Bulgarian Journal of Agricultural Science, 17 (No 2) 2011, 182-190
Agricultural Academy

FACTOR ANALYSIS OF NURSERY SEEDLING DATA IN DIFFERENT COMPOST SUBSTRATES

At. PAPAIOANNOU[1], K. KITIKIDOU[2] and D. SEILOPOULOS[1]
[1]*Aristotle University, Thessaloniki, Greece*
[2] *Dimokritos University, Orestiada, Greece*

Abstract

PAPAIOANNOU, At., K. KITIKIDOU and D. SEILOPOULOS, 2011. Factor analysis of nursery seedling data in different compost substrates. *Bulg. J. Agric. Sci.,* 17: 182-190

Analyzing units described by a mixture of sets of quantitative and categorical variables is a relevant challenge. Principal components analysis was used to include these two types of variables in order to study the correlations of a number of forest soil variables by grouping the variables in factors. Seedlings of four economically important and ecologically different species, *Quercus pubescens, Pinus maritima, Pinus nigra and Pinus brutia*, were grown in paper pots filled with either mixtures of sawdust, straw, byproducts of rice or sugar beet substrates. The variables used in the analysis were: tree species, composts, height, diameter, weight above ground and underground weight. The principal components method of factors extraction (PCA) begins by finding a linear combination of variables (a component) that accounts for as much variation in the original variables as possible. It then finds another component that accounts for as much of the remaining variation as possible and is uncorrelated with the previous component, continuing in this way until there are as many components as original variables. A few components accounted for most of the variation, and these components were used to replace the original variables. In the nursery, species responded primarily to substrate type. However, there were no interactions with nursery treatments.

Key words: Seedlings production, composts, paper pots, principal components analysis.

Introduction

Reforestation is an increasingly popular activity due to the abandonment of unproductive cattle pastures and government incentive programs that support tree planting (Schelhas et al., 1997). To promote use of native species in forestry plantations and increase profitability for farmers, species screening and provenance trials for reforestation were established (Butterfield, 1993, 1995; Butterfield and Fisher, 1994). While research efforts often focus on identification of appropriate genotypes for reforestation, more attention needs to be given to seedling quality. Despite the enormous diversity of tree species, the majority of non-industrial seedlings are produced in the same way, using 500-1500 cm^3 perforated plastic (poly) bags with soil. The soil often contains high proportions of clay, has poor structure, and is low in plant nutrients. Generally little or no organic matter is

Corresponding author: kkitikid@fmenr.duth.gr

incorporated. Due to poor substrates, plant growth is slow, extending into two nursery seasons, raising costs for the nursery, and inevitably reducing plant growth in the field. Paper pots are used because they are inexpensive and readily available. They can cause root coiling, the spiral growth of roots along the smooth sides and bottom of the bag; this root deformation can cause toppling or basal sweep several years after planting, thus greatly lowering the value of the plantation (Mason, 1985; Liegel and Venator, 1987; Sharma, 1987; Josiah and Jones, 1992).

Stumps plants are an additional stock type produced in bareroot beds. They are derived from trees usually over 1.5 m tall by trimming major portions of the stem and root system (ideally 10 cm shoot and 15 cm root remain) after lifting. They are used for several species and they are popular because they require little maintenance in the nursery and are easy for landowners to transport.

Limitations to seedling growth on abandoned pastures or agriculture fields include weed competition, soil compaction and low fertility soils. Weed growth is persistent throughout the year due to favorable environmental conditions. Manual weeding with a machete is the most common vegetation control in plantations and weeding frequency depends on the cost and availability of labor (Rheingans, 1996). The few landowners, who have sufficient economic resources and are more commonly used in agriculture than in forestry, use herbicides and fertilizers.

Early establishment is important for smallholder landowners in order to reduce weeding costs and possibly reduce time to harvest. Early plant growth is regulated by the conditions at the planting site, and by the degree to which a plant's phenotypic characteristics are adapted to a planting site (Burdett et al., 1983). High quality seedlings show substantial height growth the first year of planting, thus expressing their full genetic potential (Rose et al., 1990). They capture the site quicker, therefore allowing fuller expression of site potential (Fry and Poole, 1980). In contrast, use of poor planting stock can lower plantation survival and growth, increase site maintenance costs, and reduce confidence in reforestation. The use of the most appropriate planting stock can help overcome site limitations, while early and intensive site management can also accelerate seedling growth (Ladrach, 1992). Increasing investment in nursery stock relative to investment in site preparation can increase financial returns on overall reforestation investments (South et al., 1993).

