



RESEARCH

Open Access



# 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

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

\* Correspondence:

heheras@icloud.com

<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

## Abstract

**Background:** Eelgrass grants important ecological benefits including a nursery for waterfowl and fish species, shoreline stabilization, nutrient recycling and carbon sequestration. Upon the exacerbation of deleterious anthropogenic influences, re-establishment of eelgrass beds has mainly depended on transplantation. Productivity estimations provide valuable information for the appraisal of the restoration of ecological functions of natural populations. Assessments over early stages of transplants should preferably be nondestructive. Allometric scaling of eelgrass leaf biomass in terms of matching length provides a proxy that reduces leaf biomass and productivity estimations to simple measurements of leaf length and its elongation over a period. We examine how parameter variability impacts the accuracy of the considered proxy and the extent on what data quality and sample size influence the uncertainties of the involved allometric parameters.

**Methods:** We adapted a Median Absolute Deviation data quality control procedure to remove inconsistencies in the crude data. For evaluating the effect of parametric uncertainty we performed both a formal exploration and an analysis of the sensitivity of the allometric projection method to parameter changes. We used parameter estimates obtained by means of nonlinear regression from crude as well as processed data.

**Results:** We obtained reference leaf growth rates by allometric projection using parameter estimates produced by the crude data, and then considered changes in fitted parameters bounded by the modulus of the vector of the linked standard errors, we found absolute deviations up to 10 % of reference values. After data quality control, the equivalent maximum deviation was under 7 % of corresponding reference rates. Therefore, the addressed allometric method is robust. Even the smaller sized samples in the quality controlled dataset produced better accuracy levels than the whole set of crude data.

(Continued on next page)

(Continued from previous page)

**Conclusions:** We propose quality control of data as a highly recommended step in the overall procedure that leads to reliable allometric surrogates of eelgrass leaf growth rates. The proliferation of inconsistent replicates in the crude data points towards the importance of discarding incomplete leaves. We also recommend avoiding errors in estimating the biomass of small leaves for which precision of the used analytical scale might be an issue.

**Keywords:** Eelgrass leaf growth rates, Allometric estimation, Parametric uncertainty effects, Data quality control

## Background

Eelgrass is a relevant seagrass species that distributes worldwide in estuaries and near-shore environments. Eelgrass meadows provide habitat and foraging grounds for marine animals, buffers the shoreline from erosion, filter the water, and oxygenate the sediments. Recently at a global scale, deleterious effects derived from anthropogenic influences, have been exacerbated to such an extent that in the absence of remediation efforts, the important ecological services resulting from the permanence of eelgrass meadows could be irreversibly lost. Due to its ecological relevance eelgrass has been the subject of intense research and conservation efforts that are mainly carried out by means of transplanting plots. The measurement of biomass and productivity, provide key information for the evaluation of the overall status of a given eelgrass population. The noticeable growth form of this species makes the average rate of leaf growth per shoot-day measured over a growing interval of length  $\Delta t$  determinant of overall productivity. In what follows we denote the biomass of an individual eelgrass leaf at time  $t$  through the symbol  $w(t)$  and its corresponding length by means of  $l(t)$ . Similarly, we will symbolize the observed values of the biomass of leaves in shoots by means of  $w_s(t)$  and the average rate of leaf growth per shoot-day through  $L_g(t, \Delta t)$ . Conventional techniques for the assessment of  $L_g(t, \Delta t)$ , require intensive sampling that involves the removal of shoots, and then tedious, time consuming dry weight measurement procedures in the laboratory. Even though the elimination of eelgrass shoots linked to a typical evaluation activity does not infringe damage to natural populations, the effects of shoot removal could be severe for transplanted plots. Therefore, in an overall program that leads to eelgrass conservation it is important to include nondestructive assessment methods [1, 2].

Bivariate allometric scaling relationships between measured quantities  $X$  and  $Y$  expressed as a power function of the form  $Y = \beta X^\alpha$ , appear in a myriad of research problems in physics, biology, and earth and planetary sciences [3–11]. The parameter  $\alpha$  is named allometric exponent and  $\beta$  termed normalization constant. Particularly, Echavarría-Heras et al. [1] and Echavarría-Heras et al. [2], stressed the significance of the model

$$w(t) = \beta l(t)^\alpha, \quad (1)$$

in the adaptation of an allometric proxy for  $L_g(t, \Delta t)$ . This allometric surrogate will be here denoted by means of the symbol  $L_{ga}(\alpha, \beta, t, \Delta t)$ , and its explicit formulae derived in Appendix A.

The appropriateness of this device to provide truly nondestructive assessments depends on the time invariance of the parameters  $\alpha$  and  $\beta$ , because estimates previously fitted at a

site could be used to readily obtain surrogates for currently observed values of  $L_g(t, \Delta t)$ . Several studies show that within our geographical region the parameters  $\alpha$  and  $\beta$  can be considered as time invariant ([1, 11–14]). But even though  $\alpha$  and  $\beta$  are statistically invariant within a given region, environmental influences are expected to induce a relative extent of variability on local estimates of  $\alpha$  and  $\beta$  [11] and as it was stated by Echavarría-Heras et al. [2] this could propagate significant uncertainties on  $L_{ga}(\alpha, \beta, t, \Delta t)$  values. Indeed the results of Appendix B corroborate that these proxies could produce significantly accurate assessments for  $L_g(t, \Delta t)$  only in case the fitting of equation (1) yields highly precise estimates of the parameters  $\alpha$  and  $\beta$ .

One important factor contributing to the uncertainty of the estimates of the parameters  $\alpha$  and  $\beta$ , in equation (1) associates to the set of biological influences that could have been disregarded when assuming that eelgrass leaf biomass depends solely on linked length. Another important source of uncertainty relates to the failure of the model of equation (1) to handle environmental effects on a proper way. But the high values of determination coefficients as well as results of residual analysis corresponding to fittings of the model of equation (1) to independent data sets collected in our geographical region ([1, 2, 11, 13, 15, 16]) reveal that on spite of its simplicity, equation (1) is a paradigm that provides highly consistent estimates of the involved parameters.

But beyond uncertainties that biological or environmental influences could convey on estimates of the parameters  $\alpha$  and  $\beta$  in equation (1), there could be other factors affecting their precision. Moreover, the results of Savage et al. [17], Hui and Jackson [18], Packard and Birchard [19], illustrate the relevance of factors like measurement error, data quality, sample size, and regression method on the precision of estimates of the parameters in a bivariate allometric scaling relationship. Moreover, Hui and Jackson [18], Packard and Birchard [19] and Packard and Boardman [20] concluded that in a bivariate allometric scaling nonlinear regression provides better estimates of the allometric exponent than these commonly estimated by linear regressions. Moreover, nonlinear regression was the analysis method used to obtain the significant estimates of the parameters  $\alpha$  and  $\beta$  reported in all the eelgrass studies performed at our geographical region. Therefore, in the present settings we also considered nonlinear regression as an analysis method. However, in comparison with previously reported results ([2, 11, 15]) applying this analysis method to the present data set, resulted in an smaller value for the determination coefficient. Besides, in the current crude data, we detected a proliferation of inconsistent replicates, in the whole showing a noticeable discrepancy to the dispersion pattern that is normally expected for a scaling relationship of the form given by equation (1). Then, we assumed that on the light of the proven consistency of this model, the unusual spreading observed in the present data could on its own; provide an explanation for the reduction in determination coefficient obtained here. This view was strengthened by the fact that the discussed anomalous spreading was not observed for the data sets used in previous analysis of the fitting of a scaling relationship like equation (1) at our study site [11–16]. And, it was perhaps a lack of standardization in processing tasks that led to a relatively larger incidence of measurement error in the present data set. Therefore, we considered that the exploration of data quality effects on the precision of estimates of the parameters  $\alpha$  and  $\beta$  in equation (1) is a fundamental step in the overall procedure that leads to reliable allometric surrogates for the assessment of eelgrass leaf growth rates. Since, the aforesaid study had not been yet done; we made here an attempt to fill this gap.

Eelgrass leaf biomass is lognormally distributed, and the sizes of the different groups of replicates that make up our data are reduced. Therefore, we removed inconsistent replicates from the set of crude data by using a robust median absolute deviations cleaning procedure. We also relied on a sensitivity analysis study in order to assess the improvement in the accuracy of the  $L_{ga}(\alpha, \beta, t, \Delta t)$  proxy that may possibly be associated to whatever reduction on the uncertainties of  $\alpha$  and  $\beta$  could be derived from an improvement in data quality. Besides, we explored the corresponding effects of sample size, by randomly drawing differently sized samples out of both the crude and the processed data sets and then comparing the values of the determination coefficients for the fittings of equation (1) to the selected samples. In the Methods section we further substantiate the data cleaning approach, the procedure for estimating the effects of sample size as well as the steps of the performed sensitivity analysis. In the Results and Discussion section we elaborate on the relevance and limitations of this study. In the Conclusions section we presented the summary and potential implications of our findings.

## Methods

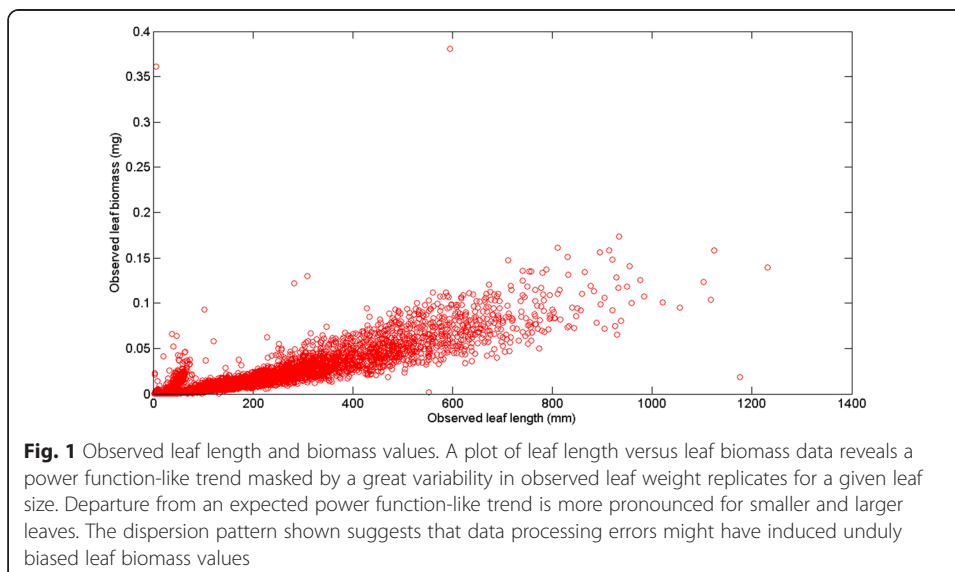
### Raw data

For the aims of the present research we assembled an extensive data set containing measurements taken on a total of 10412 individual eelgrass leaves collected in San Quintin Bay Baja California as previously reported in ([1, 2, 11–14]). Crude data includes measurements of length (mm), width (mm) and dry weight (gr).

