Article

Significances of differences between slopes: An upgrade for replicated time series

Vasco M. N. C. S. Vieira[1], **Joel Creed**[2]

[1]MARETEC, Instituto Superior Técnico, Universidade Técnica de Lisboa, Av. Rovisco Pais, 1049-001, Lisboa, Portugal

[2]Departamento de Ecologia, Instituto de Biologia Roberto Alcântara Gomes, Universidade do Estado do Rio de Janeiro - UERJ, PHLC Sala 220, Rua São Francisco Xavier 524, 20.559-900, Rio de Janeiro, RJ, Brazil

E-mail: vasco.vieira@ist.utl.pt

Abstract

In some ecology subjects the slope of the line fit between *x* and *y* variables is the focus of concern. Such is the case of self-thinning theory, developed for plant demography and later verified also occurring in algae and animals. Different slopes identify statistical populations subject to different conditions. Therefore, it is fundamental that a test is able to identify honestly significant differences between slopes. The most used tested for the overlap of the 95% confidence intervals of the bootstrapped slopes. However, Vieira and Creed (2013) demonstrated it to possess weak theoretical grounds having proposed a permutation methods alternative. Unfortunately, both were fallible upon small sample sizes and/or large data scatter. Data about self-thinning, as well as other subjects, often comes in replicated time series enabling upgrading the test algorithm to randomize sampling units only within the respective time frame. This was added to the previous software, increasing outstandingly the capacity of the permutation test in identifying both true and false differences between slopes.

Keywords slope difference; permutation; bootstrap; overlap; randomization tests; time series.

1 Introduction

Determining the significance of slope differences is a common requirement in ecological studies, namely about self-thinning, ontogeny and sexual dimorphism, among others. When analyzing data, slopes are usually estimated using model 1 linear regression where *y* is dependent on *x* and residuals are minimized vertically, i.e. along the *y* axis. However, when there is no dependency relation between variables or *x* is measured with error model 1 regression is inadequate being proposed model 2 regression alternatives. The residuals are minimized obliquely. Principal Components Analysis (PCA) was the immediate choice (Pearson, 1901; Jackson, 1991). Later, Reduced Major Axis (RMA) was found to be more accurate and reliable under broader conditions (Sokal and Rolf, 1981). The significance of the differences between slopes has long been estimated testing for

the overlap of their bootstrapped 95% confidence intervals (Creed, 1995; Scrosati and DeWreede, 1997; Creed et al., 1998; Steen and Scrosati, 2004) following the methodology proposed by Sokal and Rolf (1981). However, Vieira and Creed (2013) found the numerical random re-sampling with repetition favours the occurrence of re-combinations yielding largely diverging slopes, widening their confidence intervals and thus increasing the chances of overlooking significant differences. To overcome this problem a permutation test was proposed simulating the null hypothesis of no differences between slopes against which the original difference was tested. It is a recurring matter using bootstrap and Jackknife methods to estimate significances of metrics (Stauffer et al., 1985; Jackson, 1993; Peres-Neto et al., 2003; Peres-Neto et al., 2005; Zhang, 2011a, b). However, these are a class of randomization methods originally developed to estimate confidence intervals around metrics (Efron and Tibshirani, 1986; Manly, 1991; Lebart, 2006; Zhang, 2010) and not significance testing. Its algorithm does not even consider a null hypothesis. Therefore, better suited alternatives of permutation methods have been proposed (Manly, 1991; Dijksterhuis and Heiser, 1995; Anderson and Legendre, 1999; Peres-Neto et al., 2005; Ferrarini, 2011; Vieira, 2012; Legendre, 2013; Vieira and Creed, 2013).

