

# Estimating sapling vitality for Scots pine (*Pinus sylvestris* L.) in Russian Karelia

D.G. Oreshkin<sup>a</sup>, J.P. Skovsgaard<sup>b,\*</sup>, J.K. Vanclay<sup>c</sup>

<sup>a</sup> St. Petersburg State University, Department of Geobotany and Plant Ecology, Faculty of Biology and Soil Science, Universitetskaya nab, 7/9, St. Petersburg 199034, Russia

<sup>b</sup> Danish Forest and Landscape Research Institute, Horsholm Kongevej 11, DK-2970 Horsholm, Denmark

<sup>c</sup> Center for International Forestry Research, PO Box 6596 JKPWB, Jakarta 10065, Indonesia

## Abstract

A new method is proposed for estimating vitality or growth potential for saplings of Scots pine (*Pinus sylvestris* L.), based on height, diameter and height increment. A two-stage process was used to establish the vitality index. The logarithms of height, diameter and height increment were regressed against age, to adjust for the wide range of ages present in our data (c. 10,000 saplings with ages spanning 4-50 years). Then principal component analysis was used to obtain coefficients, which were, in turn, standardized on each axis to provide a vitality index scaled in standard deviations. This standardized scale allows the rank of an individual in the population to be assessed, and draws attention to possible outliers. The use of age-adjusted residuals ensured that the estimator was independent of age, and stable over a wide age range. The first principal component indicates if a sapling is relatively tall (weight = 0.5), thick (w = 0.5) or fast-growing (w = 0.7) for its age. Most of the information is contained in the first principal component, but the second component, which explains about 10% of the variance, appears to offer some utility as an indicator of 'acceleration' due to changing conditions. The resulting measures of vitality have been useful for research and management in the dry lichen-moss pine forest in Russian Karelia, but are specific to this species, locality and ecotype. Further research and site-specific data are necessary to adapt the system to other situations.

## 1. Introduction

Although site productivity has attracted considerable attention in forest research (e.g., Hagglund, 1981; Vanclay, 1992; Skovsgaard, 1997), the related aspect of the productivity, thriftiness or vitality of individual trees has not received similar attention in field studies, except for the specific situation regarding nursery stock (e.g., Ritchie and Dunlap, 1980). Despite their importance in phyto-sociological and forestry studies dealing with the relationship between trees and their environment, estimates of vigour are usually subjective and qualitative, rather than objective and quantitative. We use the term *vitality* (cf. Mirkin et al., 1989; Zlobin, 1989) to refer to quantitative estimates of relative vigour of saplings and small trees. Specifically, a sapling has high vitality if its current potential for survival and continued growth is good. One indication of vitality is if height increment is comparatively large for the sapling's size (both height and diameter). Here we address only growth of survivors, and do not specifically address the issue of survival.

Zlobin (1970) suggested that vitality indices may be classified into:

1. physiological-biochemical indices, including those based on potential evapo-transpiration, and on water, nutrient, chlorophyll or photosynthate concentrations in plant tissues,
2. phenological indices, which reflect vitality in relation to current environmental conditions, and
3. morphological indices, which reflect the sum of all life processes, and are widely used

because of their ease of measurement.

Zlobin (1989) further classified these into metric and allometric, and static and dynamic indicators. Metric indices are based on physical dimensions of individuals, while allometric ones draw on the interrelationships between plant dimensions. Dynamic indices include data on morphogenesis and growth.

In this paper we propose a new method for estimating a morphological vitality index for Scots pine (*Pinus sylvestris* L.) saplings. We take an empirical approach and refer to Oreshkin (1996) for a more thorough discussion of the method's relative merits compared to other procedures for estimating vitality.

In spite of a desire to maximize information about the state of individuals, it is impractical to use all possible traits because of constraints on resources and time. Thus it is important to choose traits that are information-rich, providing sufficient information about vitality with minimal cost. To meet this criterion, the selected traits should

1. capture the full range of variation evident in the attributes of interest,
2. have low correlations with other selected traits, and should
3. be well correlated with other traits of interest that remain unsampled.