### Data quality control procedures

Short [21] and Gaeckle and Short [22] estimated leaf biomass in eelgrass by using an isometric weight to length ratio, and the appropriateness of allometric methods in eelgrass research has been validated for a number of independent data sets (e.g. [1, 2, 11–16] and references therein). Particularly, Echavarría-Heras et al. [2] formally demonstrated that the conspicuous leaf architecture and growth form of eelgrass makes leaf length a reliable allometric descriptor of the associated biomass. Indeed, results show that for independent data sets collected in our geographical region the model of equation (1) always produced reliable fits ([1, 2, 11–13, 15, 16]). Moreover, the results of Solana-Arellano et al. [11] stating that the parameters associated to the scaling relationship of equation (1) can be considered invariant within a given geographical region, endorse that this model is highly consistent. This makes it reasonable to assume that the true relationship linking  $w(t)$  and  $l(t)$  should conform to a power function like dependency and consequently we could expect a typical dispersion for the linked data. In other words, the inherent spreading pattern should fit in a dominant power function trend masked by expected stochastic variability induced by intrinsic biological factors and environmental forcing. Particularly, at our study site the fitting of equation (1) produced a determination coefficient of  $R^2 = 0.90$  for the data set addressed in [2]. Meanwhile, for data set in [11] as well as for that reported in [15] the corresponding value was  $R^2 = 0.92$ , with consistency of residuals verified for all these fits. Since in comparison with these results, for the present set of crude data the fitting of the model of equation (1) produced a reduced determination coefficient ( $R^2 = 0.81$ ) we concluded that differences in determination coefficients could be explained by inconsistencies in the present data that could be linked to factors beyond

biological influences or local environmental forcing. Moreover, by using a direct survey of the dispersion pattern shown by the present crude data (Fig. 1) we observed that departure of a given replicate from an expected power function-like trend was more pronounced for leaves with smaller sizes. Since, this effect was not observed in related data sets collected in our region, we judged that these inconsistencies could be in a fundamental way tied to the considerably larger extension of the present data set, which required processing performed by several technicians and it was perhaps a lack of standardization in these tasks that explains the present proliferation of inconsistencies that should be removed in order to prevent their permanence from exerting a reduction in the precision of estimates for the parameters in the scaling relationship of equation (1). Notwithstanding, the judgment to remove an outlier or inconsistent measurement in a data set, it is necessary to be able to identify its occurrence. Data cleaning procedures tasks are commonly performed by using a mean plus or minus three standard deviations method. This technique is based on the property of a normal distribution for which 99.87 % of the data appear within this range [23], but its application in the present settings presents difficulties. It firstly assumes that the distribution of data is normal while eelgrass leaf biomass is lognormally distributed [2]. Secondly, both the mean and standard deviation are strongly impacted by outliers [24]. Thirdly, our data is composed by groups of measurements that include a reduced number of leaf biomass replicates, and the mean plus or minus three standard deviations method is very unlikely to detect outliers in small samples [25]. This fact made it also inconvenient to use alternative data cleaning techniques like studentized residuals, the hat elements, Cook's distance, or the Mahalanobis distance procedure [26]. Alternatively, the median of a group of data is totally immune to the sample size and a robust estimator of scale, reason why we adapted a Median Absolute Deviation (MAD) method [27] which provided the present criteria for the removal of inconsistencies in groups of more than 10 replicates of biomass values associated to a given leaf length. By eliciting the consistency of the allometric scaling of equation (1), we assumed that the addressed length-to-weight variation pattern should conform to a power function-like trend. Therefore, individual leaf biomass measurements



that in an initial data analysis were found to severely deviate from this pattern in sets under 10 replicates were removed from the set of unprocessed data. In summary, by analyzing the spreading of leaf biomass values in a leaf length-to-weight plot, we detected replicates that we supposed represented unduly deviations from an expected power function-like variation pattern. By applying the explained data cleaning procedures we removed these from the set of crude data, because in light of the proven consistency of the scaling relationship of equation (1) these inconsistent replicates could have been caused by errors due to a lack of standardization in leaf length or biomass measurements, or perhaps, explained by errors owed to faulty equipment for dry weight assessment or even explained by incorrect recordings.

In order to formalize the present MAD data cleaning procedure, we arranged the crude data in different groups  $G(l) = \{w_{G1}(l), \dots, w_{Gn}(l)\}$  formed by an observed leaf length  $l$  and associated leaf biomass replicates  $w_{G1}(l), \dots, w_{Gn}(l)$ . For almost all groups of replicates  $G(l)$  we observed several leaf biomass values that parted from the expected variability pattern and were considered as inconsistencies. For each group  $G(l)$  we firstly obtained its median denoted by means of the symbol  $MED\{w_{G1}(l), \dots, w_{Gn}(l)\}$  or simply by means of  $M(G(l))$  for short. Then, for each replicate  $w_{Gj}(l)$  in  $G(l)$  we calculated its absolute deviation from the group median  $\delta_{Gj}(l)$ , that is,  $\delta_{Gj}(l) = |w_{Gj}(l) - M(G(l))|$ . Similarly, we obtained the median of the set of absolute deviations denoted by the symbol  $MED\{\delta_{G1}(l), \dots, \delta_{Gn}(l)\}$ . Following, Huber [28] and also recalling that eelgrass leaf biomass values are log-normally distributed [2], we obtained the Median Absolute Deviation of a group  $G(l)$  denoted here through  $MAD(G(l))$  and given by

$$MAD(G(l)) = bMED\{\delta_{G1}(l), \dots, \delta_{Gn}(l)\}, \quad (2)$$

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

For the removal of inconsistent replicates in a group  $G(l)$  we used the decision criterion

$$M(G(l)) - T \cdot MAD(G(l)) < w_j(l) < M(G(l)) + T \cdot MAD(G(l)), \quad (3)$$

where  $T$  is the rejection threshold that following Miller [24], we set at a value of  $T = 3$ . For groups  $G(l)$  under ten replicates we applied a direct data cleaning procedure by removing replicates that we considered were severely deviated from the central power function-like trend.

### Sensitivity analysis

In order to evaluate the influences that the uncertainties of the parameters  $\alpha$  and  $\beta$  convey in the performance of  $L_{ga}(\alpha, \beta, t, \Delta t)$ , we used both a formal study and also simulation runs. The analytical exploration is presented in Appendix B. For the simulation study, initially, at equally spaced iteration (i.e., sampling) times  $t$ , we drawn uniformly distributed random numbers representing a number of shoots retrieved  $NS(t, \Delta t)$  and the number  $nl(s)$  of leaves that each one of these shoots holds. Afterwards, the lengths  $l(t + \Delta t)$  of each one the leaves in a shoot were simulated according to a fitted lognormal distribution, ( $p > 0.05$ ), with a mean of 207.24 mm and a standard error of 3.8 [2]. Similarly, the corresponding leaf length increments,  $\Delta l$ , were also extracted from a fitted lognormal distribution ( $p > 0.05$ ) with a mean of 140.04 mm and a standard error of 7.3 [2]. The values for these increments coincided with the length of the whole leaf, as it occurs in

newly produced leaves. Subsequently, we utilized the simulated leaf lengths and estimates  $\hat{\alpha}$  and  $\hat{\beta}$  of the parameters  $\alpha$  and  $\beta$  in equation (1), in order to produce the proxy growth rates  $L_{ga}(\alpha, \beta, t, \Delta t)$ . In order to produce estimates  $\hat{\alpha}$  and  $\hat{\beta}$  for the parameters  $\alpha$  and  $\beta$  we followed Hui and Jackson [18], Packard and Birchard [19] and Packard and Boardman [20] and fitted equation (1) to the present data sets using an iterative nonlinear least-squares method rather than the traditional approach of linearizing the equation through a logarithmic transformation of data. All fittings were performed using the Matlab Statistics Toolbox.

To study the sensitivity of  $L_{ga}(\alpha, \beta, t, \Delta t)$  to changes in the parameters  $\alpha$  and  $\beta$  we selected as reference values the estimates  $\hat{\alpha}$  and  $\hat{\beta}$  and produced changing values  $\alpha_{ep}$  for  $\hat{\alpha}$  and  $\beta_{eq}$  for  $\hat{\beta}$  namely,

$$\alpha_{ep} = \hat{\alpha} + \Delta\alpha_p, \tag{4}$$

$$\beta_{eq} = \hat{\beta} + \Delta\beta_q, \tag{5}$$

with the values  $\Delta\alpha_p$  and  $\Delta\alpha_q$  satisfying

$$|\Delta\alpha_p| = p \cdot stde(\hat{\alpha}) \tag{6}$$

and

$$|\Delta\beta_q| = q \cdot stde(\hat{\beta}), \tag{7}$$

where  $stde(\hat{\alpha})$  and  $stde(\hat{\beta})$  stand for the standard errors of  $\hat{\alpha}$  and  $\hat{\beta}$  respectively and  $p$  and  $q$  are numbers satisfying,  $0 < p \leq 1$  and  $0 < q \leq 1$ .

