (6)). Notice that both lines overlap



**Fig. 13** Scatter diagram of residuals of Eq. (5). This shows residuals against leaf area that associate to the fit of the regression model of Eqs. (5) and (6) allowing heteroscedasticity. Blue lines bound region 95% confidence intervals



**Fig. 14** Dispersion around response curve fitted using Eq. (5) assuming heteroscedasticity. This displays leaf dry weight against leaf area in arithmetical scale and response curve (median) estimated by the heteroscedastic regression model of Eqs. (5) and (6) (red lines). Also shown is region comprising 95% of leaf dry weight values observed for a given value of leaf area

**Table 8** Residual statistics for the fit of Eq. (8)

Minimum	1Q	Median	3Q	Maximum
-0.1004000	-0.0017230	0.0000000	0.0010230	0.36210

**Appendix 3**

**Sensitivity of  $w_m(\alpha, \beta, t)$**

This Appendix presents a study of the influences of parameter fluctuations in the precision of the  $w_m(\alpha, \beta, t)$  projections. The examination is performed on the basis of both analytical and simulation approaches. Presentation requires a formal description of variables of interest. For that aim we label a generic eelgrass shoot by using a subscript  $s$ . Correspondingly, the characters  $nl(s)$  and  $w_s(t)$  shall stand for the associated number of leaves and their combined dry weight. Similarly,  $ns(t)$  denotes the number of shoots collected at a sampling date  $t$ . And, the number of sampling date symbolized by means of  $n(t)$ .

Observed dry weight in a shoot  $s$  is represented by mean of

$$w_s(t) = \sum_{nl(s)} w(t), \tag{14}$$

where  $\sum_{nl(s)}$  indicates summation of the leaves that the shoot  $s$ , holds. Meanwhile, the matching average leaf dry weight of shoots denoted by the symbol  $w_m(t)$  is given by

$$w_m(t) = \frac{\sum_{ns(t)} w_s(t)}{ns(t)}, \tag{15}$$

where  $\sum_{ns(t)}$  indicates summation of the  $ns(t)$  shoots collected at a time  $t$ .

The character  $\langle w_m(t) \rangle$  stands for the average of  $w_m(t)$  values over the number  $n(t)$  of sampling times. It is given by

$$\langle w_m(t) \rangle = \frac{\sum_{n(t)} w_m(t)}{n(t)}, \tag{16}$$

where  $\sum_{n(t)}$  indicates summation over the number of sampling times.

We assumed that  $w(t)$  and  $a(t)$  are linked through the scaling expression of Eq. (1). Then, we can obtain an allometric proxy for  $w_s(t)$ . The symbol  $w_s(\alpha, \beta, t)$  represents this surrogate and it is given by

