Forest Ecology and Management 310 (2013) 120–133

Contents lists available at ScienceDirect

Forest Ecology and Management

journal homepage: www.elsevier.com/locate/foreco

On the evaluation of competition indices – The problem of overlapping samples

Rune Østergaard Pedersen *, Erik Næsset, Terje Gobakken, Ole Martin Bollandsås

Department of Ecology and Natural Resource Management, Norwegian University of Life Sciences, P.O. Box 1423, Ås, Norway

ARTICLE INFO

Article history:
Received 5 April 2013
Received in revised form 18 July 2013
Accepted 23 July 2013
Available online 7 September 2013

Keywords:
Plot edge bias
National forest inventory
Airborne laser scanning
Spatial autocorrelation
Competition indices

ABSTRACT

We discuss statistical concerns regarding evaluation of three types of individual tree competition indices (non-spatially, spatially explicit and based on airborne laser scanning), with special attention to the method of selection of competitors, and the spatial dependency and smoothing caused by overlapping samples of competitors. We quantify the effect of spatial autocorrelation on the effective sample size for various search methods, to reveal potential type I statistical error, for a sample of 557 plots of the Norwegian National Forest Inventory located in the Hedmark Country. Our results show that spatial autocorrelation mostly appears when competitors are selected within short search radii (3–4) m of the subject tree. However, when simultaneously accounting for the impact of spatial autocorrelation on the effective sample size between individual tree growth at breast height and competition, the effect appears to be neglect-able. This result is verified by testing if the change in the effective degrees of freedom in the Spearman rank correlation *t*-test for the Clifford et al. correction and a spatial bootstrap method, relative to the classical *t*-test effective degrees of freedom, are correlated with different measures of stand structure. This ratio showed no systematic variation across measures of plot micro and macro-scale variation like Loreyś mean height, the Gini-coefficient of tree basal area or volume per hectare. The conclusion seems indifferent to plot edge bias correction. A linear mixed model with spatial covariance structure confirmed that sample overlap does not cause serious spatial dependence. Moreover, a median based statistical test revealed a significant smoothing effect, with increasing search radii of competitors, which causes loss of variation. However, the smoothing does not decrease the ability of the competition indices to correlate with individual tree growth at breast height within search radii of 12 m, and thus it does not represent any problem for prediction.

1. Introduction

The use of competition indices (CI) has become an important part of forest management worldwide. A variety of models include competition as a parameter, for instance mortality (Eid and Tuhus, 2001), recruitment (Lexerød, 2005) and basal area growth models (Bollandsås and Næsset, 2009). The rationale behind the use of a CI is to capture the current social variation and competition pressure, which may reflect the growth conditions of the individual tree better than just the size of the tree, and is thus a term quantifying the struggle for survival which goes back to the fundamental work of Darwin. The use of competition is empirically justified through studies of self-thinning, a work pioneered by Yoda et al. (1963), leading to the establishment of the self-thinning rule, which has been the topic of numerous later publications e.g. (Westoby, 1984; Zeide, 1987). The self-thinning rule postulates that the average weight of plants per area unit is a function of the number of individuals (or accumulated density) risen to some power. The self-thinning rule allows as other stand aggregated expressions of competition like the space top height factor (Wilson, 1946) evaluation of competition on the stand level, but it does not say anything about the micro-scale variation. Thus, it cannot be used to model the growth of individual trees post micro-scale disturbances like thinning.

With time, competition has been extended as a scientific expression, reflecting the increasing understanding of the importance of competition in plant communities. Terms like the symmetry of competition (Weiner and Solbrig, 1984) and hereunder one-sided competition (Schwinning and Weiner, 1998) requires tools for quantification of competition at the single-tree level. One of the first to expand the use of competition from stand level to individual trees was Newnham (1964) who built upon the work of Staebler (1951), using the crown width of an open grown tree as a proxy for potential growth. His method attempts to model the growth of the individual tree as a function of the near spatial surroundings, which is a general characteristic of spatially explicit CIs. In contrast, non-spatially explicit CIs do not require knowledge

* Corresponding author. Tel.: +47 64965715.
E-mail address: rupe@umb.no (R.Ø. Pedersen).

0378-1127/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.foreco.2013.07.040