Therefore, the fluctuating values  $\alpha_{ep}$  and  $\beta_{eq}$  for  $\hat{\alpha}$  and  $\hat{\beta}$  are scaled through proportions  $p$  and  $q$  of their standard errors. In our analysis we also relied on a parameter change index denoted by means of  $\rho(p, q)$ , and defined through

$$\rho(p, q) = \sqrt{(\hat{\beta} - \beta_{eq})^2 + (\hat{\alpha} - \alpha_{ep})^2}. \tag{8}$$

Particularly, the maximum value that  $\rho(p, q)$  can attain as  $p$  and  $q$  vary is denoted through the symbol  $\rho_{max}(p, q)$  and is given by,

$$\rho_{max}(p, q) = \sqrt{stde(\hat{\alpha})^2 + stde(\hat{\beta})^2} \tag{9}$$

In what follows we will also use the mean value of  $\rho(p, q)$  denoted here by  $\rho_{av}(p, q)$  and given by

$$\rho_{av}(p, q) = \sqrt{stde(\alpha)^2 + stde(\beta)^2} / 2. \tag{10}$$

The simulated leaf lengths  $l(t + \Delta t)$ , corresponding increments  $\Delta l$  the  $NS(t, \Delta t)$  and  $nl(s)$  values, and the parameter estimates  $\hat{\alpha}$  and  $\hat{\beta}$  produced by means of equation (17) a reference trajectory  $L_{ga}(\hat{\alpha}, \hat{\beta}, t, \Delta, t)$ . It turns out that every picked value of  $\rho(p, q)$  yields different pairs  $(\alpha_{ep}, \beta_{eq})$ , each one associated to a couple  $(\Delta\alpha_p, \Delta\beta_q)$  that comply with the condition of equation (8). For each pair  $(\alpha_{ep}, \beta_{eq})$ , the simulated leaf and shoot data and equation (17) shape a trajectory  $L_{ga}(\alpha_{ep}, \beta_{eq}, t, \Delta t)$ . The average of the values

that these trajectories produce at each sampling time  $t$  is then calculated and its value denoted through  $\langle L_{ga}(\alpha_{ep}, \beta_{eq}, t, \Delta t) \rangle_\rho$ . Following the procedure, requires to calculate the deviations between  $\langle L_{ga}(\alpha_{ep}, \beta_{eq}, t, \Delta t) \rangle_\rho$  and  $L_{ga}(\hat{\alpha}, \hat{\beta}, t, \Delta t)$ , that are denoted by means of the symbol  $\delta L_{gap}(\alpha_{ep}, \beta_{eq}, t, \Delta t)$  and obtained through,

$$\delta L_{gap}(\alpha_{ep}, \beta_{eq}, t, \Delta t) = \left\langle L_{ga}(\alpha_{ep}, \beta_{eq}, t, \Delta t) \right\rangle_\rho - L_{ga}(\hat{\alpha}, \hat{\beta}, t, \Delta t). \tag{11}$$

We also need to compute the average through time of the  $\delta L_{gap}(\alpha_{ep}, \beta_{eq}, t, \Delta t)$  deviations whose output is represented here through the symbol  $\langle \delta L_{gap}(\alpha_{ep}, \beta_{eq}, t, \Delta t) \rangle_t$ . Finally, it is necessary to obtain the time average of the values of the reference trajectory  $L_{ga}(\hat{\alpha}, \hat{\beta}, t, \Delta t)$ . The resulting record is denoted using the symbol  $\left\langle L_{ga}(\hat{\alpha}, \hat{\beta}, t, \Delta t) \right\rangle_t$ . These statistics are then used to produce the relative deviation index  $\vartheta(\Delta\alpha_p, \Delta\beta_q)_\rho$  given by

$$\vartheta(\Delta\alpha_p, \Delta\beta_q)_\rho = \frac{\left| \left\langle \delta L_{gap}(\alpha_{ep}, \beta_{eq}, t, \Delta t) \right\rangle_t \right|}{\left\langle L_{ga}(\hat{\alpha}, \hat{\beta}, t, \Delta t) \right\rangle_t}. \tag{12}$$

The value of  $\vartheta(\Delta\alpha_p, \Delta\beta_q)_\rho$  provides a measure of the sensitivity of the reference trajectory  $L_{ga}(\hat{\alpha}, \hat{\beta}, t, \Delta t)$  to a change of tolerance  $\rho(p, q)$  on the pair  $(\hat{\alpha}, \hat{\beta})$ . Moreover, by calculating  $\vartheta(\Delta\alpha_p, \Delta\beta_q)_{max}$  defined through,

$$\vartheta(\Delta\alpha_p, \Delta\beta_q)_{max} = \max_\rho \left\{ \vartheta(\Delta\alpha_p, \Delta\beta_q)_\rho \right\}, \tag{13}$$

we can get in what percentage of  $\left\langle L_{ga}(\hat{\alpha}, \hat{\beta}, t, \Delta t) \right\rangle_t$  the maximum absolute deviation between  $L_{ga}(\alpha_{ep}, \beta_{eq}, t, \Delta t)$  and  $L_{ga}(\hat{\alpha}, \hat{\beta}, t, \Delta t)$  amounts.

**Sample size effects**

In order to explore the extent of sample size influences on  $L_{ga}(\hat{\alpha}, \hat{\beta}, t, \Delta t)$  we performed the fitting of equation (1), using differently sized samples of leaf length and biomass  $(l(t), w(t))$  data which were randomly drawn from the available datasets. For each sample of size  $n$ , we obtained the determination coefficient ( $r_n^2$ ) as well as the values of the estimators of the parameters  $\hat{\alpha}_n$  and  $\hat{\beta}_n$ , their respective standard errors  $stde_n(\hat{\alpha})$  and  $stde_n(\hat{\beta})$  and the values of the upper bound  $\rho_{max}(p, q)_n$  and  $\vartheta(\Delta\alpha_p, \Delta\beta_q)_{\rho max}$ .

**Results and discussion**

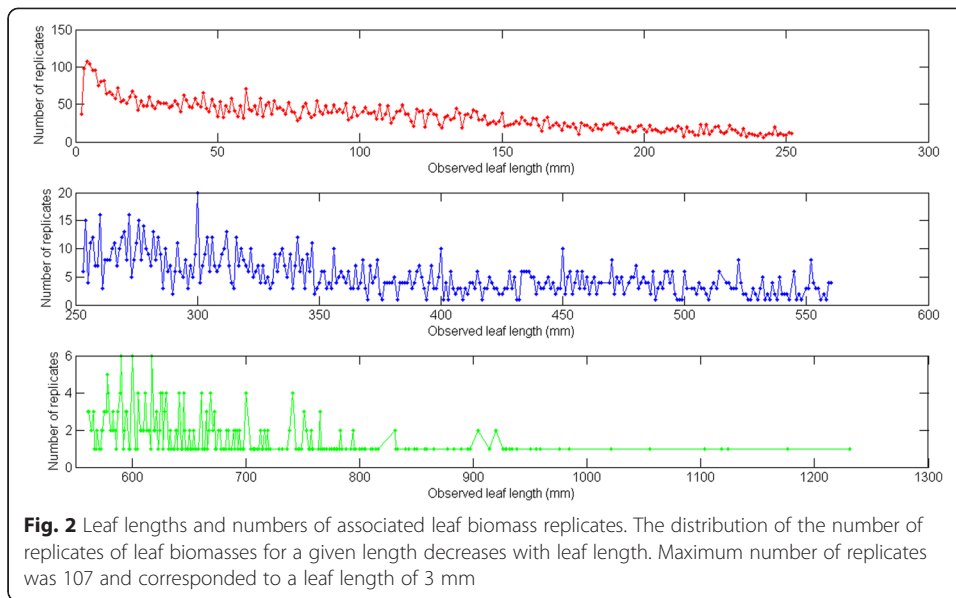
In seagrass research, allometric equations have provided useful empirical models aimed to the representation of biologically relevant traits  $Y$  in terms of an easily measured variable  $X$  (e.g. [29–33]). This is readily exemplified by equation (1) that provides a convenient representation of eelgrass leaf biomass  $w(t)$  in terms of associated length  $l(t)$ . But the realm of allometric methods in eelgrass research is not only limited to empirical formulations that provide convenient proxies for leaf biomass estimations. Indeed from a theoretical standpoint, allometric approaches have provided for instance, a framework for the evaluation of the suitability of the leaf-biomass-to length ratio as a tool for nondestructive



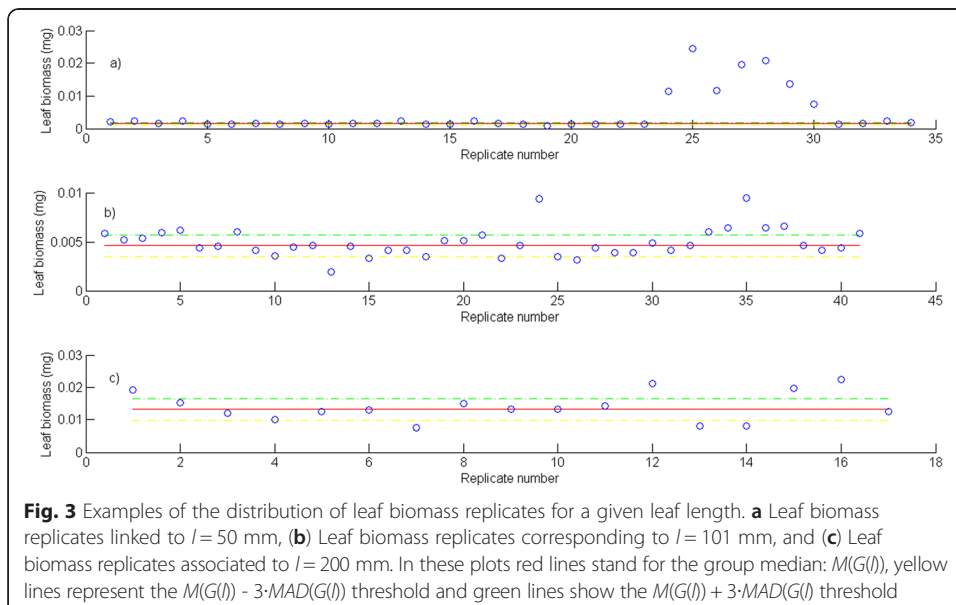
leaf biomass evaluations [15] and have also endorsed the substantiation of the plasto-chrone method of leaf growth assessments [16]. Moreover, equation (1) is the foundation of the derivation of the  $L_{ga}(\alpha, \beta, t, \Delta t)$  proxy of equation (17) a paradigm aimed to the indirect assessment of eelgrass leaf growth rates ([1, 2, 12]).