$$w_s(\alpha, \beta, t) = \sum_{nl(s)} \beta a(t)^\alpha. \tag{17}$$

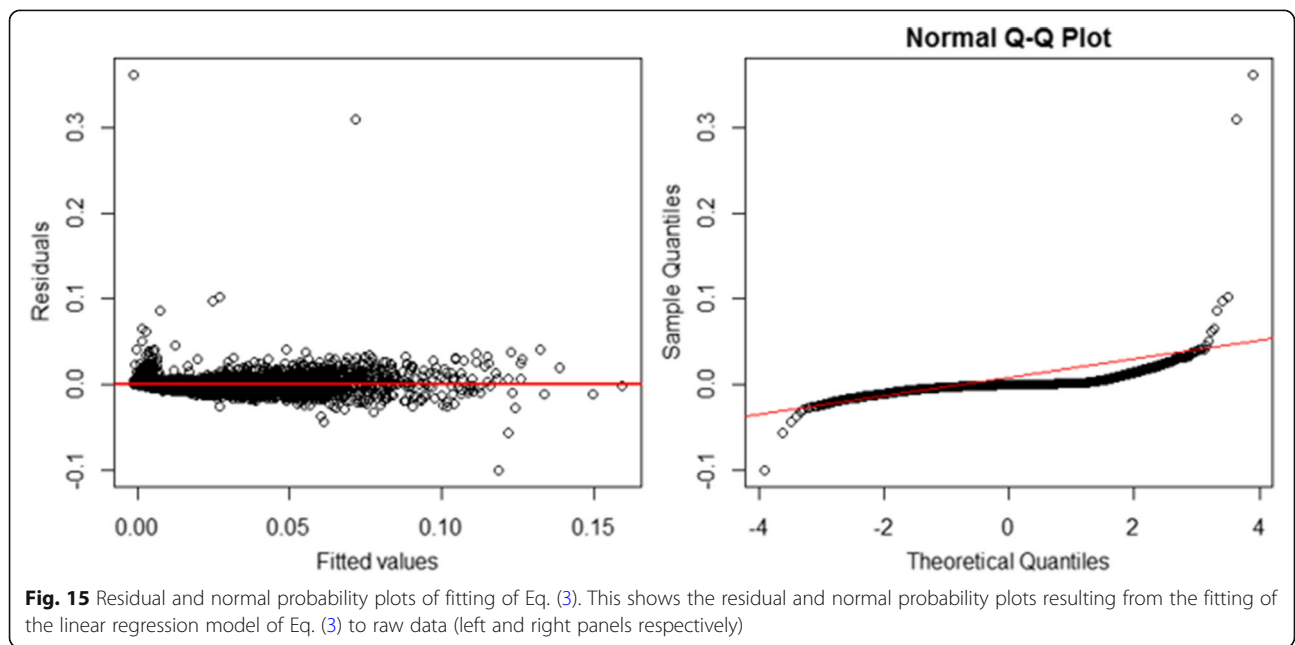
Moreover,

$$w_s(t) = w_s(\alpha, \beta, t) + \epsilon_s(\alpha, \beta, t), \tag{18}$$

that is,  $w_s(\alpha, \beta, t)$  is an approximation to the true value of  $w_s(t)$ , being  $\epsilon_s(\alpha, \beta, t)$  the error of estimation. Similarly, Eq. (17) produces an allometric proxy for  $w_m(t)$ , the average leaf dry weight in shoots at a time  $t$ . This surrogate, represented here by means of  $w_m(\alpha, \beta, t)$ , is given by

$$w_m(\alpha, \beta, t) = \frac{\sum_{ns(t)} w_s(\alpha, \beta, t)}{ns(t)}. \tag{19}$$

In turn, the average of  $w_m(\alpha, \beta, t)$  over the number  $n(t)$  of sampling times, denoted through  $\langle w_m(\alpha, \beta, t) \rangle$ , is calculated by means of





**Table 9** Fitting results for the model of Eq. (8)

Parameters	Estimate	Std. Error	t value	Pr(> t )	Confidence Interval (95%)
$\beta$	2.04e-005	7.838e-08	260.23	2e-16	(2.02e-005, 2.055e-005)
$\delta$	-0.001158	8.954e-05	-12.93	2e-16	(-0.001158, -0.0009825)

$$\langle w_m(\alpha, \beta, t) \rangle = \frac{\sum_{n(t)} w_m(\alpha, \beta, t)}{n(t)}. \tag{20}$$

Similarly,

$$w_m(t) = w_m(\alpha, \beta, t) + \epsilon_m(\alpha, \beta, t), \tag{21}$$

where  $\epsilon_m(\alpha, \beta, t)$  stands for the resulting approximating error. Moreover, the root mean squared deviation between  $w_m(t)$  and  $w_m(\alpha, \beta, t)$  is denoted by means of  $rms(\alpha, \beta)$  and given by

$$rms(\alpha, \beta) = \sqrt{\sum_{n(t)} \epsilon_m(\alpha, \beta, t)^2 / n(t)}, \tag{22}$$

again  $\sum_{n(t)}$  indicates summation over the sampling dates  $n(t)$ .

We assume changes in parameters  $\alpha$  and  $\beta$  given by factors of their standard errors. Then, a factor  $q$  associates to a shift of the value of the parameter  $\alpha$ . This deviation, denoted by means of the symbol  $\alpha_q$  is given by

$$\alpha_q = \alpha + \Delta\alpha_q, \tag{23}$$

where

$$\Delta\alpha_q = q \cdot ste(\alpha), \tag{24}$$

with  $ste(\alpha)$  symbolizing the standard error of  $\alpha$  and  $q$  satisfying.

$$|q| \leq q_{max}. \tag{25}$$

Similarly, a factor  $r$ , such that

$$|r| \leq r_{max}, \tag{26}$$

associates to a fluctuation in  $\beta$  denoted by means of  $\Delta\beta_r$  and expressed by way of

$$\Delta\beta_r = r \cdot ste(\beta), \tag{27}$$

being  $ste(\beta)$  the standard error of  $\beta$ . This returns a

**Table 10** Fitting test statistics for the model of Eq. (8)

Test	Value
Residual standard error	0.007279 on 10,410 degrees of freedom
Multiple R-squared	0.8668
Adjusted R-squared	0.8667
F-statistic	6.772e04 on 1 and 10,410 DF
p-value	2.2e-16

variation of the parameter  $\beta$ , denoted via  $\beta_r$  and represented in the form