Self-thinning was first studied in plants (Yoda et al., 1963; White and Harper, 1970; Lonsdale and Watkinson, 1983; Weller, 1987; Weiner and Wigham, 1988; Ellison, 1989; Morris, 1996; Morris, 2003) but found also occurring among algae, fish and gastropods (Hughes and Griffiths, 1988; Creed, 1995; Arenas and Fernández, 2000; Ríncon and Cerviá, 2002; Keeley, 2003; Steen and Scrosati, 2004; Rivera and Scrosati, 2008). When individuals in mono-specific even-aged stands undergo active growth while enduring strong competition, the stronger grow and the weaker are eliminated. Thus the stand gains biomass although loosing individuals. Plotting the $\log_{10}$ stand biomass against the $\log_{10}$ stand density fits well a straight line. Its slope indicates the strength of the competition, i.e. the flatter the slope the stronger the competition (Weiner and Wigham, 1988; Ellison, 1989; Morris, 1996, 2003; Keeley, 2003; Steen and Scrosati, 2004). Testing it has often been made recurring to manipulative experiments, sometimes in controlled environments (Weiner and Wigham, 1988; Ang and DeWreede, 1992; Morris, 1996; Scrosati and DeWreede, 1997; Creed et al., 1998; Arenas et al., 2002; Keeley, 2003; Steen and Scrosati, 2004; Scrosati, 2006). The methods and software by Vieira and Creed (2013) neglected the former self-thinning data when collected in time-series often replicated, thus overlooking their potential to considerably improve comparisons among different line fits. Fig.1 shows examples of replicated time series. The previous methods randomize among the entire set of replicates with sampling units loosing track of its time coordinate. Odd combinations of sampling units may concentrate randomized sets in a particular sub-area of the plot (i.e. in a narrower time frame) leading to wide scatter among the randomized slopes and thus decreasing the tests ability to detect significant differences. The objective of this work was to test the gain from randomizing sampling units only within their respective time coordinate. The use of permutation tests to estimate the significance of differences between slopes requires data to be standardized to their groups' means (Vieira and Creed, 2013). With replicated time series the data may also be standardized to the replicates' means with further gains in line fitting; although its adequacy is a matter of ecological theory.

2 Methods

2.1 The data analysis

The data had m observations divided among g groups, each with r time series replicates composed by t sequenced sampling units (su). Groups did not need to have equal numbers of replicates. Their slopes were estimated using RMA regression. The significances of the differences between slopes were estimated following two methods: (i) permutation tests estimated the original slope difference, randomized the full data

set with su loosing track of their replicate origin and estimated the new slope difference. This was repeated *n* times. Presently, *n* = 9999 randomization iterations was used. The original slope difference was compared to the distribution of the 10000 slope differences (the original plus the randomized). (ii) The groups were bootstrapped *n* times (presently, *n* = 10000) yielding the error distribution and confidence intervals for their slopes. These were found significantly different at the 5% level when their 95% confidence intervals did not overlap. A more detailed explanation about the permutation method can be found in Vieira and Creed (2013) whereas about RMA and the bootstrap/overlap method in Sokal and Rolf (1981) and Vieira and Creed (2013).

The replicated time series was the first upgrade. When these run in parallel the sampling units could (and should) be randomized within their respective time coordinate. The *Fucus serratus* data had sequences of 7su. Thus, there were 7 bins within which the respective su were randomized. For the permutation test each bin had all the su from the correspondent time coordinate in all replicates from the groups being compared. For the bootstrap/overlap method each bin had all the su from the correspondent time coordinate from all replicates in one group. This replicated time series method cannot be used when the time series does not run in parallel and rather have complementary time frames. The permutation test requires data to be centred to its mean before slope estimation. In Vieira and Creed (2013) the data could only be centred to its group mean. With replicated time series the data could alternatively be centred to its replicate's mean. This alternative standardization criterion was the second upgrade.

Matlab based software for the estimation of significance of slope differences from replicated time series data can be found in supplementary material.