Zlobin (1970) used height increment, diameter and the apical:lateral shoot length ratio to estimate vitality of spruce saplings, but for pine saplings, he adopted a simpler approach based on height increment and the number of branches in a whorl. Later, Zlobin (1981) promoted principal component analysis (PCA) for use in the selection of efficient traits (i.e., choose traits which have the largest coefficients in the eigenvectors), and for determining the principal components (PCs), which may themselves form the vitality indicators. Clearly, the number of useful PC axes depends on the variance explained by each axis. Zlobin used two PCs, and subdivided his data into three classes on each axis, creating a total of nine vitality classes which were ranked subjectively.

PCA is a useful technique for examining vitality, particularly since the traits age, height, diameter and height increment are all highly correlated. The orthogonal nature of the principal components allows different aspects of vitality to be quantified and assessed in terms of information content. We use PCA to create vitality indices that:

1. are based on few traits that reflect different aspects of vitality;
2. can be measured easily, even where many individuals are involved;
3. effectively integrate the various traits of interest;
4. are independent of age (n.b., especially when initial traits are age-dependent).

## 2. Data

The material for this study comprises a selection of Scots pine saplings in dry moss-lichen pine forests on sandy soil in Russian Karelia. Growth in these forests is slow, and the expected size of a tree at age 50 years may be as little as 10 m in height, and 10 cm in diameter (Fig. 1).

Data were collected near the Suistamo railway station (in 1994) and near the town of Pitkaranta (in 1995). Within sixteen plots, 7032 saplings were sampled, spanning a wide range of tree age, size and stand density (Fig. 1). For each sapling, height (cm), stem diameter (mm) at 20 cm above ground, and height increment (mm) during the previous year were measured. The age of the individuals was assessed by the number of nodes. Height increment was measured as the distance between the nodes corresponding to the previous year, because measurement did not coincide with the end of the growth season, and the terminal shoot could continue to grow after the measurement was taken. Seedlings less than 20 cm in height were omitted from this sample. The data were supplemented with 2666 observations taken by Yastrebov and Poznanskaya (1993) in the same vicinity. The combined data set included 9698 specimens ranging in age from 4 to 50 years.