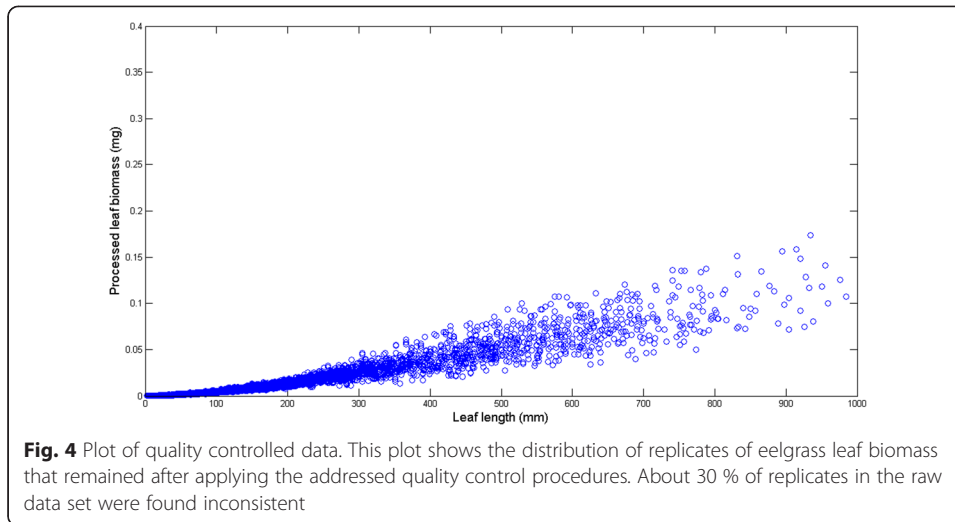
The effectiveness of the  $L_{ga}(\alpha, \beta, t, \Delta t)$  construct for providing truly accurate and non-destructive assessments for the observed average rate of leaf growth per shoot-day  $L_g(t, \Delta t)$  depends on both, the time invariability of the parameters  $\alpha$  and  $\beta$ , and on the accuracy of their estimates. From a general standpoint, the variability of the allometric exponent  $\alpha$  has been the focus of theoretical and empirical studies because it often seems to have a constant value specific to a particular biological relationship (e.g. [34–39]). On the other hand, several studies have provided evidence that support certain variability in the exponent of allometric scaling laws (e.g. [40–45]). Accordingly, the value of the normalization constant  $\beta$  is thought to be characteristic of species or populations [46]. And the variability observed in the normalization constant, explained as a differential response to environmental conditions [34, 35, 47–49]. Particularly, for eelgrass, Solana-Arellano et al. [11] analyzed independent data set collected in different geographical regions to conclude that no universal values can be found for the allometric parameter  $\alpha$  in equation (1), suggesting as well that this scaling relationship might be considered static, thus implying that local factors determine the extent of the variability of the actual values of the parameters  $\alpha$  and  $\beta$ . On the presence of this variability associated uncertainties are expected to spread inaccuracies in eelgrass leaf growth rates produced by means of equation (17).

Figure 1 displays the variability of observed leaf lengths and linked biomasses. The survey of the crude data reveals a notorious proliferation of replicates associated to an observed leaf length. Data was arranged into 755 different groups  $G(l) = \{w_{G1}(l), \dots, w_{Gn}(l)\}$ . Figure 2 displays the variation of the number of replicates for a given leaf size. We can observe that the number of replicates in groups  $G(l)$  decreases with leaf size. This is probably due to the fact that separation from shoots by drag forces is more pronounced for longer and older leaves than it is for small and younger leaves. Figure 3 shows examples of the dispersion within  $G(l)$  groups. Moderated differences in replicated leaf biomass values associating to a given length can be explained by the inherent stochastic variability along with normal systematic errors introduced by data processing. Nevertheless, an exploration of the dispersion pattern of leaf biomass values in the present crude data reveals replicates representing unduly deviations from the expected power function-like trend. These inconsistent measurements are more visible for leaves under 100 mm long and also for those longer than 800 mm. For groups with 10 or more replicates and according to the present MAD data quality control criterion, whenever an individual leaf biomass replicate  $w_{Gj}(l)$  differed in magnitude from the group median by more than 3 times the MAD it was considered as an inconsistency and it was rejected [50]. Similarly, the power function-like trend data cleaning criteria described in the methods section permitted the removal of inconsistent leaf biomass values in groups with less than 10 replicates. Figure 4 presents a plot of the leaf biomass values that remained after applying data quality control. This procedure retained a representative number of replicates showing the normally expected variability of eelgrass leaf biomass values ([1, 2, 11–16]). Moreover, in order to exemplify differences in quality among equivalent data sets we applied the present data cleaning procedure to equivalent data reported in Echavarría-Heras et al. [12]. For easy of presentation we chose leaf length class intervals covering the extent of observed leaf lengths in



both data sets. For each one of the lengths of in these class intervals we applied the present MAD procedure and calculated the percentage of discarded replicates. Then we obtained the average percentage of discarded leaves that associate to each length class interval. Figure 5 compares the average percentages of discarded replicates per leaf length class for the present data and for the Echavarría-Heras et al. [12] data set. For leaf lengths lying in the first class interval ( $0 \leq l(t) < 50$ ) the average percentage of rejected replicates in the present data set was 30 % while the Echavarría-Heras et al. [12] data reported no inconsistent replicates for the same interval. And generally, the percentages of discarded replicates in the present data set were markedly bigger than those associated to the Echavarría-Heras et al. [12] data. This comparison can explain why the fitting the model

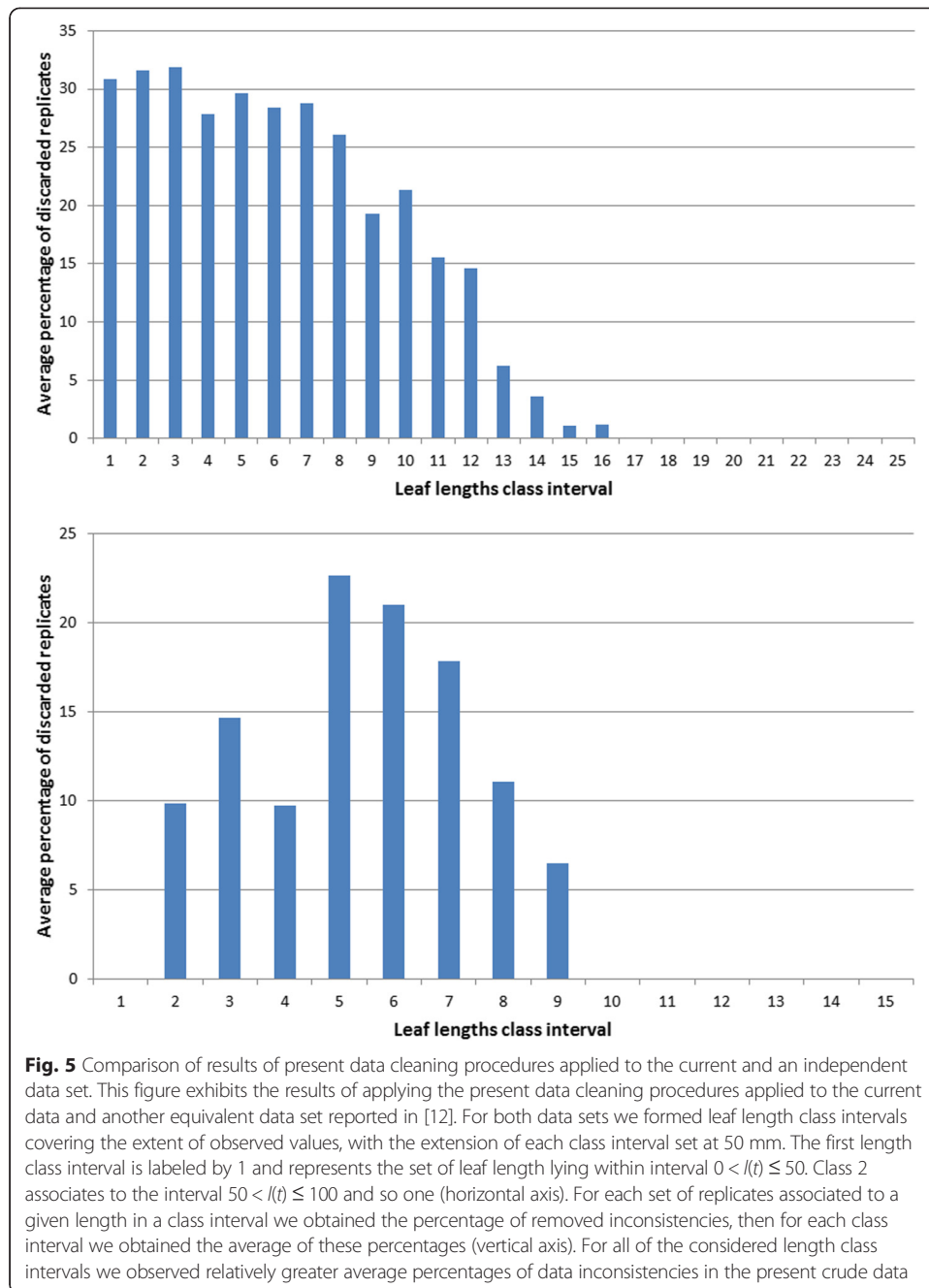




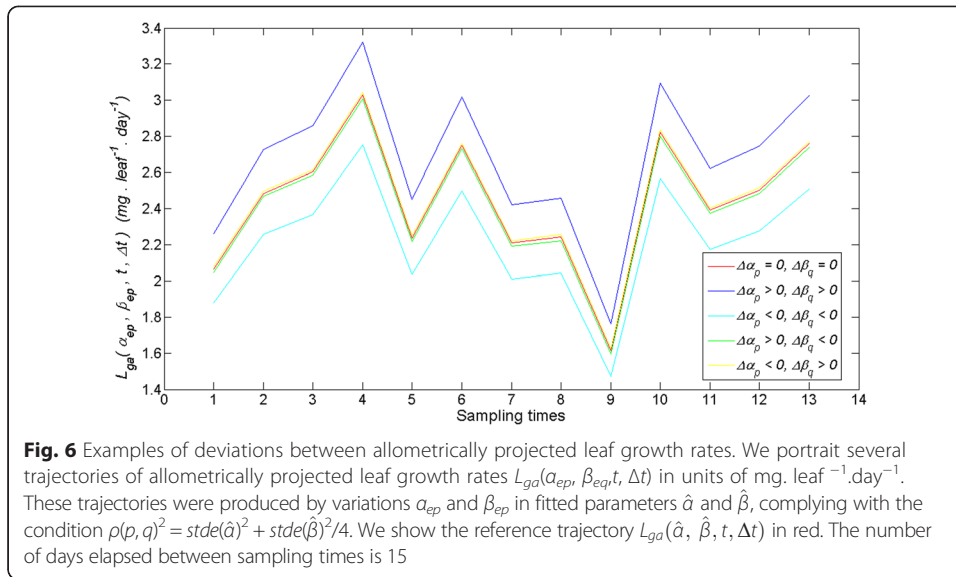
of equation (1) to the present crude data set produced an smaller determination coefficient than that reported in Echavarría-Heras et al. [12].