2.2 The data

Juveniles of *Fucus serratus* were reared in 10×15cm plates on a controlled environment at 4 densities (650, 1334, 2000 and 2668 plants/m^2) during 171 days (Creed et al., 1998). Each treatment had 4 replicated parallel time series. All plates were sampled for total biomass (wet weight) at days 0, 19, 40, 71, 109, 139 and 171. Macrostages of *Laminaria digitata* were reared on 10×15cm plates with initial densities of 650, 1334, 2000, 2668 and 5186 plants/m^2 (Creed et al 1998). Each treatment had 4 replicated parallel time series. The density and stand wet biomass in each plate were accessed for 11 sampling times at approximately 19 day intervals. Presently, only the first 9 sampling times were used as during the last 2 the plates were suffering from both decreasing densities and biomass loss, i.e., there was no self-thinning. Experiments were carried out in culture tanks at Port Erin Marine Laboratory, Isle of Man. At the end a wet:dry weight relation was established.

3 Results

The self-thinning lines were estimated with the *x* and *y* variables standardized to their groups' mean and alternatively to their replicates' mean. As expected the slopes became flatter with increasing initial densities (Table 1). *Fucus serratus* reared at the sparser treatment (650 plants/m^2 initial densities) did not even show any mortality. As there was no change in density the line fit yielded a ∞ slope. The exception was treatment B (1334 plants/m^2 initial densities) from the *Laminaria digitata* data, for reasons unknown. Using different standardization methods yielded conspicuously different slopes. This was particularly true for the *Laminaria digitata* case where it was evident standardizing to the replicates' means widen the range of estimated slopes (Table 1). This latter standardization method also yielded better line fits (Fig. 1).

Tuning the permutation test more elaborate, i.e. starting with the simple permutation test, passing to the replicated time series upgrade with the group standardization and finally to the replicate standardization, always increased the tests ability to detect true differences (Table 2). With the *Fucus serratus* data the slopes were always significantly different. Nevertheless, significances become stronger as the test was more elaborately tuned. The sparser treatment was not used due to its 'no self-thinning' yielding ∞ slope. The simple

Table 1 Slopes obtained for the initial density treatments (groups) of *Fucus serratus* and *Laminaria digitata* self-thinning data according to the standardization method. (A) 650, (B) 1334, (C) 2000, (D) 2668 and (E) 5186 plants/m^2 initial densities.

Species	Standardization	A	B	C	D	E
Fucus serratus	Std. by group	-∞	-6,031	-2,453	-1,063	
	Std. by rep.	-∞	-5,810	-2,259	-1,102	
Laminaria digitata	Std. by group	-13,546	-6,252	-10,646	-9,539	-5,176
	Std. by rep.	-32,030	-7,744	-13,449	-12,4379	-5,079

Table 2 Significance of differences between slopes obtained by the permutation method for the *Fucus serratus* self-thinning data. (B) 1334, (C) 2000 and (D) 2668 plants/m^2 initial densities.

	Simple randomization		Time series randomization			
Standardization:	by group		by group		by replicate	
Groups:	C	D	C	D	C	D
B	0.0096	0.-------	0.0046	0.-------	0.0002	0.-------
C		0.0044		0.0011		0.0001

Table 3 Significance of differences between slopes obtained by the permutation method for the *Laminaria digitata* self-thinning data. Significant differences in **bold**. (A) 650, (B) 1334, (C) 2000, (D) 2668 and (E) 5186 plants/m^2 initial densities.

Groups	B	C	D	E
	Simple randomization			
	Standardized by group			
A	**0.0002**	0.177	0.2097	**0**
B		**0.0067**	0.0637	0.379
C			0.6543	**0**
D				**0.0099**
	Time series randomization			
	Standardized by group			
A	**0**	**0.0387**	0.1399	**0**
B		**0.0015**	**0.0484**	0.3416
C			0.6615	**0.0006**
D				**0.0056**
	Time series randomization			
	Standardized by replicate			
A	**0**	**0.0027**	**0.0063**	**0**
B		**0.0099**	**0.0445**	**0.0113**
C			0.7261	**0**
D				**0**