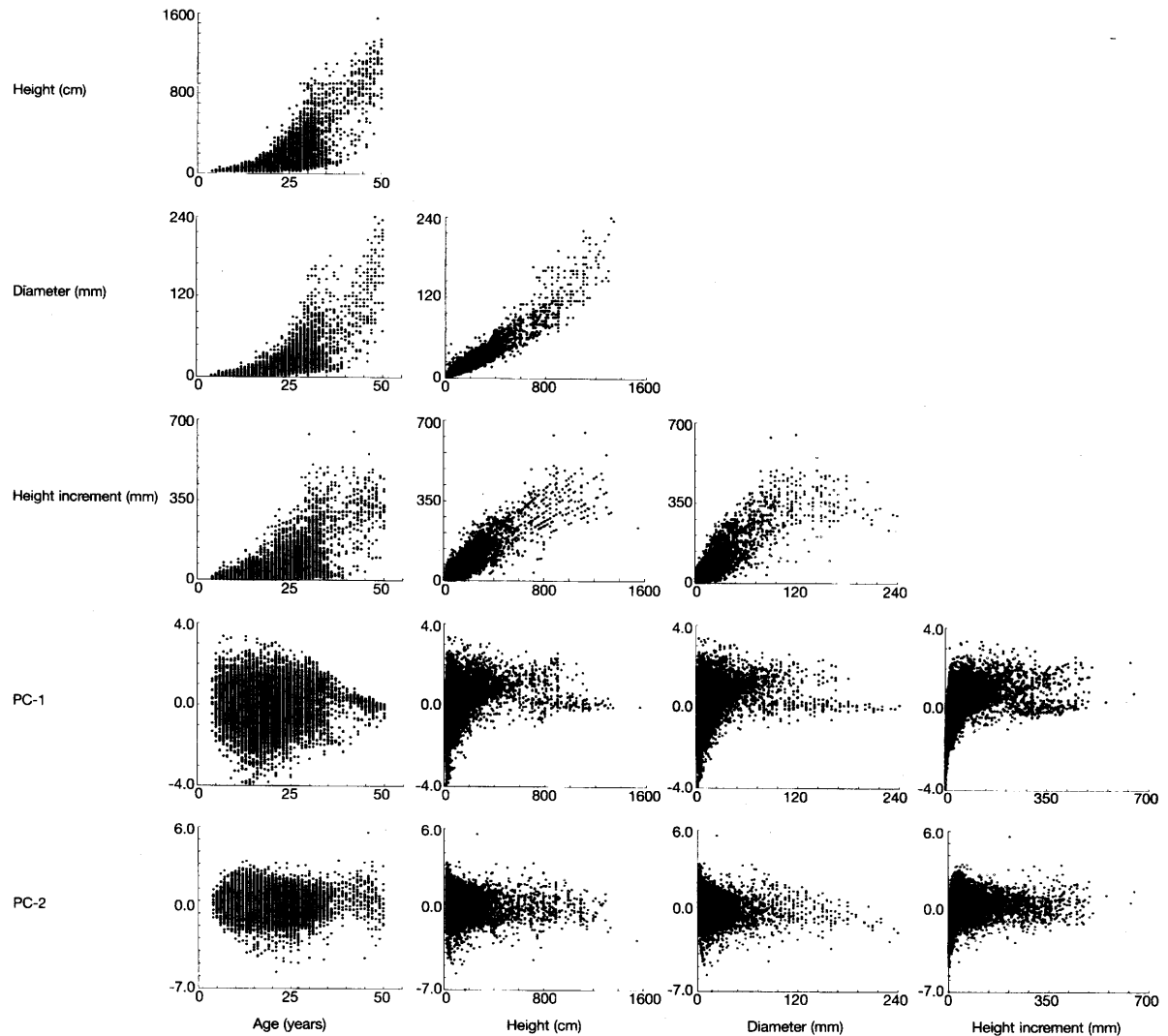


Fig. 1. Scatterplot matrix of raw data and PC scores showing the extent of the data and the correlation between sapling characteristics.

### 3. Methods

Efficient estimation of vitality faces a dilemma: that the best indicators are impractical because of measurement difficulties, while traits that are easily measured address only one aspect of a multidimensional phenomenon. One solution explored here, is to use PCA to combine several simple indicators into a more holistic index.

The trait loadings (i.e., the coefficients of the eigenvectors) of the PCAs can be examined to infer the biological significance of the index. Using PCA, several traits may be combined into a single vitality estimate. Many morphological traits like height, diameter and height increment depend on age. Thus, trait loadings and the directions of the PCs may depend on age, and should be adjusted for age in the estimation procedure (since age per se is not an indicator of vitality). Hence, we adopted the following procedure for preparing vitality indices:

1. describe the relationship between average trait value and age;
2. compute principal components within each age class;
3. analyze trait loadings and interpret the PC axes; 4. investigate age-dependence in direction of the axes and in the magnitude of the loadings.

Because we wished to examine the stability of the PCs across a large range of ages, we standardized the axes to ensure comparable results. In particular, if the variance within PC axes depends on age, coordinates will not be comparable at different ages. Thus we standardized our vitality index,  $v$ , using the following equation:

$$v_i = Z_i/S_{it}$$

where  $v_i$  is the vitality estimate on the  $i$ th PC,  $z_i$  is the coordinate of the individual on that axis, and  $s_{it}$  is the standard deviation at the age  $t$ . The scaling by standard deviation simplifies comparisons, and draws attention to extreme deviations which may be due to sampling errors. This procedure is comparable to the use of standardized residuals, but may be more appropriate where the variance is non-homogeneous.

#### 4. Results

The size distribution of saplings in our data was skewed, and was normalized using a logarithmic transformation (zero values of height increment were changed to 1 mm). The natural logarithm of height, diameter and height increment was closely and consistently correlated with age throughout the range 4 to 50 years (Fig. 2, Table 1). The data show some heteroscedacity, with standard deviations (after log transformation) increasing with age up to 39 years (Fig. 3). However, we used unweighted linear regression, and residuals were used in the subsequent PCA.