Observations show that the order relationship  $0 < \beta < \alpha$  holds [1, 2, 11, 12] and that it is also reasonable to consider the variation ranges of  $\Delta\alpha$  and  $\Delta\beta$  satisfying:  $|\Delta\alpha| < \alpha$  and  $|\Delta\beta| < \beta$ . By embracing these assumptions in Appendix B we derived explicit formulae for the deviations  $\delta L_{ga}(\Delta\alpha, \Delta\beta, t, \Delta t)$  that changed values of the parameters  $\alpha$  and  $\beta$  having the form  $\alpha_e = \alpha + \Delta\alpha$  and  $\beta_e = \beta + \Delta\beta$  induce in the reference values  $L_{ga}(\alpha, \beta, t, \Delta t)$ . These are given by equations (21). By assuming that the observed order relationship for the parameters  $\alpha$  and  $\beta$  holds and also by considering the aforesaid expected variation ranges for  $\Delta\alpha$  and  $\Delta\beta$ , we found that the proxies  $L_{ga}(\alpha, \beta, t, \Delta t)$  for the average rate of leaf growth per shoot-day, will be overestimated by  $L_{ga}(\alpha_e, \beta_e, t, \Delta t)$  values, primarily when the ordered pair  $(\Delta\alpha, \Delta\beta)$  lies on the domain  $\Delta\alpha > 0$  and  $\Delta\beta > 0$ . Nevertheless, as it is explained in Appendix B, each positive value of  $\Delta\beta$ , can be associated to a set of negative values of  $\Delta\alpha$  for which  $L_{ga}(\alpha, \beta, t, \Delta t)$  will be also overestimated by  $L_{ga}(\alpha_e, \beta_e, t, \Delta t)$ . Correspondingly,  $L_{ga}(\alpha_e, \beta_e, t, \Delta t)$  underestimates  $L_{ga}(\alpha, \beta, t, \Delta t)$  whenever the pair  $(\Delta\alpha, \Delta\beta)$  is placed inside the region  $-\alpha < \Delta\alpha < 0$  and  $-\beta < \Delta\beta < 0$ , but this time, each value of  $\Delta\beta$  satisfying  $-\beta < \Delta\beta < 0$  can be associated to a set of positive values of  $\Delta\alpha$  for which  $L_{ga}(\alpha, \beta, t, \Delta t)$  will be also underestimated by  $L_{ga}(\alpha_e, \beta_e, t, \Delta t)$ .

By fitting of equation (1) to the crude leaf length and weight data we obtained  $r^2 = 0.81$  and estimates,  $\hat{\alpha} = 1.32367$  with  $stde(\hat{\alpha}) = 0.0143$  for  $\alpha$  and of  $\hat{\beta} = 0.000014$  with  $stde(\hat{\beta}) = 1.4e-6$  for  $\beta$ . Therefore, according to equation (8) we have  $0 \leq \rho(p, q) \leq 0.0143$ . Using the  $\hat{\alpha}$  and  $\hat{\beta}$  estimates we obtained  $\rho_{max}(p, q) = 0.0143$  (cf. Eq. (9)) and  $\rho_{av}(p, q) = 0.0072$  (cf. Eq. (10)). By means of the simulated leaf and shoot data and equation (17) we estimated the  $L_{ga}(\hat{\alpha}, \hat{\beta}, t, \Delta t)$  reference trajectory. Afterwards for fixed values of the parameter change index  $\rho(p, q)$  we used equations (6) and (7) to produce values for the  $\Delta\alpha_p$  and  $\Delta\beta_p$  increments complying with the condition of equation (8). This allowed the characterization of the associated  $L_{ga}(\alpha_{ep}, \beta_{ep}, t, \Delta t)$  trajectories. Figure 6 displays trajectories obtained for the case in which  $\rho(p, q)$  takes its average value  $\rho_{av}(p, q)$ . These simulation outputs are consistent with the results of the appendix setting domains of

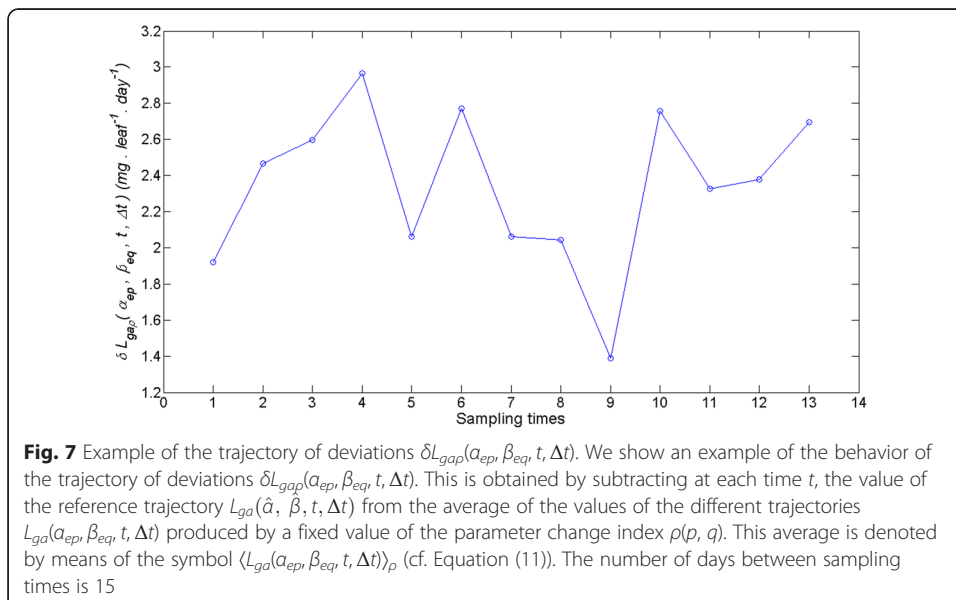


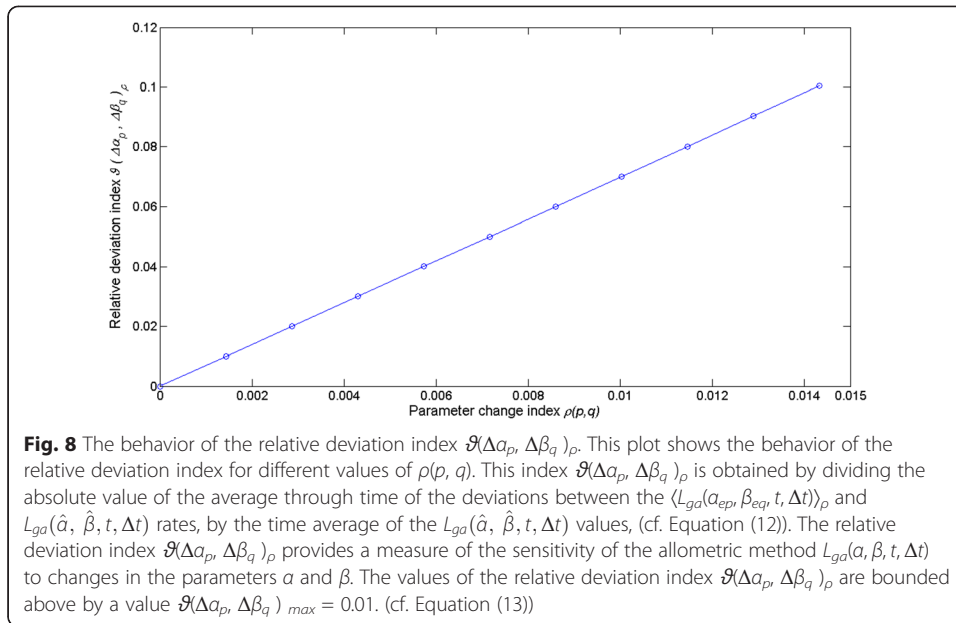
variation of the pair  $(\Delta\alpha, \Delta\beta)$  where  $L_{ga}(\alpha, \beta, t, \Delta t)$  values are overestimated or underestimated. Similarly, Fig. 7 shows the time variation of the average deviations  $\langle \delta L_{ga}(\Delta\alpha_p, \Delta\beta_q, t, \Delta t) \rangle_p$  obtained for  $\rho(p, q) = \rho_{av}(p, q)$ . Moreover, Fig. 8 shows that whenever  $0 \leq \rho(p, q) \leq \rho_{max}(p, q)$  the values of the relative deviation index  $\vartheta(\Delta\alpha_p, \Delta\beta_q)_\rho$  are bounded above by a value  $\vartheta(\Delta\alpha_p, \Delta\beta_q)_{max} = 0.01$ , that is, the maximum absolute deviation between  $L_{ga}(\alpha_{ep}, \beta_{eq}, t, \Delta t)$  and  $L_{ga}(\hat{\alpha}, \hat{\beta}, t, \Delta t)$  amounts to 10 % of  $\langle L_{ga}(\hat{\alpha}, \hat{\beta}, t, \Delta t) \rangle_t$ . In turn, when fitting of equation (1) to the quality controlled data we obtained  $r^2 = 0.91$  and estimates  $\hat{\alpha} = 1.403$  with  $stde(\hat{\alpha}) = 0.0120$  for  $\alpha$  and of  $\hat{\beta} = 8.462e-006$



with  $\text{stde}(\hat{\beta}) = 6.4300e-007$ , for  $\beta$ . Therefore, according to equation (9) this time we have  $\rho_{max}(p, q) = 0.01$ . In turn, performing the sensitivity study of equations (4) through (13) we obtained  $\vartheta(\Delta\alpha_p, \Delta\beta_q)_{max} = 0.07$ , which amounts to 7 % of  $\langle L_{ga}(\hat{a}, \hat{\beta}, t, \Delta t) \rangle_t$  (see Fig. 9).