$$\beta_r = \beta + \Delta\beta_r. \tag{28}$$

**Qualitative exploration of sensitivity of  $w_m(\alpha, \beta, t)$**

A qualitative study is intended to set domains where reference estimations  $w_m(\alpha, \beta, t)$  are underestimated or overestimated by changing values  $w_m(\alpha_q, \beta_r, t)$ . For that aim, we firstly set reasonable local variation ranges for the numerical deviations  $\alpha_q$  and  $\beta_r$ . Fittings of Eq. (1) to eelgrass leaf dry weight and area data demonstrate that the inequality  $0 < \beta < \alpha$  holds [30]. As stated by Eqs. (24) and (27), letting  $|q| < ste(\alpha)/\alpha$  along with  $|r| < ste(\beta)/\beta$  set domains  $|\Delta\alpha_q| \leq \alpha$  and  $|\Delta\beta_r| \leq \beta$ . This enables suitable domain to pursue a qualitative study of the effects of parametric changes in  $w_s(\alpha, \beta, t)$  (cf. Eq. (17)). For that aim, let's consider deviations  $\delta w_s(\alpha_q, \beta_r, t)$  defined through

$$\delta w_s(\alpha_q, \beta_r, t) = w_s(\alpha_q, \beta_r, t) - w_s(\alpha, \beta, t). \tag{29}$$

Then,

$$\delta w_s(\alpha_q, \beta_r, t) = \sum_{n(t)} \left( (\beta + \Delta\beta_r) a(t)^{\alpha + \Delta\alpha_q} - \beta a(t)^\alpha \right), \tag{30}$$

and factoring the term  $\beta a(t)^\alpha$

$$\delta w_s(\alpha_q, \beta_r, t) = \sum_{n(t)} \beta a(t)^\alpha \mu_a(\Delta\alpha_q, \Delta\beta_r, a(t)), \tag{31}$$

where

$$\mu_a(\Delta\alpha_q, \Delta\beta_r, a(t)) = \frac{\pi(\Delta\alpha_q)}{\phi(\Delta\beta_r)} - 1 \tag{32}$$

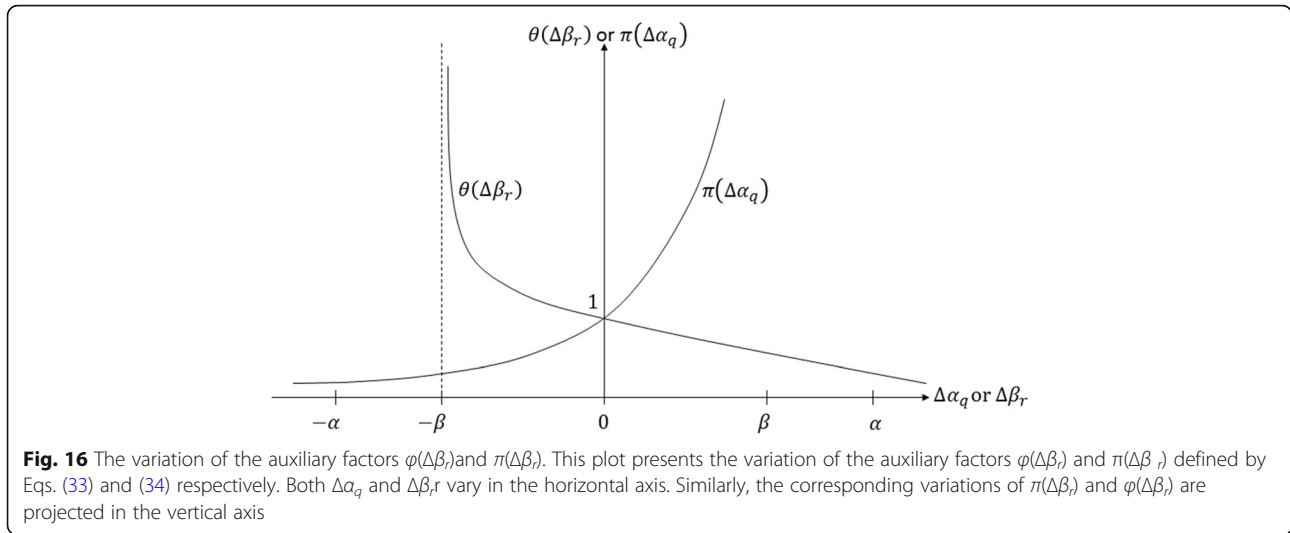
with

$$\phi(\Delta\beta_r) = \frac{\beta}{\beta + \Delta\beta_r} \tag{33}$$

and

$$\pi(\Delta\alpha_q) = a(t)^{\Delta\alpha_q}. \tag{34}$$

Now, since  $a(t)$  is positive, for fixed values of the parameters  $\alpha$  and  $\beta$  in the domain  $|\Delta\alpha_q| < \alpha$ , the factor  $\pi(\Delta\alpha_q)$  remains positive and increasing. Respectively, for  $\Delta\beta_r$  varying in the interval  $|\Delta\beta_r| < \beta$ , the factor  $\phi(\Delta\beta_r)$  is positive and decreasing. Furthermore, for  $\Delta\alpha_q = \Delta\beta_r = 0$ , we have  $\pi(\Delta\alpha_q) = \phi(\Delta\beta_r) = 1$  (Fig. 16).