In order to capitalize on advances made in reforestation with native species, seedling production techniques should be improved. Assessing outplanting performance must be an integral part of defining and adjusting target seedling characteristics. The objective of this research was to determine, for four widely planted species with contrasting ecological characteristics, how different substrates affect seedling growth in the nursery and early growth in the field.

Materials and Methods

Field study

The nursery was located at Nea Chalkidona (Northern Greece). Four rapidly growing species that are commonly planted in the area were studied: *Quercus pubescens*, *Pinus maritima*, *Pinus nigra* and *Pinus brutia*. For each species, 14 treatments (composts) were replicated. The composts composition and ingredients ratio are given in Table 1.

Seeds were collected from phenotypically superior trees and each paper pot contained seed from a single mother tree. Seed was directly sown in the paper pots between March and April 1994, depending on species.

Each replicate of each treatment contained 480 plants (120 from each species). Paper pots were spaced so that the density of all seedlings was equal.

At the end of the growing period, trees height, diameter, weight above ground and underground weight were measured. Tree mortality due to accidental cattle grazing or herbicide was treated as missing data.

Table 1
Substrate properties for nursery

No	Mixture				Ratio			
1	Sawdust	Straw	Forest soil		45	45	10	
2	Sawdust	Rice canes	Forest soil		45	45	10	
3	Rice peels	Straw	Forest soil		45	45	10	
4	Rice peels	Rice canes	Forest soil		45	45	10	
5	Sugar beets	Straw	Forest soil		45	45	10	
6	Sugar beets	Rice canes	Forest soil		45	45	10	
7	Sugar beets	Straw	Rice peels	Forest soil	40	25	25	10
8	Sugar beets	Rice canes	Rice peels	Forest soil	40	25	25	10
9	Sugar beet canes	Sawdust	Forest soil		45	45	10	
10	Sugar beet canes	Sugar beets	Forest soil		45	45	10	
11	Rice peels	Sugar beet mud	Forest soil		55	40	5	
12	Sugar beets	Sugar beet mud	Forest soil		55	40	5	
13	Sawdust	Rice canes	Rice peels	Forest soil	40	40	10	10
14	Rice peels	Rice canes	Sugar beet mud	Forest soil	40	40	10	10

Data analysis

The variables used in the analysis were:

Variable 1: Species (1= *Quercus pubescens*, 2=*Pinus maritima*, 3=*Pinus nigra*, 4=*Pinus brutia*)

Variable 2: Composts (1-14, according to Table 1)

Variable 3: Height (Seedling height in cm)

Variable 4: Diameter (Seedling height in cm)

Variable 5: Weight1 (weight above ground in g)

Variable 6: Weight2 (underground weight in g)

Principal components (PCA) and common factor (MLA for maximum likelihood and IPA for iterated principal axis) analyses are methods of decomposing a correlation or covariance matrix. Although principal components and common factor analyses are based on different mathematical models, they can be used on the same data and both usually produce similar results. Mathematicians and psychometricians have known about the factor indeterminacy problem for decades. For a historical review of the issues, see Steiger (1979); for a general review, see Rozeboom (1982). For further information refer Harman (1976), Mulaik (1972), Gnanadesikan (1977), or Mardia, Kent, and Bibby (1979), Afifi, May, and Clark (2004), Clarkson and Jennrich (1988), or Dixon (1992). Factor analysis is often used in exploratory data analysis to:

- Study the correlations of a large number of variables by grouping the variables in “factors” so that variables within each factor are more highly correlated with variables in that factor than with variables in other factors.
- Interpret each factor according to the meaning of the variables.
- Summarize many variables by a few factors. The scores from the factors can be used as input data for *t* tests, regression, ANOVA, discriminate analysis, and so on. Often the users of factor analysis are overwhelmed by the gap between theory and practice. In this chapter, we try to of-

Table 2
Mean height for all species-compost combinations

		Species			
		Quercus pubescens	*Pinus maritima*	*Pinus nigra*	*Pinus brutia*
		Mean	Mean	Mean	Mean
Compost	Compost 1	14.07	8.796	5.487	9.471
	Compost 2	11.625	6.404	4.471	6.108
	Compost 3	11.483	9.454	4.663	7.033
	Compost 4	10	6.85	4.529	4.988
	Compost 5	12.515	7.492	4.075	6.25
	Compost 6	11.533	8.215	4.408	7.213
	Compost 7	11.653	7.546	4.696	7.446
	Compost 8	11.348	7.907	4.3	7.583
	Compost 9	14	7.858	4.521	8.463
	Compost 10	10.745	8.608	5.025	8.471
	Compost 11	8.236	7.667	4.096	6.942
	Compost 12	9.154	6.817	3.333	7.754
	Compost 13	8.853	8.027	3.833	7.283
	Compost 14	8.723	7.958	4.079	5.467

fer practical hints. It is important to realize that you may need to make several passes through the procedure, changing options each time, until the results give you the necessary information for your problem. If you understand the component model, you are on the way toward understanding the factor model, so let us begin with the former.