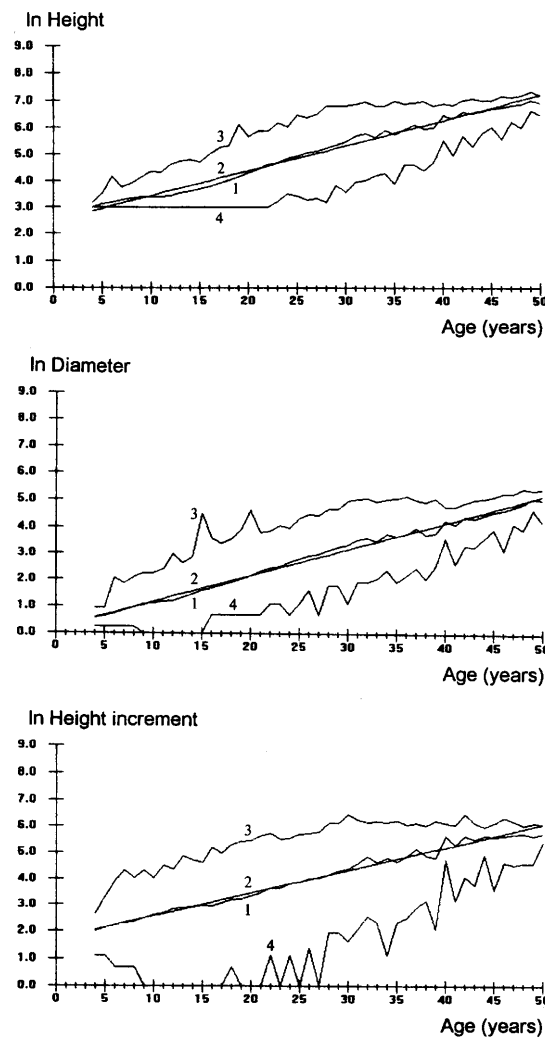
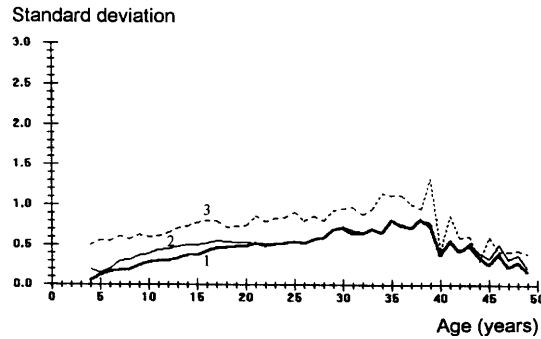


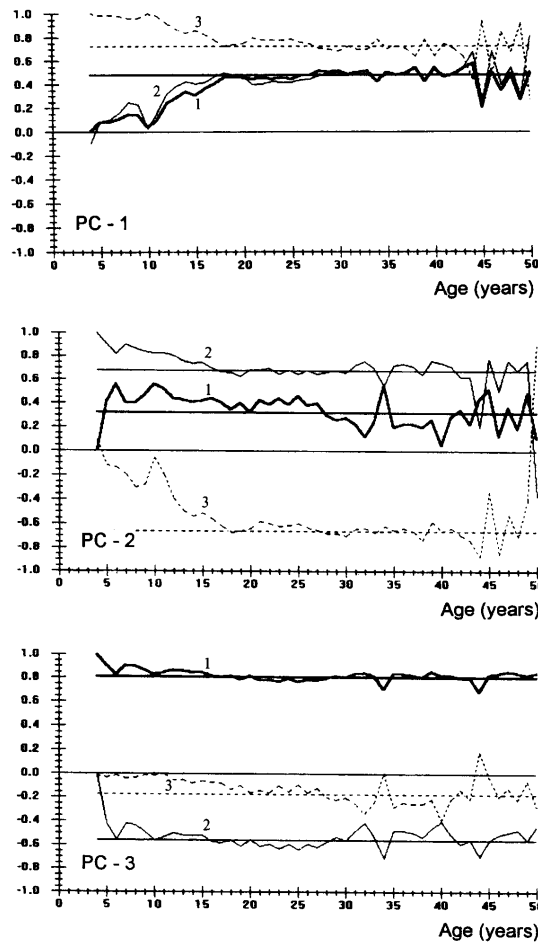
Fig. 2. Natural logarithm of height, diameter, and height increment vs. age. Legend: 1= average, 2 = linear regression on age, 3 = maximum values, 4 = minimum values.

**Table 1.** Parameters of the linear regressions of sapling trait (height, diameter and height increment; lines labelled "2" in Fig. 2) and its standard deviation against sapling age (t); [*r* is the correlation coefficient]

Trait	$\beta_0$	$\beta_1$	<i>r</i>	$\gamma_0$	$\gamma_1$
1. ln(height)	2.4944	0.0943	0.972	0.3166	0.0310
2. ln(diameter)	0.1590	0.1007	0.974	0.5855	-0.0076
3. ln(height increment)	1.7042	0.0870	0.957	0.2321	-0.0020



**Fig. 3.** Standard deviations of ln height (1), ln diameter (2), and ln height increment (3) vs. age.



**Fig. 4.** Trait loadings vs. age for first (PC-1), second (PC-2) and third (PC-3) principal component. Legend: 1= height, 2= diameter, 3 = height increment.