permutation test applied to the *Laminaria digitata* data recognized 5 out of 10 slope differences as significant (Table 3). This proportion increased to 7 out of 10 when the replicated time series upgrade with the group standardization was used, and to 9 out of 10 when the replicate standardization was used instead. These 9 slope

differences turned more significant as the test was tuned more elaborate. On the contrary, the remaining 1 slope difference turned more non-significant. This dynamic demonstrated the increased capacity of the permutation test in separating true from false differences as it was more elaborately tuned. The bootstrap/overlap method did not show this capacity (Table 4). When the simple randomization method was applied only 3 out of 10 slope differences were recognized as significant. But when the replicated time series upgrade was used all slope differences were recognized as significant, which is known not to be true.

Table 4 Overlap of the slopes' 95% confidence intervals for the *Laminaria digitata* self-thinning data. Slope differences are significant (**bold**) only when there is no overlap. (A) 650, (B) 1334, (C) 2000, (D) 2668 and (E) 5186 plants/m^2 initial densities.

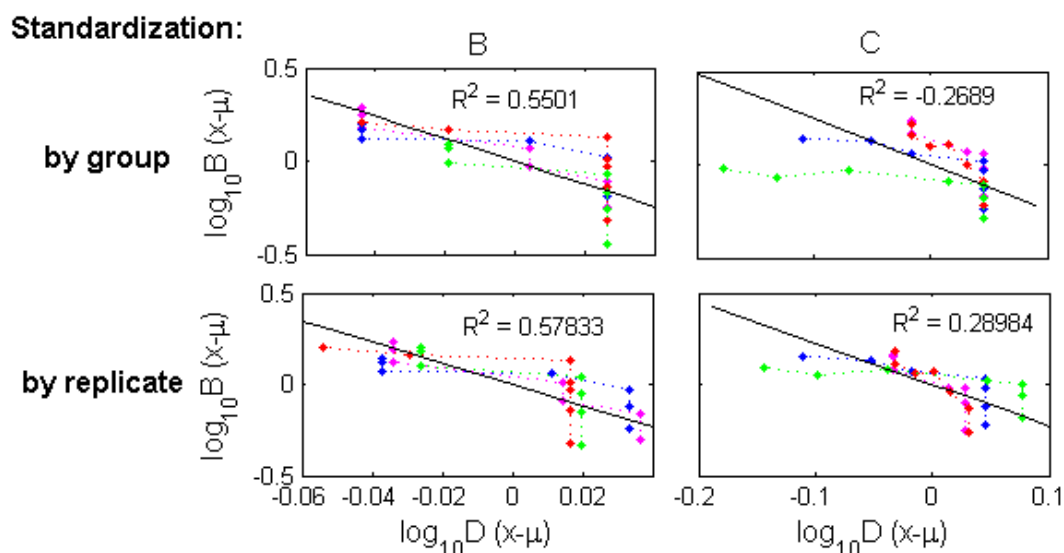| Groups	B	C	D	E
	Simple randomization			
A	1	1	1	1
B		**0**	0.2008	0.7023
C			1	**0**
D				**0**
	Timeseries randomization			
A	**0**	**0**	**0**	**0**
B		**0**	**0**	**0**
C			**0**	**0**
D				**0**

Fig. 1 Self-thinning lines for *Fucus serratus* estimated from data standardized by groups or by replicates. (B) 1334 and (C) 2000 plants/m^2 initial densities.

4 Discussion

Randomization methods, both permutations and bootstraps, have long been applied to time series analysis (Theiler et al., 1992; Hjellvik and Tjøstheim, 1995; Schreiber and Schmitz, 1996; Kugiumtzis, 2002; Politis 2003). They may or not preserve marginal distributions and/or correlations depending on the test objective. Nevertheless, the objective always falls upon testing the intrinsic properties of the own time series trajectory. This work developed a comparison between slopes taken from populations sampled at short lasting time series. It is an entirely different subject demanding a specific approach although sharing the permutations and bootstraps technical solutions.