Results and Discussion

In this study, we need to be able to appraise the species development to determine the likely competition for our composts.

Mean growth for all species-compost combinations is given in Tables 2-5.

The best species-compost combination seems to be *Quercus pubescens*-Compost1 (sawdust – straw – forest soil, 45-45-10), concerning height, diameter and weight above ground development, while Compost10 (sugar beet canes – sugar beets – forest soil, 45-45-10) gave best results concerning underground weight, for the same species. Generally, species responded primarily to substrate type, with *Quercus pubescens* having best growth.

Communalities indicate the amount of variance in each variable that is accounted for. Initial communalities are estimates of the variance in each variable accounted for by all components or factors. For principal components extraction, this is always equal to 1.0 for correlation analyses. Extraction communalities are estimates of the variance in each variable accounted for by the components. The communalities in this table are all high, which indicates that the extracted components represent the variables well. If any communality is very low in a principal components extraction, we may need to extract another component. The extraction communalities for this solution (Table 6) are acceptable, although the lower values of Species show that this variable doesn't fit as well

Table 3
Mean diameter for all species-compost combinations

		Species			
		Quercus pubescens	*Pinus maritima*	*Pinus nigra*	*Pinus brutia*
		Mean	Mean	Mean	Mean
Compost	Compost 1	6.625	1.943	1.609	2.206
	Compost 2	4.956	2.106	1.486	1.769
	Compost 3	4.512	1.71	1.475	1.912
	Compost 4	3.832	1.522	1.133	1.067
	Compost 5	4.337	1.604	1.25	1.489
	Compost 6	4.839	1.422	1.303	1.727
	Compost 7	5.112	1.806	1.389	1.857
	Compost 8	4.435	1.309	1.336	1.888
	Compost 9	5.122	1.188	1.356	1.78
	Compost 10	4.142	1.496	1.371	2.092
	Compost 11	4.338	1.289	1.29	1.829
	Compost 12	4.286	1.663	1.473	1.949
	Compost 13	3.796	1.451	1.385	2.105
	Compost 14	3.708	1.444	1.131	1.382

Table 4
Mean weight above ground for all species-compost combinations

		Species			
		Quercus pubescens	*Pinus maritima*	*Pinus nigra*	*Pinus brutia*
		Mean	Mean	Mean	Mean
Compost	Compost 1	2.197	0.736	0.286	0.791
	Compost 2	1.359	0.344	0.192	0.359
	Compost 3	1.49	0.361	0.223	0.539
	Compost 4	1.215	0.242	0.093	0.109
	Compost 5	1.655	0.414	0.131	0.304
	Compost 6	1.852	0.275	0.183	0.477
	Compost 7	1.733	0.47	0.208	0.493
	Compost 8	1.861	0.247	0.272	0.734
	Compost 9	2.04	0.229	0.226	0.621
	Compost 10	1.811	0.332	0.248	0.701
	Compost 11	1.134	0.241	0.196	0.582
	Compost 12	1.577	0.405	0.227	0.553
	Compost 13	1.028	0.296	0.235	0.631
	Compost 14	0.979	0.263	0.09	0.238

Table 5
Mean underground weight for all species-compost combinations

		Species			
		Quercus pubescens	*Pinus maritima*	*Pinus nigra*	*Pinus brutia*
		Mean	Mean	Mean	Mean
Compost	Compost 1	4.828	0.205	0.198	0.437
	Compost 2	3.449	0.226	0.261	0.37
	Compost 3	4.044	0.176	0.242	0.501
	Compost 4	4.029	0.211	0.181	0.16
	Compost 5	4.335	0.219	0.216	0.557
	Compost 6	4.787	0.221	0.373	0.608
	Compost 7	6.169	0.31	0.355	0.748
	Compost 8	5.692	0.211	0.448	0.982
	Compost 9	5.777	0.243	0.267	0.607
	Compost 10	6.16	0.133	0.213	0.62
	Compost 11	3.585	0.185	0.255	0.594
	Compost 12	5.785	0.301	0.302	0.478
	Compost 13	3.328	0.249	0.175	0.533
	Compost 14	3.712	0.159	0.13	0.213

Table 6
Communalities

	Initial	Extraction
Species	1	0.522
Compost	1	0.959
Height, cm	1	0.396
Diameter, cm	1	0.723
Weight above ground, g	1	0.868
Underground weight, g	1	0.872

as the others.