**Fig. 16** The variation of the auxiliary factors  $\phi(\Delta\beta_r)$  and  $\pi(\Delta\alpha_q)$ . This plot presents the variation of the auxiliary factors  $\phi(\Delta\beta_r)$  and  $\pi(\Delta\alpha_q)$  defined by Eqs. (33) and (34) respectively. Both  $\Delta\alpha_q$  and  $\Delta\beta_r$  vary in the horizontal axis. Similarly, the corresponding variations of  $\pi(\Delta\beta_r)$  and  $\phi(\Delta\alpha_q)$  are projected in the vertical axis

From Eq. (31) the proxies  $w_s(\alpha, \beta, t)$  will be overestimated by shifting projections  $w_s(\alpha_q, \beta_r, t)$  whenever the inequality

$$\mu_a(\Delta\alpha_q, \Delta\beta_r, a(t)) > 0 \tag{35}$$

is satisfied. And from Eq. (32) this last inequality leads to

$$\pi(\Delta\alpha_q) > \phi(\Delta\beta_r). \tag{36}$$

But, for  $-\alpha \leq \Delta\alpha_q \leq 0$  the factor  $\pi(\Delta\alpha_q)$  is continuous and increases monotonically from a value  $\pi(-\alpha) > 0$ , until it reaches a value of  $\pi(0) = 1$ , while the factor  $\phi(\Delta\beta_r)$  takes positive and arbitrarily large values whenever  $\Delta\beta_r$  approaches  $-\beta$ , and decreases monotonically towards a value of  $\phi(0) = 1$ . Thus, the continuity of  $\phi(\Delta\beta_r)$  and  $\pi(\Delta\alpha_q)$  implies the order relationship  $\pi(\Delta\alpha_q) < \phi(\Delta\beta_r)$ , in the domain  $-\alpha \leq \Delta\alpha_q \leq 0$  and  $-\beta < \Delta\beta_r \leq 0$ . In the other hand in the domain  $0 \leq \Delta\alpha_q \leq \alpha$  and  $0 \leq \Delta\beta_r \leq \beta$ , both  $\pi(\Delta\alpha_q)$  and  $\phi(\Delta\beta_r)$  steer away from one being  $\pi(\Delta\alpha_q)$  increasing and  $\phi(\Delta\beta_r)$  decreasing. Therefore, by continuity the statement of inequality (36) will hold only in the domain  $\Delta\alpha_q > 0$  and  $\Delta\beta_r > 0$ .

Similarly, reference values  $w_s(\alpha, \beta, t)$  will be underestimated by fluctuating ones  $w_s(\alpha_q, \beta_r, t)$  whenever, the statement

$$\mu_a(\Delta\alpha_q, \Delta\beta_r, a(t)) < 0 \tag{37}$$

holds, or equivalently whenever

$$\pi(\Delta\alpha_q) < \phi(\Delta\beta_r), \tag{38}$$

which as it has been discussed above, only holds whenever the increments  $\Delta\alpha_q$  and  $\Delta\beta_r$  vary in

the domain  $-\beta < \Delta\beta_r < 0$  and  $-\alpha < \Delta\alpha_q < 0$ .

Since, the  $w_s(\alpha, \beta, t)$  projections are always positive, the  $w_m(\alpha, \beta, t)$  reference projections will be overestimated or underestimated by changing  $w_m(\alpha_q, \beta_r, t)$  values in the same domains of variation of  $\Delta\alpha_q$  and  $\Delta\beta_r$ , where the  $w_s(\alpha, \beta, t)$  projections are overestimated or underestimated by  $w_s(\alpha_q, \beta_r, t)$  values.