**Table 2.** Average trait loadings on PC axes, for calculating vitality. Notice that the coefficients on PC-2 have been negated (cf. Fig. 4) so that positive values indicate greater vitality

Axis	PC-1	PC-2	PC-3
1. ln(height)	0.48687	-0.31679	0.81400
2. ln(diameter)	0.47563	-0.68548	-0.55126
3. ln(height increment)	0.73262	0.65556	-0.18307

To investigate age-dependence, residuals were stratified into 47 one-year age classes, and PCs were computed for each class. The PC axes remained stable throughout most of the sampled age range (Fig. 4). The instability evident after age 43, may be related to the small sample size and limited number of suppressed saplings in the sample. Consistent changes in trait loadings are evident only for saplings younger than 18 years, especially in the first and second PCs. This is attributed to the absence of saplings under 20 cm in our sample, which truncates the variance (Fig. 2a), and to the measurement error inevitable when increments approach the resolution of the measuring instrument (1 mm).

Average trait loadings were calculated from loadings derived from ages 22 to 43 years (Table 2). These average loadings maintained the dimensions (standard deviations) and orthogonal nature of the original vectors. These provided the basis for the vitality estimate:

$$v_j = \sum_{i=1}^3 \frac{f_{ij} [\ln x_i - (\beta_{0i} + \beta_{1i}t)]}{\gamma_{0i} + \gamma_{1i}t}$$

where  $v_j$  is the  $j$ 'th component of the vitality index,  $x_i$  is the initial trait (height, diameter and diameter increment respectively),  $f$  is trait loading (Table 2),  $t$  is age (years), and the  $\beta$ s and  $\gamma$ s are the parameters of the regression of trait and standard deviation respectively, on age (Table 1). This equation provides three independent and orthogonal estimates of sapling vitality, ranked in decreasing order of information content. These should be independent of age, height and diameter, but some age-dependence may exist for trees younger than c. 20 years (Fig. 4).

Preliminary tests at the stand level show negative correlations between stand basal area and the stand average value of the first vitality component (-0.894,  $p < 0.05$ ). The second and third axes exhibited smaller correlations, consistent with expectations that the first PC captures most of the information about average vitality.

## 5. Discussion

The method we propose provides an index of vitality in three dimensions. The first dimension has positive loadings for all traits. Since an increase in any trait will increase the vitality estimate on this axis, saplings which are tall, thick, or fast-growing (in height) for their age will be assigned high vitality. Thus, the first axis may be interpreted as overall vitality. Note that the height increment loading (0.7) exceeds those for height and diameter (both c. 0.5), and thus has a greater influence on the vitality estimate. The first axis explains, on average, 88.0% of height variance, 79.4% of diameter variance and 92.5% of height increment variance.

The coefficients on the second PC axis were negated, so that positive values would indicate higher vitality, consistent with the other PCs. After the sign changes (Table 2; cf. Fig. 4), this PC has a positive loading for height increment and negative loadings for height and diameter. These latter two traits effectively integrate growth conditions throughout the plant's life, in contrast to height increment which characterises growth during the current season. Thus, this second axis emphasizes current growth rather than periodic or mean growth, and may be viewed as an indicator of 'acceleration', when a sapling's circumstances change. This axis explains, on average, only 5.8% of height variance, 17.5% of diameter variance and 7.4% of height increment variance.

The third axis explains very little of the total variance (6.2% of height, 3.1% of diameter, and 0.1% of height increment). The loading for height increment is negligible, so it essentially contrasts the dimensions of height and diameter. Thus it appears to be an analogue of vitality estimators based on the height-diameter ratio (Larin and Pautov, 1989; Yastrebov and Poznanskaya, 1993).

Compared to other, similar approaches our method has the advantage of accounting for how the traits in question change with age under different environmental conditions. This is especially important for studies of natural forests, where saplings of similar size may span a wide range of ages. Yastrebov and Poznanskaya (1993) attempted to circumvent age-dependency by using mean annual increment in height and diameter, but their solution is not entirely satisfactory, since height and diameter increments are also age-dependent. Kuuluvainen et al. (1993) calculated a linear regression of sapling height increment on height, and used the residuals from this regression to estimate growth potential for seedlings of Scots pine in eastern Finland.