of the spatial location of neighboring trees (Munro, 1974). The introduction of spatially explicit CIs is founded on the belief that improvements in quantification of competition can be obtained by including spatial information of the competitors in the near surroundings of the subject tree. Variables like distances (Hegyi, 1974) and angels (Pukkala, 1989) between subject trees and competitors are included in the various CIs used in the literature, and Shi and Zhang (2003) applied local indicators of spatial autocorrelation as a measure of competition. Therefore, the spatially explicit CIs require a definition of a neighborhood, which means a zone or a sampling procedure, to identify the competing trees. This may for example be a radius around each subject tree (Hegyi, 1974; Braathe, 1984; Pukkala and Kolstroem, 1987), the m nearest competitors (Soares and Tomé, 1999) or a selection of trees proportional to size selected by a relascope (Daniels, 1976). Recently the idea of using a circular search radius around each tree was extended by Pedersen et al. (2012) to consider search radii in various heights above ground around the subject tree. The choice of sampling procedure should reflect the desire to collect those competing trees which explain the largest possible part of the vigor of the subject tree. Hence, the selection of competitors should be based on a biological rationale, and to some extent economic constrains. How can we evaluate the effectiveness of the selection procedures? Two widely used methods are the Pearson correlation coefficient and the use of a growth model (Tomé and Burkhart, 1989; Biging and Dobbertin, 1992; Soares and Tomé, 1999). The CIs are ranked according to their correlation with the basal area growth, or the improvement in some statistical measure like R^2 of the growth model. Alternatively, statistical inferences are derived for the parameters of a growth model including competition. The estimate of e.g. the Pearson correlation coefficient will be inflated by spatial autocorrelation (Chun and Griffith, 2013). In addition, the asymptotic test statistics like *t*-tests used to assess its significance assumes independent pairs of observations. Spatial autocorrelation will cause violation of this latter assumption (García, 1992; Fox et al., 2001; Fox et al., 2008) and increase the type I statistical error. This means that *p*-values will be too small, resulting in overoptimistic numbers of statistical significant tests (Schabenberger and Gotway, 2005). The magnitude of this effect is dependent on factors like stand age, species composition, if the stand is even or un-even aged and a natural or planted forest. In addition, the sampling of competing trees is also of importance, because individual trees share competitors in space.

Some authors have studied the sampling of competitors. Ledermann and Stage (2001) showed how the marginal contribution of adding a competitor changes with increasing distance between competitor and subject tree for some commonly used CIs. Miina and Pukkala (2000) used a simulation to optimize the search radius for various indices around the subject tree. The authors calculated the log-likelihood estimate for a selected growth function at various search distances in order to find the one that gave the best predictions of growth under different tree species and age combinations. Strictly speaking, this method is only valid if the effective sample size (ESS) is taken into consideration. The ESS shows how many observations the sample contains when adjusting for the spatial/longitudinal autocorrelation (Cressie, 1991). Many have noticed the lack of improvement when including distance in the CIs (Daniels et al., 1986; Biging and Dobbertin, 1995), and Stage and Ledermann (2008) attribute this to the small plot sizes used in some studies, but could this be due to spatial autocorrelation, and search criteria of competitors? Furthermore, how does the sampling procedure relate to spatial autocorrelation? For CIs based on airborne laser scanning (ALS) and for the ones applied elsewhere in the literature no investigation of the statistical properties of the sample of competing trees on the whole plot or in the search radius around each tree has been made, and an empirical study is needed to quantify the implications of spatial autocorrelation in practice. Anselin (1988) calls the spatial autocorrelation induced by splitting objects which were initially overlapping for non-hierachical aggregation, and it is according to Tiefelsdorf (1998) one of three important reasons for spatial autocorrelation. The two remaining reasons are miss-specified regression (modeling issue) and an underlying spatial process. The problem studied in the current paper is highly similar to non-hierarchical aggregation because of the shared competitors. One difference though is that the competitors do not affect different subject trees in the same manner because of the mathematical weighting done by the CIs.

Reed and Burkhart (1985) investigated how spatial autocorrelation varied in a stand of Loblolly Pine for different levels of competition, and found significant spatial autocorrelation for individual tree basal area using Moranś I and Gearyś C. However, the authors did not consider the sampling frame of the spatially explicit CIs utilized, nor did they use a bivariate measure of correlation such as the Pearson correlation coefficient. They investigated the 2 variables independently with the purpose of applying the estimated level of spatial autocorrelation of individual tree basal area in a stand generator in order to study its relationship with competition pressure and other stand characteristics. Furthermore, studies in boreal forests (Kuuluvainen et al., 1996) have revealed that spatial autocorrelation is present, and some studies have documented a positive spatial autocorrelation (Kenkel et al., 1989), which is in accordance with Anselin. Specifically Wyszomirski and Weiner (2009) showed by means of simulation how any departure from a Poison distributed "forest" may induce positive spatial autocorrelation in plant sizes.

The term plot edge bias is used to describe how competing trees outside the plot area, may influence the growth of the trees on the plot. If the effect is not considered it will lead to biased estimates of competition (Radtke and Burkhart, 1998). Therefore different plot edge bias corrections have been used in the literature (Pretzsch, 2009). It would be interesting to see how plot edge bias correction and spatial autocorrelation are related.

It is the objective of the current article to:

(1) Investigate the validity of the comparison between different CIs which has routinely been conducted in the literature when not considering spatial autocorrelation. Especially reveal potential differences between non-spatially and spatially explicit CIs.
(2) Investigate if the sampling procedure of competitors may influence commonly used statistical measures, so that the conclusions of an analysis of potential growth predictions from CIs may be invalid. In particular, quantify the effect in type I and II statistical error, i.e., the effect of the non-hierarchical aggregation.
(3) Test the effect of plot edge bias correction on the performance of CIs, with special attention on the relation between plot edge bias correction and spatial autocorrelation.

2. Materials

2.1. Field data

Two different datasets were used in the current study. The main study of competition was conducted in Hedmark Country (HC) located in South-Eastern Norway (27,340 km^2) (see also (Gobakken et al., 2012)). The Aurskog-Høland (AH) data was used solely for modeling of crown width (CW) since this parameter was not attainable in the HC data and it is described in electronic online Appendix C. CW was needed for quantification of competition when using ALS (described in Section 3.1.1).