#### Exploration of sensitivity of $w_m(\alpha, \beta, t)$ by simulation methods

In order to perform a simulation study of sensitivity, we considered a combined range of variability for  $\alpha$  and  $\beta$  given by the modulus of the vector of parametric changes  $(\Delta\alpha_q, \Delta\beta_r)$ . This is denoted by means of the symbol  $\theta(q, r)$  and calculated through

$$\theta(q, r) = \sqrt{\Delta\alpha_q^2 + \Delta\beta_r^2}. \tag{39}$$

Equivalently, using Eqs. (24) and (27) yield

$$\theta(q, r) = \sqrt{(q \cdot ste(\alpha))^2 + (r \cdot ste(\beta))^2}. \tag{40}$$

And, setting  $q = 1$  and  $r = 1$  give  $(\Delta\alpha_q, \Delta\beta_r) = (ste(\alpha), ste(\beta))$ . This produces vectors of shifting parametric values with modulus determined by the standard errors of estimates. Denoting by means of the symbol  $\theta_{ste}$  the associated range of variability we have

$$\theta_{ste} = \sqrt{ste(\alpha)^2 + ste(\beta)^2}. \tag{41}$$

By virtue of Eq. (39), every fixed value of  $\theta(q, r)$  returns different pairs  $(\Delta\alpha_q, \Delta\beta_r)$  of increments, each one giving rise to a couple of shifting parameter values  $(\alpha_q, \beta_r)$ . Besides, for each ordered pair of changing parameters, leaf data in shoots and Eq. (19) return a shifting trajectory, that is denoted through the symbol  $w_m(\alpha_q, \beta_r, t)$ . The average of the values taken by the fluctuating trajectories at a

sampling date  $t$  is denoted by  $\langle w_m(\alpha_q, \beta_r, t) | \theta \rangle$ . The procedure estimates the deviations between the observed  $w_m(t)$  and the average  $\langle w_m(\alpha_q, \beta_r, t) | \theta \rangle$  trajectory for each time  $t$ . This is denoted by  $\delta w_{m\theta}(\alpha_q, \beta_r, t)$  and calculated by way of

$$\delta w_{m\theta}(\alpha_q, \beta_r, t) = w_m(t) - \langle w_m(\alpha_q, \beta_r, t) | \theta \rangle. \quad (42)$$

The average of the  $\delta w_{m\theta}(\alpha_q, \beta_r, t)$  deviation values over all sampling times  $n(t)$ , represented through the symbol  $\langle \delta w_{m\theta}(\alpha_q, \beta_r, t) \rangle$ , is given by

$$\langle \delta w_{m\theta}(\alpha_q, \beta_r, t) \rangle = \frac{\sum_{n(t)} \delta w_{m\theta}(\alpha_q, \beta_r, t)}{n(t)}, \quad (43)$$

where,  $\sum$  indicates summation over the involved  $n(t)$  sampling dates. These statistics along with, the average of  $w_m(t)$  values over the number  $n(t)$  of sampling times are used to produce a relative deviation index represented by means of the symbol  $\vartheta(\theta)$  and given by

$$\vartheta(\theta) = \frac{|\langle \delta w_{m\theta}(\alpha_q, \beta_r, t) \rangle|}{\langle w_m(t) \rangle}. \quad (44)$$

The value of  $\vartheta(\theta)$  stands a measure of the sensitivity of the reference trajectory  $w_m(\alpha, \beta, t)$  to a change of tolerance  $\theta(q, r)$  on the values of  $\alpha$  and  $\beta$ . Moreover,  $\vartheta(\theta)$  assess in what percentage of  $\langle w_m(t) \rangle$  the absolute deviation between  $w_m(\alpha_q, \beta_r, t)$  and  $w_m(t)$  amounts.

### Application to present data

The results of Appendix 1 show that there are not differences in parameter estimates of the nonlinear regression model of Eqs. (5) and (6) assuming heteroscedasticity or homoscedasticity. Therefore, in order to explore the sensitivity of  $w_m(\alpha, \beta, t)$  by means of the simulation code of Eqs. (39)–(44), we take as a reference parameter estimates acquired from the raw data using the homoscedastic case. Parameter estimates were  $\hat{\alpha} = 1.104$  with  $ste(\hat{\alpha}) = 5.101 \times 10^{-3}$  for  $\alpha$  and of  $\hat{\beta} = 8.71 \times 10^{-6}$  with  $ste(\hat{\beta}) = 3.53 \times 10^{-7}$  for  $\beta$  (Table 1). Then, letting  $q = 1$  and  $r = 1$  in Eqs. (24) and (27) allows to consider a combined range of parametric change set by the standard errors of estimates. Eq. (41) returned  $\theta_{ste} = 5.101 \times 10^{-3}$ . Afterwards, for every fixed value of  $\theta(q, r)$  we determined all ordered pairs  $(\Delta\alpha_q, \Delta\beta_r)$  complying with condition set by Eq. (39). This returned fluctuating values  $\alpha_q$  for  $\alpha$  and  $\beta_r$  for  $\beta$ , given one by one by Eqs. (23) and (28). Observed leaf area data and shifting parameter values, shaped the associated changing  $w_m(\alpha_q, \beta_r, t)$  trajectories. Moreover, while  $0 \leq \theta(q, r) \leq \theta_{ste}$  the values of the relative deviation index  $\vartheta(\tau)$  satisfied  $0 \leq \vartheta(\theta) \leq 0.0205$  (Fig. 9). Thus, for a parametric variation range set by the standard errors of estimates the maximum value of the absolute deviation  $\vartheta(\theta)$  between  $w_m(\alpha_q, \beta_r, t)$  and  $w_m(t)$  amounts to approximately 2% of the average of  $w_m(t)$  taken