The variance explained by the initial solution, extracted components, and rotated components is displayed in Table 7. This first section of the table shows the Initial Eigenvalues. The Total column gives the eigenvalue, or amount of variance in the original variables accounted for by each component.

The % of Variance column gives the ratio, expressed as a percentage, of the variance accounted for by each component to the total variance in all of the variables. The Cumulative % column gives the percentage of variance accounted for by the first n components. For example, the cumulative percentage for the second component is the sum of the percentage of variance for the first and second components. For the initial solution, there are as many components as variables, and in a correlations analysis, the sum of the eigenvalues equals the number of components. We have requested that eigenvalues greater than 1 be extracted, so the first two principal components form the extracted solution. The second section of the table shows the extracted components. They explain nearly 72% of the variability in the original ten variables, so we can considerably reduce the complexity of the data set by using these components, with only a 28% loss of information. The rotation maintains the cumulative percentage of variation explained by the extracted components, but that variation

Table 7
Total Variance Explained

Component	Initial eigenvalues			Extraction sums of squared loadings			Rotation sums of squared loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	3.315	55.253	55.253	3.315	55.253	55.253	3.311	55.186	55.186
2	1.024	17.073	72.326	1.024	17.073	72.326	1.028	17.14	72.326
3	0.715	11.911	84.238						
4	0.58	9.662	93.899						
5	0.298	4.974	98.873						
6	0.068	1.127	100						

Table 8
Rotated Component Matrix

	Component	
	1	2
Species	-0.713	-0.118
Compost	-0.014	0.979
Height, cm	0.595	-0.206
Diameter, cm	0.847	-0.068
Weight above ground, g	0.93	-0.059
Underground weight, g	0.931	0.074

Table 9
KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy		0.743
Bartlett's Test of Sphericity	Approx. Chi-Square	14211.168
	df	15
	Sig.	0

is now spread more evenly over the components. The changes in the individual totals suggest that the rotated component matrix will be easier to interpret than the unrotated matrix.

The scree plot (Figure 1) helps us to determine the optimal number of components or confirms the choice of components. The eigenvalue of each component in the initial solution is plotted.

Generally, we want to extract the components on the steep slope. The components on the shallow slope contribute little to the solution. The last big drop occurs between the second and third components, so using the first two components is an easy choice.

The rotated component matrix (Table 8) helps us to determine what the components represent.

The first component is most highly correlated with Weights (above ground and underground). Height is a better representative, however, because it is less correlated with the other component. The second component is most highly correlated with Compost.

Table 9 shows two tests that indicate the suitability of our data for structure detection.

This table shows two tests that indicate the suitability of your data for structure detection. The Kaiser-Meyer-Olkin Measure of Sampling Adequacy is a statistic that indicates the proportion of variance in our variables that might be caused by underlying factors. High values (close to 1.0) generally indicate that a factor analysis may be useful with our data. If the value is less than 0.50, the results of the factor analysis probably won't be very useful. Bartlett's test of sphericity tests the hypothesis that our correlation matrix is an identity matrix, which would indicate that our variables are unrelated and therefore unsuitable for structure detection. Small values (less than 0.05) of the

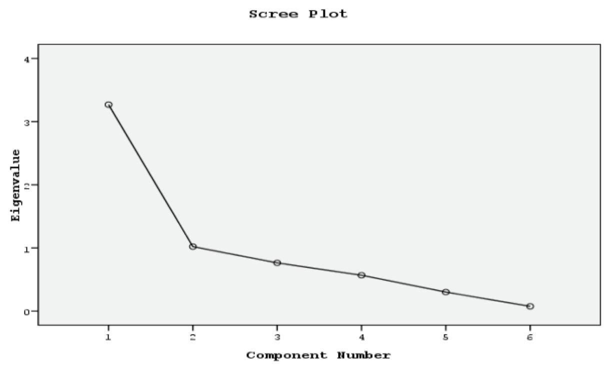

Fig. 1. Scree plot

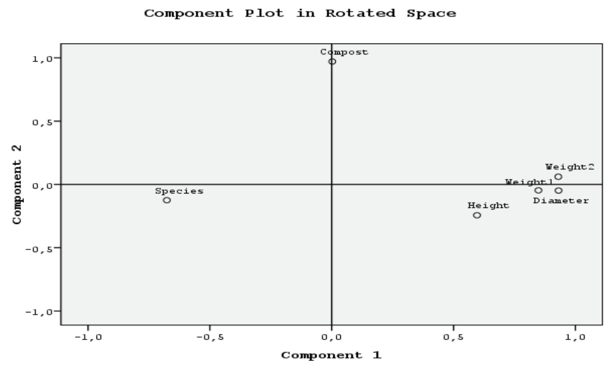

Fig. 2. Factor loadings plot

significance level indicate that a factor analysis may be useful with your data.