The equation coefficients given in Tables 1 and 2 are specific to Scots pine in the dry moss-lichen pine forest in Russian Karelia. Further analyses and sitespecific data are necessary to adapt this system to other species and ecotypes. While the general procedure should be generalizable to other sites, the specific equations and transformations need not apply elsewhere. For the particular data set used in this analysis, the logarithmic transformation was appropriate, and allowed a simple linear regression with age. We do not advocate that this relationship will be generally applicable: in the present study, this

relationship is used to adjust traits according to tree age (mainly in the range 5-40 years), and we emphasize that it should not be used for long-term growth forecasts.

Comparison of the age-height and age-height increment relationships (Table 1) reveals that the height increment recorded for 1993-1995 is about half that expected in the longer term. We offer no explanation for this phenomenon. There are several possible causes, and further research is required to elucidate the most plausible explanation.

## 6. Conclusion

The ability of principal component analysis to combine and rank information from several traits into orthogonal vectors makes the estimation of vitality more tractable. We advocate the use of linear regression to adjust selected traits for sapling age, as this provides more stable principal components with relatively constant coefficients across a wide age range. However, the method remains empirical, specific to particular situations, and dependent upon relatively large amounts of data.

## Acknowledgements

This study was made possible by financial support from the International Science Foundation (grants R6E000 and R6E300) and from the Fundamental Natural Science Competition Centre of St. Petersburg State University. Thanks are due to A.B. Yastrebov (St. Petersburg) for providing supplementary data, and for helpful comments on the manuscript. We are grateful to the participants of expeditions for their help with sampling. Per B. Brockhoff, Jens Lund (both Copenhagen) and two referees provided helpful criticism of the manuscript.

## References

- Hagglund, B., 1981. Evaluation of forest site productivity. *Forestry Abstracts* 42 (11), 515-527.
- Kuuluvainen, T., Hokkanen, T.J., Jarvinen, E., Pukkala, T., 1993. Factors related to seedling growth in a boreal Scots pine stand: a spatial analysis of a vegetation-soil system. *Can. J. For. Res.* 23, 2101-2109.
- Larin, V.B., Pautov, Ju.A., 1989. The development of young coniferous forests on slashes. Nauka, Leningrad. 143 pp. (In Russian).
- Mirkin, B.M., Rozenberg, G.S., Naumova, L.G., 1989. Dictionary of notions and terms of modern phytocoenology. Nauka, Moscow. 222 pp. (In Russian).
- Oreshkin, D.G., 1996. The vitality estimation of pine (*Pinus sylvestris* L.) saplings. In: Skovsgaard, J.P., Johannsen, V.K. (Eds.), Modelling regeneration success and early growth of forest stands. Proceedings from the IUFRO conference, held in Copenhagen, 10-13 June 1996. Danish Forest and Landscape Research Institute, Horsholm, pp. 85-96.
- Ritchie, G.A., Dunlap, J.R., 1980. Root growth potential: its development and expression in forest tree seedlings. *N.Z. J. For. Sci.* 10, 218-248.
- Skovsgaard, J.P., 1997. Management of Sitka spruce without thinnings. An analysis of stand structure and volume production of unthinned stands of Sitka spruce (*Picea sitchensis* (Bong.) Carr.) in Denmark. Danish Forest and Landscape Research Institute, The Research Series, Vol. 19. 525 pp.
- Vanclay, J.K., 1992. Assessing site productivity in tropical moist forests: a review. *For. Ecol. Manage.* 54, 257-287.
- Yastrebov, A.B., Poznanskaya, A.A., 1993. The analysis of influence of tree stand on saplings in Karelian pine forests. *Botanichesky Zhurnal* 78 (5), 123-132. (In Russian).
- Zlobin, Ju.A., 1970. Quality estimation of tree saplings. *Lesovedenie* 3, 96. (In Russian).
- Zlobin, Ju.A., 1981. On viability degrees of plants. *Zhurnal Obschey Biologii* 42 (4), 429-505. (In Russian).
- Zlobin, Ju.A., 1989. Theory and practice of vitality estimation in plant coenopopulations. *Botanichesky Zhurnal* 74 (6), 769781. (In Russian).