over all sampling dates ( $\langle w_m(t) \rangle$  cf. Eq. (16)). Moreover, for  $\theta(q, r)$  in a range  $0 \leq \theta \leq 2\theta_{ste}$  we have  $0 \leq \vartheta(\theta) \leq 0.031$ . This leads to a maximum absolute deviation between  $w_m(\alpha_q, \beta_r, t)$  and  $w_m(t)$  of around 3% of  $\langle w_m(t) \rangle$ . This shows the robustness of the  $w_m(\alpha, \beta, t)$  projection method.

In turn, fitting the model of Eq. (1) to the quality controlled data using nonlinear regression, produced estimates,  $\hat{\alpha} = 1.132$  with  $ste(\hat{\alpha}) = 3.954 \times 10^{-3}$  for  $\alpha$  and of  $\hat{\beta} = 6.956 \times 10^{-6}$  with  $ste(\hat{\beta}) = 2.202 \times 10^{-7}$  for  $\beta$  (Table 1). Letting  $p = 1$  and  $q = 1$  in Eq. (40) gives  $\theta_{ste} = 3.954 \times 10^{-3}$ , therefore,  $0 \leq \theta(q, r) \leq \theta_{ste}$  implies  $0 \leq \vartheta(\theta) \leq 0.003$  (Fig. 10). Hence, the maximum absolute-relative deviation between  $w_m(\alpha_q, \beta_r, t)$  and  $w_m(t)$  amounts to only 0.03% of the  $\langle w_m(t) \rangle$  average. On addition, a range  $0 \leq \theta \leq 2.8\theta_{ste}$  leads to  $0 \leq \vartheta(\theta) \leq 0.03006$ . This leads to a maximum absolute deviation between  $w_m(\alpha_q, \beta_r, t)$  and  $w_m(t)$  of around 3% of  $\langle w_m(t) \rangle$ . Comparing with results for raw data, we can ascertain that data quality induces a relative gain in the precision of the  $w_m(\alpha, \beta, t)$  projections.

### Acknowledgements

We thank Francisco Javier Ponce Isguerra for art work.

### Funding

All funding for the present research were obtained from institutional grants

### Availability of data and materials

Data and simulation codes will be available from corresponding author upon request.

### Authors' contributions

HEH conceived, designed performed and supervised the whole research. CLR performed formal derivations simulation and numerical tasks. EVD performed the statistical analysis and interpretation of data. NCC revised the manuscript critically both at statistical and formal levels. All authors read and approved the final manuscript.

### Ethics approval and consent to participate

Not applicable.

### Consent for publication

Not applicable.

### Competing interests

The authors declare that they have no competing interest.

### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

### Author details

<sup>1</sup>Centro de Investigación Científica y de Estudios Superiores de Ensenada, Carretera Ensenada-Tijuana No. 3918, Zona Playitas, Código Postal 22860, Apdo. Postal 360 Ensenada, B.C., Mexico. <sup>2</sup>Centro de Investigación en Matemáticas, A.C. Jalisco s/n, Mineral Valenciana, Código Postal 36240 Guanajuato Gto, Mexico. <sup>3</sup>Instituto Tecnológico de Tijuana, Calzada Tecnológico S/N, Fracc. Tomas Aquino, Código Postal 22414 Tijuana, Baja California, Mexico.