Two upgrades were introduced into the previous method and software increasing outstandingly the permutation test's capacity to identify both true and false slope differences: (i) randomizing within bins containing only sampling units from the same time coordinate and (ii) shifting the standardization method to use the replicate's mean. From a strictly numerical point of view conjugating both was undoubtedly the best option. Nevertheless, the adequacy of standardizing to the replicate's mean is the object of theoretical statistical ecology consideration. In this example, when standardizing by replicates, the equal initial densities were displaced differently along the *x* axis, which may be argued incongruent. On the other hand, this was comparable to estimating a slope for each replicate and then averaging slopes. The fact remains that the test's performance increased outstandingly. The bootstrap/overlapped method was always an inadequate choice: applied with the simple randomization method it overlooked many truly significant differences (i.e. yielding false negatives) which had already been shown in Vieira and Creed (2013); whereas applied with the replicated time series randomization method it detected false significant differences (i.e. yielding false positives). Small sample sizes and/or large data scatter had been shown to render the simple randomization method keen to false negatives, irrespective of applied with permutation tests or the bootstrap/overlap tests (Vieira and Creed, 2013). Below a threshold it was impossible to detect significant differences. The present method and software solved this problem for the cases where data is organized into replicated time series. Many works on self-thinning applied older-dated tests to replicated time series of small sample sizes often finding slopes not being significantly different (Weiner and Whigham, 1988; Creed, 1995; Morris, 1996; Scrosati and DeWreede, 1997; Creed et al., 1998; Arenas and Fernández, 2000; Steen and Scrosati, 2004). These and similar cases should be reassessed by this new method as it demonstrated an outstanding capacity to identify true and false differences between slopes the previous methods did not have.

References

Anderson MJ, Legendre P. 1999. An empirical comparison of permutation methods for tests of partial regression coefficients in a linear model. Journal of Statistical Computation and Simulation, 62: 271-303

Arenas F, Fernández C. 2000. Size structure and dynamics in a population of *Sargassum muticum* (Phaeophyceae). Journal of Phycology, 36: 1012-1020

Dijksterhuis GB, Heiser WJ. 1995. The role of permutation tests in exploratory multivariate data analysis. Food Quality and Preference, 6: 263-270

Creed JC. 1995. Spatial dynamics of a *Himanthalia elongate* (Fucales, Phaeophyta) population. Journal of Phycology, 31: 851-859

Creed JC, Kain JM, Norton TA. 1998. An experimental evaluation of density and plant size in two large brown seaweeds. Journal of Phycology, 34: 39-52

Efron B, Tibshirani R. 1986. Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. Statistical Science, 1(1): 54-75

Ellison AM. 1989. Morphological determinants of self-thinning in plant monocultures and a proposal concerning the role of self-thinning in plant evolution. Oikos, 54: 287-293

Ferrarini A. 2011. A fitter use of Monte Carlo simulations in regression models. Computational Ecology and Software, 1(4): 240-243

Hjellvik V, Tjøstheim D. 1995. Nonparametric tests of linearity for time series. Biometrika. 82(2): 351-368

Hughes RN, Griffiths CL. 1988. Self-thinning in barnacles and mussels: the geometry of packing. The American Naturalist, 132(4): 484-491.

Jackson JE. 1991. A User's Guide to Principal Components. John Wiley & Sons, New York, USA

Jackson DA. 1993. Stopping rules in principal components analysis: a comparison of heuristical and statistical Approaches. Ecology, 74(8): 2204-2214

Keeley ER. 2003. An experimental analysis of self-thinning in juvenile steelhead trout. Oikos, 102: 543-550

Kugiumtzis D. 2002. Surrogate data test for nonlinearity using statically transformed autoregressive process. Physical Review E, 66: 025201

Lebart L. 2007. Which bootstrap for principal axes methods? Pages 581-588 in P. Brito, G. Cucumel, P. Bertrand and F. de Carvalho, editors. Selected Contributions in Data Analysis and Classification. 581-588, Springer, Berlin, Heidelberg, Germany