The factor loadings plot (Figure 2) is a visual representation of the rotated factor matrix. If the relationships in the matrix are complex, this plot may be easier to interpret.

Using a principal axis factors extraction, we have uncovered two latent factors that describe relationships between our variables. Components showed that there is no significant difference in plants growth (height, diameter, weights) between composts or tree species.

References

Afifi, A., S. May and V. Clark, 2004. Computer-aided multivariate analysis, 4th ed. New York: *Chapman & Hall.*

Burdett, A., D. Simpson and C. Thompson, 1983. Root development and plantation establishment. *Plant and Soil*, **71**: 103-110.

Butterfield, R., 1993. Tropical Timber Species Growth in the Atlantic Lowlands of Costa Rica and Wood Variation of Two Native Species. Ph.D. dissertation, North Carolina State Univ. Raleigh. 76 pp.

Butterfield, R., 1995. Promoting biodiversity: Advances in evaluating native species for reforestation. *Forest Ecology and Management,* **75:** 111-121.

Butterfield, R. and R. Fisher, 1994. Untapped potential. Native species for reforestation. *J. of Forestry*, **92** (6): 37-40.

Clarkson, D. and R. Jennrich, 1998. Quartic rotation criteria and algorithms. *Psychometrika*, **53:** 251–259.

Dixon, W., 1992. BMDP statistical software manual. Berkeley: *University of California Press.*

Fry, G. and B. Poole, 1980. Evaluation of planting stock quality several years after planting (a discussion). *New Zealand J. of Forestry Science*, **10**: 299-300.

Gnanadesikan, R., 1997. Methods for statistical data analysis of multivariate observations. 2nd ed. New York: *John Wiley & Sons.*

Harman, H., 1976. Modern factor analysis, 3rd ed. Chicago: *University of Chicago Press.*

Josiah, S. and N. Jones, 1992. Root Trainers in Seedling Production Systems for Tropical Forestry and Agroforestry. Land Resources Series No. 4. The World Bank Asia Technical Department. 40 pp.

Ladrach, W. 1992. Plantation establishment techniques in tropical America. *Tree Planter's Notes*, **43** (4): 125-132.

Liegel, L. and C. Venator, 1987. A Technical Guide for Forest Nursery Management in the Caribbean and Latin America. USDA, Puerto Rico. 156 pp.

Mardia, K., J. Kent and J. Bibby, 1979. Multivariate analysis. London: *Academic Press.*

Mason, E., 1985. Causes of juvenile instability of Pinus radiata in New Zealand. *New Zealand J. of Forestry Science*, **15** (3): 263-280.

Mulaik, S., 1972. The foundations of factor analysis.

New York: *McGraw-Hill.*

Rheingans, R., 1996. The Ecological, Economic and Political Effects of Tree Plantation Inter-cropping on Small Farm Forestry in the Atlantic Lowlands of Costa Rica. Ph.D. diss., Yale Univ. 195 pp.

Rose, R., W. Carlson and P. Morgan, 1990. The target seedling concept, pp. 1-8. In: Rose, R., Campbell, S. and Landis, T. (Eds.) Target Seedling Symposium: Proceedings, Combined Meeting of the Western Forest Nursery Associations. Aug. 13-17.

Rozeboom, W., 1982. The determinacy of common factors in large item domains. *Psychometrika,* **47**: 281–295.

Sharma, R., 1987. Some observations on coiling of roots in nursery raised plants. *J. of Tropical Forestry,* **3**: 207-212.

Shelhas, J., T. Jantzi, C. Kleppner, K. O'Connor and T. Thatcher, 1997. Meeting farmers' need through forest stewardship. *J. of Forestry,* (2): 33-38.

South, D., R. Mitchell, B. Zutter, J. Balneaves, B. Barber, D. Nelson and J. Zwolinski, 1993. Integration of nursery practices and vegetation management: Economic and biological potential for improving regeneration. *Can. J. For. Res*., **23**: 2038-2092.

Steiger, J., 1979. Factor indeterminacy in the 1930's and 1970's: some interesting parallels. *Psychometrika*, **44:** 157–167.

Received May, 3, 2010; accepted for printing December, 12, 2010.