Received: 11 September 2017 Accepted: 22 January 2018

Published online: 06 March 2018

## References

- Costanza R, d'Arge R, de Groot R, Farber S, Garso M, Hannon B, et al. The value of the world's ecosystem services and natural capital. *Nature*. 1997; 387:253–60.
- Short FT, Neckles HA. The effects of global climate change on seagrasses. *Aquatic Bot*. 1999;63:169–96.
- Grech A, Chartrand-Miller K, Erfemeijer P, Fonseca M, McKenzie L, Rasheed M, et al. A comparison of threats, vulnerabilities and management approaches in global seagrass bioregions. *Environ Res Lett*. 2012;7:024006.
- Unworth KFR. Seagrass meadows in a globally changing environment. *Marine Poll Bull*. 2014;83(2):383–6.
- Flindt MR, Rasmussen EK, Valdemarsen T, Erichsen A, Kaas H, Canal-Vergés P. Using a GIS-tool to evaluate potential eelgrass reestablishment in estuaries. *Ecol Model*. 2016;338:122–34.
- Hughes BB, Eby R, Dyke EV, Tinker MT, Marks CI, Johnson KS, et al. Recovery of a top predator mediates negative eutrophic effects on seagrass. *Proc Natl Acad Sci*. 2013;110:15313–8.
- van Katwijk MM, Thorhaug A, Marbà N, Orth RJ, Duarte CM, Kendrick GA, et al. Global analysis of seagrass restoration: the importance of large-scale planting. *J Appl Ecol*. 2016;53:567–78.
- Short FT, Kostenb S, Morganc PA, Malone S, Moore GE. Impact of climate change on submerged and emergent wetland plants. *Aquat Bot*. 2016;135: 3–17.
- Holmquist JG, Powell GVN, Sogard SM. Decapod and stomatopod assemblages on a system of seagrass-covered mud banks in Florida bay. *Mar Biol*. 1989;100:473–83.
- Montague CL, Ley JA. A possible effect of salinity fluctuation on abundance of benthic vegetation and associated fauna in northeastern Florida bay. *Estuar Coast Shelf Sci*. 1993;16:703–17.
- Plummer ML, Harvey CJ, Anderson LE, Guerry AD, Ruckelshaus MH. The role of eelgrass in marine community interactions and ecosystem services: results from ecosystem-scale food web models. *Ecosystems*. 2013;16(2):237–51.
- Blackburn TH, Nedwell DB, Weibe WJ. Active mineral cycling in a Jamaican seagrass sediment. *Mar Ecol Prog Ser*. 1994;110:233–9.
- Park SR, Li WT, Kim SH, Kim JW, Lee KS. A comparison of methods for estimating the productivity of *Zostera marina*. *J Ecol Field Biol*. 2010;33(1): 59–65.
- Terrados J, Borum J. Why are seagrasses important? goods and services provided by seagrass meadows. In: Borum J, Duarte CM, Krause-Jensen D, Greve TM, editors. *European seagrasses: an introduction to monitoring and management: EU project Monitoring and Managing of European Seagrasses (M&MS)*; 2004. p. 8–10. [http://www.seagrasses.org/handbook/european\\_seagrasses\\_high.pdf](http://www.seagrasses.org/handbook/european_seagrasses_high.pdf).
- Newell IER, Koch EW. Modeling seagrass density and distribution in response to changes in turbidity stemming from bivalve filtration and seagrass sediment stabilization. *Estuar Coast Shelf Sci*. 2004;27(5):793–806.
- Christianen MJA, van Belzen J, Herman PMJ, van Katwijk MM, Lamers LPM, van Leent PJM, et al. Low-canopy seagrass beds still provide important coastal protection services. *PLoS One*. 2013;8(5):e62413.
- Liu X, Zhou Y, Yang H, Ru S. Eelgrass detritus as a food source for the sea cucumber *Apostichopus Japonicus Selenka* (Echinodermata: Holothuroidea) in coastal waters of North China: an experimental study in flow-through systems. *PLoS One*. 2013;8(3):e58293.
- Kennedy H, Beggins J, Duarte CM, Fourqurean JW, Holmer M, Marbà N, et al. Seagrass sediments as a global carbon sink: isotopic constraints. *Global Biogeochem Cycles*. 2010;24(4):GB4026.
- Duarte CM, Sintes T, Marba N. Assessing the CO2 capture potential of seagrass restoration projects. *J Appl Ecol*. 2013;50:1341–9.
- Short FT, Coles RG, Pergent-Martini C. Global seagrass distribution. In: Short FT, Coles RG, editors. *Global seagrass research methods*. Amsterdam: Elsevier Science B.V; 2001. p. 6.
- Phillips RC. In: Odum HT, Copeland BJ, Mc Mahan EA, editors. *Temperate grass flats in coastal ecological Systems of the United States*. 2nd ed. Washington DC: Conservation Foundation; 1974. p. 244–99.
- Jacobs RPWM. Distribution and aspects of the production and biomass of eelgrass, *Zostera marina* L. at Roscoff, France. *Aquat Bot*. 1979;7:151–72.
- Williams RB. Nutrient level and phytoplankton productive in the estuary. In: Chabreck RA, editor. *Proceedings of the coastal marsh and estuary management symposium*. Baton Rouge: Louisiana State University; 1973. p. 59.
