

---

# 12 Spatial Stochastic Modelling of Cone Production from Stone Pine (*Pinus pinea* L.) Stands in the Spanish Northern Plateau

Nikos Nanos,<sup>1</sup> Rafael Calama,<sup>2</sup> Nieves Cañadas,<sup>3</sup>  
Carlos García<sup>4</sup> and Gregorio Montero<sup>2</sup>

---

## Abstract

The spatial structure of the mean cone production from even-aged stone pine stands of the Northern Plateau of Spain has been studied. Available data consisted of 123 five-tree plots, where cone crop was collected during a 5-year period (1996–2000). The experimental variogram for the mean cone crop has shown that cone production is a variable with a high percentage of spatially structured variance and a range of spatial correlation of approximately 2000 m.

A kriging map for the mean stand production (over the 5-year period) has been built. Additionally we used conditional simulation to study the spatial variability and to quantify the uncertainty of predictions. This method permits the construction of uncertainty and probability maps. Geostatistical prediction of cone production at a regional scale has been shown to be a useful tool for modelling a low-value natural resource such as stone pine cones, without needing additional covariate measurements.

## Introduction

Stone pine (*Pinus pinea* L.) is a typical Mediterranean species occupying in Spain an area of about 400,000 ha (more than 50% of the species' total area worldwide). Stone pine stands play an important role as soil protection elements in the semiarid zones of the Mediterranean basin (Ximénez de Embún, 1960) due to the ability of the species to occupy sandy areas, in both coastal and continental dune systems. Stone pine stands have been used and transformed by human activities from time immemorial. Among the products obtained from these stands, firewood, bark, wood and pinyon (the stone pine edible nut) are the most important. Stone pine pinyon is a highly prized food, widely used in pastry making. Cones are collected in natural or artificial stands, as well as in seed orchards and grafted plantations from genetically

<sup>1</sup> Unidad de Anatomía, Fisiología y Genética Forestal, ETS de Ingenieros de Montes, Spain

<sup>2</sup> Centro de Investigación Forestal CIFOR. Instituto Nacional de Investigación y Tecnología Agraria y Alimentaria, Spain

Correspondence to: rcalama@inia.es

<sup>3</sup> Dirección General del Medio Natural, Consejería de Medio Ambiente, Spain

<sup>4</sup> Servicio Territorial de Medio Ambiente. Junta Castilla y León

selected individuals. Despite the high cost of harvesting (cones are usually harvested manually), the actual increasing value of pinyon justifies its commercial utilization.

Forest management of these stands has been influenced by the main production objective (either wood or pinyon). The guidance of forest management to one or other product has been conditioned by the price and the demand for the corresponding product. The silvicultural practice had to be flexible enough in order to change the objective of the management as a response to market oscillations. Forest managers need useful, simple and flexible tools that can help them to decide on the principal forest product, depending on the market prices and the predicted stand productivity. Forest growth and yield models are nowadays among the most powerful available tools for forest management.

## Cone Production Modelling

Forest edible fruit production has seldom been modelled, due to its great complexity and variability. Among the numerous factors that control cone production we can mention genetic variability, site factors (climatic, edaphic, etc.), forest stand factors (stand density, basal area), single tree (breast height diameter, crown size, etc.) and other exogenous factors (pests, rodents, robbery, etc.).

Stone pine flowering and fruiting is a 3-year process. Interannual variability in the production is very pronounced. Despite the interest shown in stone pine stands as fruit producers, most of the existing research work on cone production is merely descriptive. These works generally associate a silvicultural intervention with the mean cone production for a given area. Prediction models have been carried out only in few studies, especially in Italy (Cappelli, 1958; Pozzera, 1959; Castellani, 1989) and Spain.

In Spain, the Department of Silviculture of CIFOR-INIA has been carrying out a research programme for the development of sustainable management schedules for stone pine stands. One of the main objectives of this research is the modelling of cone crop production. Within this project more than 400 experimental plots have been installed in the four most important stone pine regions of Spain. In these plots annual cone crop is collected and measured. As a result of this programme, several research projects, including cone production analysis, have been developed.

García Güemes (1999), in his silviculture simulation model for *P. pinea* in Valladolid, includes an analysis of the mean cone production per hectare. Cone production was found to be positively correlated with both mean square diameter and basal area and negatively correlated with stand density. Maximum cone production is found at ages ranging between 80 and 100 years. Montero *et al.* (2000) used data from stone pine stands of Western Andalucía to study the influence of age and density in cone production. In this case, significant differences in production were not found among different stand density classes. Cañadas (2000) developed a single-tree growth and yield model for even-aged stands of stone pine growing in the Central Range region. Individual tree cone crop was positively correlated to squared breast height diameter of the tree and dominant height of the stand, while stand density showed a negative correlation.

## Justification for a geostatistical prediction for cone production

Current prediction models for stone pine cone production are deterministic. Actual knowledge of the factors that control cone production is very limited. Generally,

single-tree models are considered to be better at predicting cone production (compared with whole-stand models). The inclusion in single-tree models of stand variables, such as stand density, as well as environmental factors – edaphic, climatic and topographical – improves the prediction ability of the model. However, single-tree models are not very useful in practice, since numerous covariates need to be measured.

In any case, the prediction ability of cone production models is generally very low. Attempts at predicting annual crop from these models are useless, since crop cycle variability is not included in the models. The low value of the stone pine cones does not justify a large monetary investment in predicting the production for a given area.