Legendre P. 2013. Model II Regression user guide, R Edition. http://cran.r-project.org/web/packages/lmodel2/vignettes/mod2user.pdf

Lonsdale WM, Watkinson AR. 1983. Plant geometry and self-thinning. Journal of Ecology, 71: 285-297

Manly BJF. 1991. Randomization and Monte Carlo Methods in Biology. Chapman & Hall, London, UK

Morris EC. 1996. Effect of localized placement of nutrients on root competition in self-thinning populations. Annals of Botany, 78: 353-364

Morris EC. 2003. How does fertility of the substrate affect intraspecific competition? Evidence and synthesis from self-thinning. Ecological Research, 18(3): 287-305

Pearson K. 1901. On lines and planes of closest fit to systems of points in space. Philosophical Magazine, 2: 559-572

Peres-Neto PR, Jackson DA, Somers KM. 2003. Giving meaningful interpretation to ordination axes: assessing loading significance in principal component analysis. Ecology, 84: 2347-2363

Peres-Neto, PR, Jackson DA, Somers KM. 2005. How many principal components? Stopping rules for determining the number of non-trivial axes revisited. Computational Statistics and Data Analysis, 49: 974-997

Politis DN. 2003. The impact of bootstrap methods on time series analysis. Statistical Science, 18(2): 219-230

Rincón PA, Lobón-Cerviá J. 2002. Nonlinear self-thinning in a stream-resident population of brown trout (*Salmo trutta*). Ecology, 83(7): 1808-1816

Rivera M, Scrosati R. 2008. Self-Thinning and size inequality dynamics in a clonal seaweed (*Sargassum lapazeanum*: Phaeophyceae). Journal of Phycology, 44: 45-49

Schreiber T, Schmitz A. 1996. Improved surrogate data for nonlinearity tests. Physical Review Letters, 77: 635-638

Sokal RR, Rohlf FJ. 1981. Biometry: the Principles and Practice of Statistics in Biological Research (2nd ed). W.H. Freeman and company. New York, USA

Stauffer DF, Garton EO, Steinhorst RK. 1985. A comparison of principal components from real and random data. Ecology, 66(6): 1693-1698

Steen H, Scrosati R. 2004. Intraspecific competition in *Fucus serratus* and *F. evanescens* (Phaeophyceae: Fucales) germlings: effects of settlement density, nutrient concentration, and temperature. Marine Biology,

144: 61-70

Theiler J, Eubank S, Longtin A, et al. 1992. Testing for nonlinearity in time series: the method of surrogate data. Physica D, 58: 77-94

Vieira VMNCS. 2012. Permutation tests to estimate significances on Principal Components Analysis. Computational Ecology and Software, 2(2): 103-123

Vieira VMNCS, Creed J. 2013. Estimating significances of differences between slopes. Computational Ecology and Software, 3(3): 44-52

Weiner J, Whigham DF. 1988. Size variability and self-thinning in wildrice (*Zizania aquatica*). American Journal of Botany, 75(3): 445-448

Weller DE. 1987. A reevaluation of the -3/2 power rule of plant self-thinning. Ecological Monographs, 57: 23–43

White J, Harper JL. 1970. Correlated change in plant size and number in plant populations. Journal of Ecology, 58: 467-485

Yoda K, Kira T, Ogawa H, et al. 1963. Self-thinning in overcrowded pure stands under cultivated and natural conditions (Intraspecific competition among higher plants. XI). Journal of Biology Osaka City University, 14: 107-129

Zhang WJ. 2010. Computational Ecology: Artificial Neural Networks and Their Applications. World Scientific, Singapore

Zhang WJ. 2011a. A Java program to test homogeneity of samples and examine sampling completeness. Network Biology, 1(2): 127-129

Zhang WJ. 2011b. A Java algorithm for non-parametric statistic comparison of network structure. Network Biology, 1(2): 130-133