- Novak AB, Plaisted HK, Hays CG, Hughes RA. Limited effects of source population identity and number on seagrass transplant performance. *PeerJ*. 2017;5:e2972.
- Matheson FE, Reed J, Dos Santos VM, Mackay G, Cummings VJ. Seagrass rehabilitation: successful transplants and evaluation of methods at different spatial scales. *N Z J Mar Freshwater Res*. 2017;51(1):96–109.
- Fishman JR, Orth RJ, Marion S, Bieri J. A comparative test of mechanized and manual transplanting of eelgrass, *Zostera marina*, in Chesapeake Bay. *Restor Ecol*. 2004;12:214–9.
- Olsen JL, Coyer JA, Chesney B. Numerous mitigation transplants of the eelgrass *Zostera marina* in southern California shuffle genetic diversity and may promote hybridization with *Zostera pacifica*. *Biol Conserv*. 2014;176:133–43.
- Li WT, Kim JH, Park JI, Lee KS. Assessing establishment success of *Zostera marina* transplants through measurements of shoot morphology and growth. *Estuar Coast Shelf Sci*. 2010;88(3):377–84.
- Hauxwell J, Cebrián J, Valiela I. Eelgrass *Zostera marina* loss in temperate estuaries: relationship to land-derived nitrogen loads and effect of light limitation imposed by algae. *Mar Ecol Prog Ser*. 2003;247:59–73.
- Echavarría-Heras HA, Leal-Ramírez C, Villa-Diharce E, Cazarez-Castro NR. The effect of parameter variability in the allometric projection of leaf growth rates for eelgrass (*Zostera Marina* L.) II: the importance of data quality control procedures in bias reduction. *Theor Biol Med Model*. 2015;12:30.
- Solana-Arellano ME, Echavarría-Heras HA, Leal-Ramírez C, Kun-Seop L. The effect of parameter variability in the allometric projection of leaf growth rates for eelgrass (*Zostera marina* L.). *Lat Am J Aquat Res*. 2014; 42(5):1099–108.
- Hui D, Jackson RB. Uncertainty in allometric exponent estimation: a case study in scaling metabolic rate with body mass. *J Theor Biol*. 2007;249:168–77.
- Packard GC, Birchard GF. Traditional allometric analysis fails to provide a valid predictive model for mammalian metabolic rates. *J Exp Biol*. 2008;211:3581–7.
- Echavarría-Heras HA, Lee KS, Solana-Arellano ME, Franco-Vizcaino E. Formal analysis and evaluation of allometric methods for estimating above-ground biomass of eelgrass. *Ann Appl Biol*. 2011;159(3):503–15.
- Echavarría-Heras HA, Leal-Ramírez C, Villa-Diharce E, Castillo O. Using the value of Lin's concordance correlation coefficient as a criterion for efficient estimation of areas of leaves of eelgrass from noisy digital images. *Source code Biol Med* 2014; 9(1):29.
- Leal-Ramírez C, Echavarría-Heras HA, Castillo O. Exploring the suitability of a genetic algorithm as tool for boosting efficiency in Monte Carlo estimation of leaf area of eelgrass. In: P Melin, O Castillo, J Kacprzyk, editors. *Design of intelligent systems based on fuzzy logic, neural networks and nature-inspired optimization*. Studies in computational intelligence. England: Springer. 2015;601:291–303.
- Leys C, Ley C, Klein O, Bernard P, Licata L. Detecting outliers: do not use standard deviation around the mean, use absolute deviation around the median. *J Exp Soc Psychol*. 2013;49(4):764–6.
- Lin LI. A concordance correlation coefficient to evaluate reproducibility. *Biometrics*. 1989;45:255–68.
- McBride GB. A proposal for strength-of-agreement criteria for Lin's concordance correlation coefficient. NIWA client report: HAM2005–062. Hamilton: National Institute of Water & Atmospheric Research; 2005.
- Paiwatin Y. *Statistical Modeling and Inference Using Likelihood*. In: In all Likelihood. UK: Oxford University Press; 2013.
- Casella G, Berger RL. *Statistical inference*. 2nd ed. UK: Duxbury Press; 2001.
- Newman MC. Regression analysis of log-transformed data: statistical bias and its correction. *Environ Toxicol*. 1993;12:1129–33.
- Duan N. Smearing estimate: a nonparametric retransformation method. *J Am Stat Assoc*. 1983;78:605–10.
- Packard GC, Boardman TJ, Birchard GF. Allometric equations for predicting body mass of dinosaurs: a comment on Cawley and Janacek. *J Zool*. 2010;282:221–2.
- Lai J, Yang B, Lin D, Kerkhoff AJ, Ma K. The allometry of coarse root biomass: log-transformed linear regression or nonlinear regression? *PLoS One*. 2013; 8(10):e77007.
- Packard GC. Misconceptions about logarithmic transformation and the traditional allometric method. *Zoology*. 2017;123:115–20.
- Huber PJ. *Robust statistics*. New York: Wiley; 1981.
- Miller J. Reaction time analysis with outlier exclusion: bias varies with sample size. *Q J Exp Psychol*. 1991;43(4):907–12.