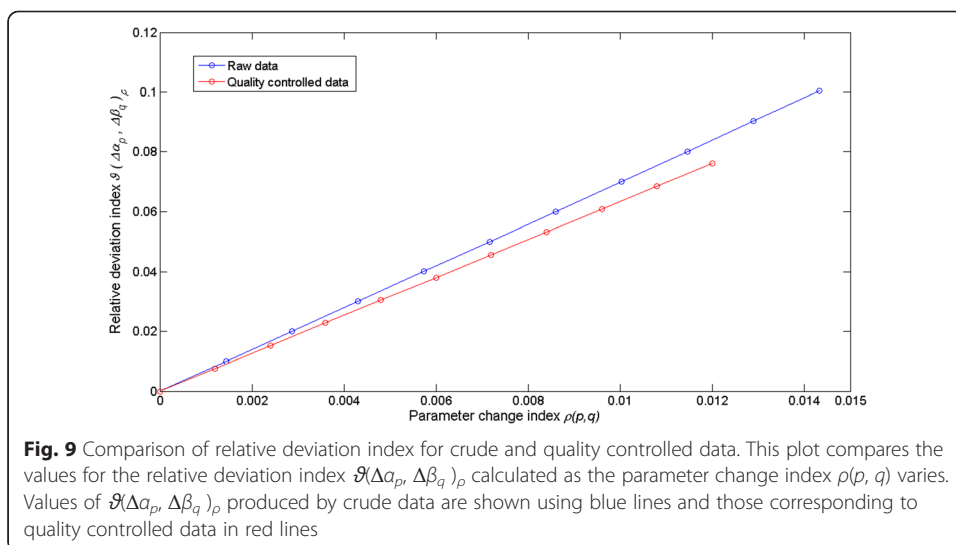
This sensitivity exploration, sets the accuracy of the allometric method of equation (17) to be mainly dependent on the extent of the upper bound  $\rho_{max}(p, q)$  of the parameter change index  $\rho(p, q)$ . This means that an improvement in the quality of the fit of equation (1) reducing the magnitude of the upper bound for  $\rho(p, q)$  will certainly increase the accuracy of the proxies  $L_{ga}(\alpha, \beta, t, \Delta t)$ . This study also demonstrates the

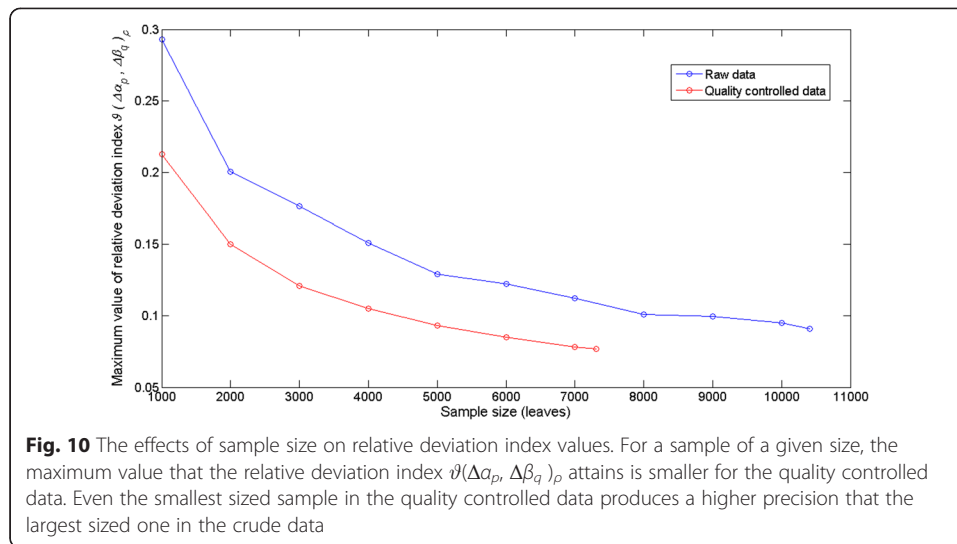




importance of data quality control in the reduction of  $\rho_{max}(p, q)$ . Therefore, within the realm of parametric uncertainties as determined by of the values of  $stde(\hat{\alpha})$  and  $stde(\hat{\beta})$  obtained for the quality controlled data set, the addressed projection method can be considered robust relative to numerical differences in the estimators  $\hat{\alpha}$  and  $\hat{\beta}$ .

Figure 10 presents the variation of  $\vartheta(\Delta\alpha_p, \Delta\beta_q)_{max}$  depending on sample size for both the set of raw data and that resulting after quality control procedures. As expected Fig. 10 shows that for both data sets  $\vartheta(\Delta\alpha_p, \Delta\beta_q)_{max}$  decreases as sample size increases, but these plots also reveal that the effect of sample size on the accuracy of the  $L_{ga}(\alpha, \beta, t, \Delta t)$  proxy is more pronounced for the raw data set. Moreover, even the





smallest sized sample in the quality controlled data set ( $n = 1000$ ) produces a smaller value for  $\vartheta(\Delta\alpha_p, \Delta\beta_q)_{max}$  than that induced by the largest sized sample in the raw data set ( $n = 10412$ ). This adds on to the aforesaid on the prominent role of data quality control in improving the accuracy of eelgrass leaf growth rate assessments obtained by means of  $L_{ga}(\alpha, \beta, t, \Delta t)$ .

## Conclusions

The results of the present study highlight the important role that the precision of parameter estimates linked to the allometric model of equation (1) plays in the overall suitability of the allometric proxy for eelgrass leaf growth rates  $L_{ga}(\alpha, \beta, t, \Delta t)$  given by equation (17). The basic allometric model of equation (1) has been consistently identified using different data sets, and a property of invariance for the involved parameters statistically verified. But on spite of model consistency, the evaluation of the extent on what factors such as data quality sampling size, and analysis method can influence parameter estimates is an important entry in the comprehensive analysis of an allometric scaling relationship like equation (1). Moreover, the recommendation by Hui and Jackson [12] stating that for variables that are measured with errors, it is important to obtain accurate estimates with repeated measurements on similar individuals grown under similar conditions highlights the relevance of controlling factors that could affect data quality in our settings. Since the present results were obtained following the recommendation of Hui and Jackson [18], Packard and Birchard [19] and Packard and Boardman [20] concerning analysis method, and since sample size in our study was optimal, it is reasonable to assume that errors in data processing could explain a drop in the determination coefficient for the fitting of equation (1) to crude data collected at our study site from  $R^2 = 92$  obtained by Solana-Arellano et al. [11] and Echavarría-Heras et al. [15] to a value of  $R^2 = 81$  produced by the present data set. And the fact that after the data cleaning procedures performed on the present data we obtained for the resulting determination coefficient a value of  $R^2 = 91$  seems

to endorse our judgment. Therefore, our results seem to confirm that data quality can be raised as a crucial factor explaining precision of estimates for the parameters  $\alpha$  and  $\beta$  in equation (1). Moreover, the present data quality control procedure reveals the remarkable influence of inconsistencies in reducing precision of allometric projections. Indeed, as it is shown in Fig. 9 the accuracy of the  $L_{ga}(\alpha, \beta, t, \Delta t)$  proxy improved after processing data irregularities. Our contribution provides guidelines that could prevent the proliferation of undesirable data quality related effects on the accuracy of allometric projections of eelgrass leaf growth rates. It was perhaps a lack of standardization that could explain a relatively reduced quality in the present data set. In other words the unduly deviations from an expected power function-like spreading could be probably tied to errors in leaf length estimation, faulty equipment for dry weight assessment, or even due to the incorrect recording of measurements. Nevertheless, the spread displayed in Fig. 4, shows notorious remnant variability for leaves longer than 500 mm, even after the application of data quality control procedures. This suggests that a data quality approach must be mandatory not only during the processing stage but even since the gathering step of forming a data set. Indeed for leaves with the smaller sizes, that normally yield reduced biomasses we might expect estimation errors that are imputable to the precision of the analytical scale used to obtain leaf dry weight estimations, along with the reading and/or recording of the output. But longest leaves remain exposed longer to environmental effects such as drag forces or herbivore. These influences could remove sections of the leaves rendering underestimated biomasses for a given leaf length. Therefore, special care must be also taken when processing longer an older leaves which are more often damaged or trimmed to such an extent that the biomass they report do not necessarily associate to the true length to weight relationship. We anticipate that the overlooking of these important issues as well as other data processing mistakes like rounding off errors or even incorrect data recording could explain the anomalous proliferation of inconsistencies (about 30 %) found in the present crude data set. Our findings also reveal that a data quality approach could optimize sampling endeavors, this because as shown in Fig. 10, even the smallest sized sample of the quality controlled data could entail better model identification results than the whole unprocessed data set. In summary, the present findings show how data quality control can be conceived as a fundamental step in the overall procedure that leads to reliable nondestructive allometric proxies of eelgrass leaf growth rates. And, it is also worth to point out, that it can be misleading to implicitly assume that a given set of data should conform to an idealized allometric relationship and then use a data cleaning procedure to ensure the fit. It can be misleading because the dispersion pattern itself could act as an indicator of eelgrass population status. But, we must emphasize that in accomplishing the present approach we are not neglecting the fact that there could be other unknown influences explaining the dispersion observed in the present crude data, but we are in fact acknowledging that the proven consistency and invariance of the addressed allometric relationship for eelgrass leaf biomass, is expected to provide a dependable account of pertinent causal influences, being these linked to the biology of the plant or associated to environmental forcing, thus highlighting the relevance of data quality influences in determining an unexpected variation pattern in the data as the one detected here.



**Appendix A**

In what follows a subscript  $s$  will be used to label a generic eelgrass shoot, holding a number  $nl(s)$  of leaves with combined biomass  $w_s(t)$ . In time, if  $\Delta w_l(t, \Delta t)$  stand for the increment in biomass that is gained by an individual leaf over a growing interval  $[t, t + \Delta t]$ , then denoting by means of  $L_{sg}(t, \Delta t)$  the resulting average growth rate of the leaves on the shoot  $s$  we will have,

$$L_{sg}(t, \Delta t) = \frac{\sum_{nl(s)} \Delta w_l(t, \Delta t)}{\Delta t},$$

and consequently if  $L_g(t, \Delta t)$  denotes the linked average rate of leaf growth per shoot-day, we then have,

$$L_g(t, \Delta t) = \frac{\sum_{NS(t, \Delta t)} L_{sg}(t, \Delta t)}{NS(t, \Delta t)}, \tag{14}$$

where  $\sum_{NS(t, \Delta t)}$  indicates summation of the shoots collected over the marking interval  $[t, t + \Delta t]$  and being  $NS(t, \Delta t)$  their number. Moreover, as it is thoroughly explained by Echavarría-Heras et al. [2], we can use equation (1) in order to derive an allometric approximation for  $L_{sg}(t, \Delta t)$ , which we here denote through the symbol  $L_{sga}(\alpha, \beta, t, \Delta t)$  and formally express by

$$L_{sga}(\alpha, \beta, t, \Delta t) = \frac{\sum_{nl(s)} \Delta w_{la}(t, \Delta t)}{\Delta t}, \tag{15}$$

where  $\Delta w_{la}(t, \Delta t)$  is an allometric proxy for  $\Delta w_l(t, \Delta t)$ .

It turns out that equation (1) yields,

$$\Delta w_{la}(t, \Delta t) = \beta l(t + \Delta t)^\alpha - \beta l(t)^\alpha$$

then, factoring  $l(t + \Delta t)^\alpha$  and taking into account that  $l(t + \Delta t) = l(t) + \Delta l$  where  $\Delta l$  stands for the increment in leaf length gained over the interval  $[t, t + \Delta t]$  we get,

$$\Delta w_{la}(t, \Delta t) = \beta l(t + \Delta t)^\alpha (1 - (1 - \rho_l(t, \Delta t))^\alpha)$$

where

$$\rho_l(t, \Delta t) = \frac{\Delta l}{l(t + \Delta t)}.$$

And by letting

$$\delta(t, \Delta t) = (1 - (1 - \rho_l(t, \Delta t))^\alpha)$$

we equivalently have

$$\Delta w_{la}(t, \Delta t) = \beta l(t + \Delta t)^\alpha \delta(t, \Delta t).$$

Therefore, from (15) we obtain,

$$L_{sga}(\alpha, \beta, t, \Delta t) = \frac{\sum_{nl(s)} \beta l(t + \Delta t)^\alpha \delta(t, \Delta t)}{\Delta t} \tag{16}$$

Similarly, if  $L_{ga}(\alpha, \beta, t, \Delta t)$  denotes the allometric proxy for  $L_{sga}(t, \Delta t)$ , this is given by

$$L_{ga}(\alpha, \beta, t, \Delta t) = \frac{\sum_{NS(t, \Delta t)} L_{sga}(\alpha, \beta, t, \Delta t)}{NS(t, \Delta t)}, \tag{17}$$

and in turn we have

$$L_g(t, \Delta t) = L_{ga}(\alpha, \beta, t, \Delta t) + \epsilon_{ga}, \tag{18}$$

with the term  $\epsilon_{ga}$  standing for the involved approximation error. It is worth to point out that the derivation of the results of equation (18) could have been done using leaf area  $a(t)$  in place of  $l(t)$ .

**Appendix B**

In this appendix we firstly derive explicit formulae for the deviations  $\delta L_{ga}(\Delta\alpha, \Delta\beta, t, \Delta t)$  that a change  $\Delta\alpha$  and  $\Delta\beta$  in  $\alpha$  and  $\beta$  will induce in  $L_{ga}(\alpha, \beta, t, \Delta t)$ . Using the derived forms we obtain domains of variation of  $\Delta\alpha$  and  $\Delta\beta$  where  $L_{ga}(\alpha, \beta, t, \Delta t)$  are underestimated or overestimated. Without loss of generality, we will assume through that the inequality  $0 < \beta < \alpha$  holds, as it regularly occurs in fittings of equation (1) to eelgrass leaf biomass and length data. Also, for the sake of mathematical tractability, without sacrificing pertinence of the analysis, regarding the increments  $\Delta\alpha$  and  $\Delta\beta$  we will assume that  $|\Delta\alpha| \leq \alpha$  and  $|\Delta\beta| < \beta$ .

By definition above we have,

$$\delta L_{sga}(\Delta\alpha, \Delta\beta, t, \Delta t) = L_{sga}(\alpha + \Delta\alpha, \beta + \Delta\beta, t, \Delta t) - L_{sga}(\alpha, \beta, t, \Delta t) \tag{19}$$

And from equation (17) one gets

$$\delta L_{ga}(\Delta\alpha, \Delta\beta, t, \Delta t) = \frac{\sum_{NS(t, \Delta t)} \delta L_{sga}(\Delta\alpha, \Delta\beta, t, \Delta t)}{NS(t, \Delta t)} \tag{20}$$

Now equation (16) yields,

$$\delta L_{sga}(\Delta\alpha, \Delta\beta, t, \Delta t) = \frac{\sum_{nl(s)} \beta l(t + \Delta t)^\alpha \delta(t, \Delta t) \mu_{Lg}(\Delta\alpha, \Delta\beta, l(t + \Delta t))}{\Delta t}, \tag{21}$$

With

$$\mu_{Lg}(\Delta\alpha, \Delta\beta, l(t)) = \left( \left( \frac{\beta + \Delta\beta}{\beta} \right) \left( 1 - \left( 1 - \frac{\Delta l}{l(t + \Delta t)} \right)^{\alpha + \Delta\alpha} \right) \right) \delta(t, \Delta t)^{-1} l(t + \Delta t)^{\Delta\alpha} - 1, \tag{22}$$

which can be equivalently written as

$$\mu_{Lg}(\Delta\alpha, \Delta\beta, l(t)) = \frac{\varphi(\Delta\alpha)}{\theta(\Delta\beta)} - 1, \tag{23}$$

where

$$\theta(\Delta\beta) = \frac{\beta}{\beta + \Delta\beta}, \tag{24}$$

and

$$\varphi(\Delta\alpha) = \frac{\left(1 - \left(\frac{l(t)}{l(t+\Delta t)}\right)^{\alpha+\Delta\alpha}\right) l(t+\Delta t)^{\Delta\alpha}}{\left(1 - \left(\frac{l(t)}{l(t+\Delta t)}\right)^\alpha\right)} \tag{25}$$

Now by assumption  $0 < \beta < \alpha$  and  $|\Delta\alpha| \leq \alpha$ , and since observations show that the statement that  $l(t)$  being time increasing is consistent, for  $\Delta t > 0$  we necessarily have  $l(t + \Delta t) > l(t)$  and therefore, the factor  $\varphi(\Delta\alpha)$  as given above increases in its entire domain and also remains positive. Again the case  $\Delta\alpha = \Delta\beta = 0$  will set  $\theta(\Delta\beta) = \varphi(\Delta\alpha) = 1$  (Fig. 11). Moreover,  $L_{ga}(\alpha, \beta, t, \Delta t)$  will be overestimated by  $L_{ga}(\alpha + \Delta\alpha, \beta + \Delta\beta, t)$  whenever the inequality

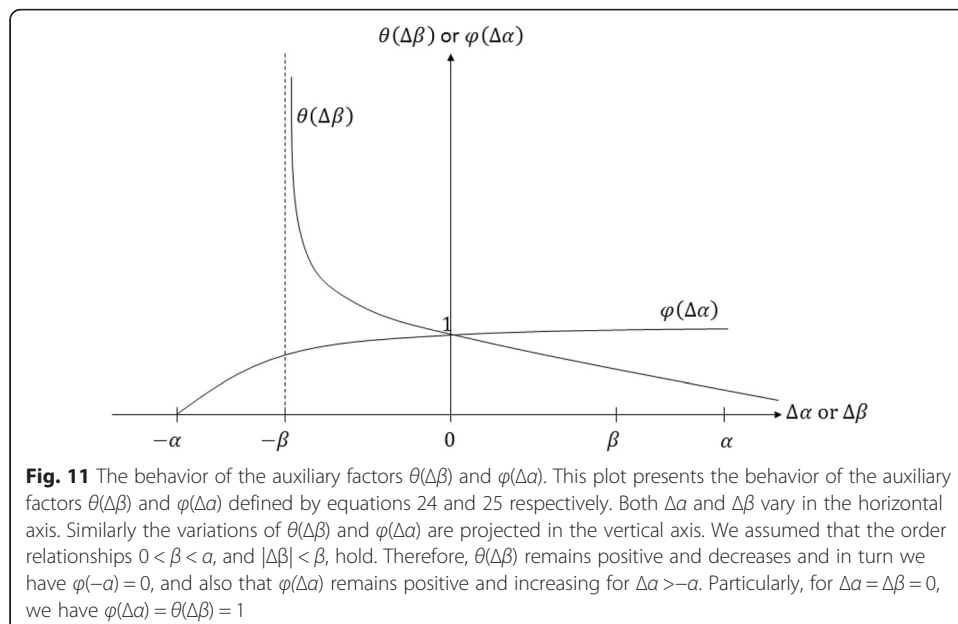
$$\mu_{Lg}(\Delta\alpha, \Delta\beta, l(t)) > 0 \tag{26}$$

holds, but since we also assumed that  $|\Delta\beta| < \beta$ , as it has been already pointed out, the factor  $\theta(\Delta\beta)$  remains positive, by virtue of (23) we have that the statement in (26) implies

$$\varphi(\Delta\alpha) > \theta(\Delta\beta). \tag{27}$$

Now for  $\Delta\alpha \geq 0$ ,  $\varphi(\Delta\alpha)$  satisfies  $\varphi(\Delta\alpha) \geq 1$  and increases, while for  $\Delta\beta \geq 0$  the factor  $\theta(\Delta\beta)$  satisfies  $\theta(\Delta\beta) \leq 1$  and decreases, therefore the statement (27) will clearly hold in the domain  $\Delta\alpha > 0$  and  $\Delta\beta > 0$ . Notwithstanding, for  $\Delta\beta = \Delta\beta^*$  fixed, and such that  $0 < \Delta\beta^* < \beta$ , we have  $\theta(\Delta\beta^*) \geq 1/2$ , then since for  $-\alpha \leq \Delta\alpha < 0$ , we have  $0 \leq \varphi(\Delta\alpha) \leq 1$ , and  $\varphi(\Delta\alpha)$  is continuous there, we can exhibit a value  $\Delta\alpha^*$ , that satisfies,  $-\alpha \leq \Delta\alpha^* < 0$ , for which we have  $\varphi(\Delta\alpha^*) = \theta(\Delta\beta^*)$ , and since  $\varphi(\Delta\alpha)$  increases whenever  $\Delta\alpha^* \leq \Delta\alpha \leq 0$ , we will have  $\varphi(\Delta\alpha) \geq \theta(\Delta\beta^*)$ . Therefore, to each positive value of  $\Delta\beta$ , we can associate a set of negative values of  $\Delta\alpha$  for which inequality (27) also holds.

Now,  $L_{ga}(\alpha + \Delta\alpha, \beta + \Delta\beta, t)$  will underestimate  $L_{ga}(\alpha, \beta, t, \Delta t)$ , whenever



$$\mu_{Lg}(\Delta\alpha, \Delta\beta, l(t)) < 0 \quad (28)$$

is satisfied, but again since  $|\Delta\beta| < \beta$ , from (23) the statement above implies

$$\varphi(\Delta\alpha) < \theta(\Delta\beta) \quad (29)$$

Now for  $-\alpha \leq \Delta\alpha \leq 0$ , the factor  $\varphi(\Delta\alpha)$  vanishes at  $\Delta\alpha = -\alpha$  and then increases monotonically towards a value  $\varphi(0) = 1$ . Meanwhile, as it has been discussed before for  $-\beta < \Delta\beta \leq 0$  the term  $\theta(\Delta\beta)$  decreases monotonically towards a value  $\theta(0) = 1$ . Therefore, inequality (29) will particularly hold in the domain  $\Delta\alpha < 0$  and  $\Delta\beta < 0$ . This time, for  $\Delta\beta = \Delta\beta_\tau$  fixed and satisfying  $-\beta < \Delta\beta_\tau < 0$ , we have  $\theta(\Delta\beta_\tau) > 1$  and since  $\varphi(\Delta\alpha)$  is continuous, there exists  $\Delta\alpha_\tau$  positive, for which we have  $\varphi(\Delta\alpha_\tau) = \theta(\Delta\beta_\tau)$  and since  $\varphi(\Delta\alpha)$  increases and changes continuously over the domain  $|\Delta\alpha| < \alpha$ , particularly for  $0 \leq \Delta\alpha < \Delta\alpha_\tau$  we will have,  $1 \leq \varphi(\Delta\alpha) < \theta(\Delta\beta_\tau)$ . Therefore, each negative value of  $\Delta\beta$  can be associated to a set of positive values of  $\Delta\alpha$  for which inequality (29) also holds.

#### Competing interests

The authors declare that they have no competing interests.

#### Authors' contributions

H conceived, designed and performed the research. C performed formal derivations and simulation and numerical tasks. E performed the statistical analysis and interpretation of data. N supervised the whole research and revised the manuscript critically both at statistical and formal levels. All authors read and approved the final manuscript.

#### 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, Guanajuato Gto. Código Postal 36240, Mexico. <sup>3</sup>Instituto Tecnológico de Tijuana, Calzada Tecnológico S/N, Fracc. Tomas Aquino, Tijuana, Baja California Código Postal 22414, Mexico.

Received: 9 June 2015 Accepted: 12 November 2015

Published online: 01 December 2015

#### References

- 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, Solana-Arellano ME, Leal-Ramírez C, Franco-Vizcaino E. An allometric method for measuring leaf growth in eelgrass, *Zostera marina*, using leaf length data. *Bot Mar*. 2013;56(3):275–86.
- Newman MEJ. Power laws, Pareto distributions and Zipf's law. *Contemporary Phys*. 2005;46:323–51.
- Marquet PA, Quiñones RA, Abades S, Labra F, Tognelli M. Scaling and power-laws in ecological systems. *J Exp Biol*. 2005;208:1749–69.
- West GB, Brown JH. The origin of allometric scaling laws in biology from genomes to ecosystems: towards a quantitative unifying theory of biological structure and organization. *J Exp Biol*. 2005;208:1575–92.
- Harris LA, Duarte CM, Nixon SW. Allometric laws and prediction in estuarine and coastal ecology. *Estuar Coast*. 2006;29:343–7.
- Filgueira R, Labarta U, Fernández-Reiriz MJ. Effect of condition index on allometric relationships of clearance rate in *Mytilus galloprovincialis* Lamarck. 1819. *Rev Biol Mar Oceanogr*. 2008;43(2):391–8.
- Kaitaniemi P. How to derive biological information from the value of the normalization constant in allometric equations. *PLoS One*. 2008;3(4):e1932.
- Martin RD, Genoud M, Hemelrijk CK. Problems of allometric scaling analysis: examples from mammalian reproductive Biology. *J Exp Biol*. 2005;208:1731–47.
- González-Trilla G, Borro MM, Morandeira NS, Schivo F, Kandus P, Marcovecchio J. Allometric scaling of dry weight and leaf area for spartina densiflora and spartina alterniflora in two southwest Atlantic Saltmarshes. *J Coast Res*. 2013;29(6):1373–81.
- Solana-Arellano ME, Echavarría-Heras HA, Leal-Ramírez C, Lee KS. 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.
- Echavarría-Heras HA, Solana-Arellano ME, Franco-Vizcaino E. An allometric method for the projection of eelgrass leaf biomass production rates. *Math Biosci*. 2010;223(1):58–65.
- Solana-Arellano E, Echavarría-Heras H, Díaz-Castañeda V, Flores-Uzeta O. Shoot biomass assessments of the marine phanerogam *Zostera marina* for two methods of data gathering. *Am J Plant Sci*. 2012;3:1541–5.
- Solana-Arellano E, Borbón-González DJ, Echavarría-Heras H. A general allometric model for blade production in *Zostera marina* L. *Bull South Cal Acad Sci*. 1998;97:39–49.
- Echavarría-Heras H, Solana-Arellano E, Lee KS, Hosokawa S, Franco-Vizcaino E. An evaluation of leaf biomass: length ratio as a tool for nondestructive assessment in eelgrass (*Zostera marina* L.). *Sci World J*. 2012;2012:543730.

16. Echavarría-Heras H, Solana-Arellano E, Leal-Ramírez C, Castillo O. Using allometric procedures to substantiate the plastochrone method for eelgrass leaf growth assessments *Theoretical Biology and Medical Modelling*. 2013; 10:34. <http://www.tbiomed.com/content/10/1/34>
17. Savage VM, Gillooly JF, Woodruff WH, West GB, Allen AP. The predominance of quarter-power scaling in biology. *Funct Ecol*. 2004;18:257–82.
18. 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.
19. 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.
20. Packard GC, Boardman TJ. Model selection and logarithmic transformation in allometric analysis. *Physiol Biochem Zool*. 2008;81:496–507.
21. Short FT. Effects of sediment nutrients on seagrasses: literature review and mesocosm experiment. *Aquat Bot*. 1987;27(1):41–57.
22. Gaeckle JL, Short FT. Plastochrone method for measuring leaf growth in eelgrass, *Zostera marina* L. *Bull Mar Sci*. 2002;71(3):1237–46.
23. Howell DC. *Statistical methods in human sciences*. New York: Wadsworth; 1998.
24. Miller J. Reaction time analysis with outlier exclusion: bias varies with sample. *Q J Exp Psychol*. 1991;43(4):907–12.
25. Cousineau D, Chartier S. Outliers detection and treatment: a review. *Int J Psychol Res*. 2010;3(1):58–67.
26. Stevens JP. Outliers and Influential data points in regression analysis. *Psychol Bull*. 1984;95(2):334–44.
27. 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.
28. Huber PJ. *Robust statistics*. New York: John Wiley; 1981.
29. McRoy CP. Standing stock and other features of eelgrass (*Zostera marina*) populations on the coast of Alaska. *J. Fish. Res. Bd. Can*. 1970; 27:1811–1821.
30. Patriquin DG. Estimation of growth rate, production and age of the marine angiosperm, *Thalassia testudinum*. *Mar Biol*. 1973;15(1):35–46.
31. Jacobs RPWM. Distribution and aspects of the production and biomass of eelgrass, *Zostera marina* L. at Roscoff France. *Aquat Bot*. 1979;7:151–72.
32. Hamburg SP, Homann PS. Utilization of growth parameters of eelgrass, *Zostera marina*, for productivity estimates under laboratory and in situ conditions. *Mar Biol*. 1986;93:299–303.
33. Duarte CM. Allometric scaling of seagrass form and productivity. *Mar Ecol Prog Ser*. 1991;77:289–300.
34. Winter JE. A critical review on some aspects of filter feeding in lamellibranchiate bivalves. *Haliotis*. 1976;7:71–87.
35. Bernard FR. Physiology and mariculture of some northeastern Pacific bivalve molluscs. *Can Spec Pub Fish Aquat Sci*. 1983;63:1–24.
36. Jones HD, Richards OG, Southern TA. Gill dimensions, water pumping rate and body size in the mussel *Mytilus edulis* L. *J Exp Mar Biol Ecol*. 1992;155:213–37.
37. Enquist BJ, Brown JH, West GB. Allometric scaling of plant energetic and population density. *Nature*. 1998;395:163–5.
38. Niklas KJ, Enquist BL. Invariant scaling relationships for interspecific plant biomass production rates and body size. *Proc Natl Acad Sci*. 2001;98:2922–7.
39. West GB, Woodruff VH, Brown HJ. Allometric scaling of metabolic rate from molecules and mitochondria to cells and mammals. *Proc Natl Acad Sci*. 2002;99:2473–8.
40. Riisgård HU. No foundation of a '3/4 power scaling law' for respiration in biology. *Ecol Lett*. 1998;1:71–3.
41. Atanasov AT, Dimitrov DB. Changes of the power coefficient in the 'metabolism-mass' relationship in the evolutionary process of animals. *Biosystems*. 2002;66:65–71.
42. Bokma F. Evidence against universal metabolic allometry. *Funct Ecol*. 2004;18:184–7.
43. Muller-Landau HC, Condit RS, Chave J, Thomas SC, Bohlman SA, Bunyavejchewin S, et al. Testing metabolic ecology theory for allometric scaling of tree size, growth and mortality in tropical forests. *Ecol Lett*. 2006;9:575–88.
44. Reich PB, Tjoelker MG, Machado JL, Oleksyn J. Universal scaling of respiratory metabolism, size and nitrogen in plants. *Nature*. 2006;439:457–61.
45. White CR, Phillips NF, Seymour RS. The scaling and temperature of vertebrate metabolism. *Biol Lett*. 2006;2:125–7.
46. Niklas KL. *Plant allometry: the scaling of form and process*. Chicago: University of Chicago; 1994.
47. West GB, Brown JH, Enquist BJ. A general model for the origin of allometric scaling laws in biology. *Science*. 1997; 276:122–6.
48. Gillooly JF, Brown JH, West GB, Savage VM, Charnov EL. Effects of size and temperature on metabolic rate. *Science*. 2001;293:2248–51.
49. Brown JH, Gupta L, Li BL, Milne BT, Restrepo C, West GB. The fractal nature of nature: power laws, ecological complexity and biodiversity. *Phil Trans R Soc Lond B*. 2002;357:619–26.
50. Miller JN. Rouges and suspects: how to tackle outliers. *Amc Technical Briefs No. 39*, 2009. [http://www.rsc.org/images/tackle-outliers-technical-brief-39\\_tcm18-214866.pdf](http://www.rsc.org/images/tackle-outliers-technical-brief-39_tcm18-214866.pdf).