The use of geostatistical models to predict cone production shows several advantages over the existing regression models. Geostatistical estimations are generally much cheaper to obtain, since no additional covariates need to be measured. This property makes geostatistics an interesting tool for modelling a low-value natural resource such as cones. Kriging and simulation techniques give an unbiased estimation for the mean stand production as well as the spatial uncertainty of these estimates. Models for predicting the stand production at a regional scale could provide a helpful framework for regional forest management and planning, since these models allow the identification of the most productive areas. Finally, the extension of geostatistics to include time as an additional dimension could provide a useful approach for predicting the stand production at both the spatial and temporal scales (Stein *et al.*, 1998), giving the possibility for modelling inter-annual variability and crop cycles.

Several variables have been successfully predicted with geostatistics: site index (Hock *et al.*, 1993), total tree height (Samra *et al.*, 1989), timber volume (Holmgren and Thuresson, 1997) and stand density (Mandallaz, 2000). Yields of non-timber forest products, such as resin, have also been modelled with the use of geostatistical tools (Nanos *et al.*, 2001).

The aim of the present work is the spatial analysis of the average cone production per hectare (mean production of the years 1996–2000) as well as the construction, following kriging and simulation techniques, of maps for the average production and the associated spatial uncertainty.

## Materials and Methods

### Data

Data consisted of 123 plots established in even-aged stands and distributed all over the stone pine public forests of southern Valladolid (Fig. 12.1). Plots were circular, of variable radius, and included five trees. Among other variables, we measured the diameter at breast height, crown projection, total height and the height to the crown base. Cones were collected each autumn during 5 consecutive years (1996–2000), a period that is considered as a productive cycle (Ximénez de Embún, 1960).

Cones were classified into two groups according to amount of damage observed (as a consequence of insect attack: *Pissodes validirostris* or *Dioryctria mende-cella*). For each group, cones were counted and weighted. In this work, the studied variable was the mean weight of undamaged cones per hectare computed as the average value over a 5-year period by plot.

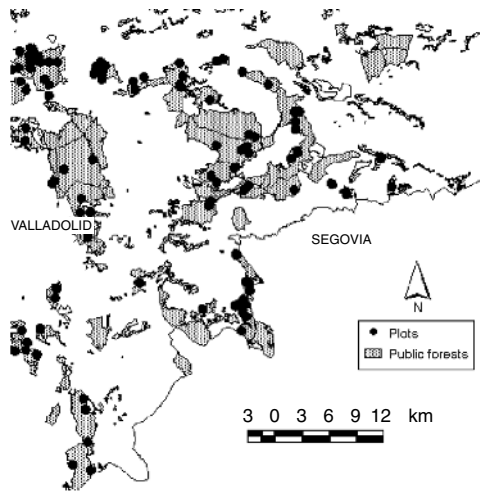


Fig. 12.1. Location of study area and plots.

### Geostatistical methods

A classical geostatistical analysis can be divided into two parts: *structural analysis* and *spatial prediction*. The aim of the first is to quantify the spatial correlation for a given attribute. When spatial correlation among observations is detected, one usually proceeds to the second step. Spatial prediction allows the study and estimation of the value of the attribute at unsampled locations within the study area.

#### *Structural analysis*

Let  $Z(x)$  denote a random function defined over the domain  $D$  of  $R^2$ , and sampled over a set of  $i = 1, 2, \dots, n$  points. In the geostatistical framework, the set of observations is considered as a single realization of the random function  $Z(x)$ . The spatial structure of a single realization is usually described by the semivariance  $\hat{\gamma}(h)$  for the lag  $h$ :

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{a=1}^{N(h)} [z(u_a) - z(u_a + h)]^2 \quad (1)$$

where  $N(h)$  is the number of pairs of data locations a vector  $h$  apart, while  $z(u_a)$  and  $z(u_a + h)$  are measurements at locations  $u_a$  and  $u_a + h$ , respectively. A plot of the semivariance versus the distance  $h$  is called a (experimental) variogram and it describes the spatial behaviour of the function.

Typically, the semivariance exhibits an ascending behaviour near the origin ( $h = 0$ ), while at larger separation distances it levels off at a maximum value called the *sill* of the variogram. The distance at which the sill is reached is called the *range* of the variogram, while the term *nugget* is used for the semivariance value at a distance  $h = 0$ .

#### *Detecting anisotropy*

Anisotropy directions were detected by variogram maps (Isaaks and Srivastava, 1989), which were constructed from 18 directional variograms, computed with a  $40^\circ$

angular tolerance and 4000 m bandwidth. A gridding algorithm was then applied to the resulting variogram values and the directions of maximum and minimum continuity were judged visually.

### *Variogram modelling*

Semivariogram modelling is needed for the posterior spatial prediction and for the conditional simulation. A basic semivariogram model should be a positive definite function, in order to guarantee the existence and singularity of the solution for the kriging system.

Anisotropy was corrected by rotating clockwise the coordinate system so as to identify the main axes of anisotropy and linearly transforming the rotated coordinates according to the anisotropic variogram model (Isaaks and Srivastava, 1989; Goovaerts, 1997). Kriging was finally performed in the transformed coordinate system using a global search neighbourhood.

### *Spatial prediction*

The existence of a model of spatial dependence allows us to tackle the problem of estimating attribute values at unsampled locations. To do this, we used the data obtained from the sampled plots, and the semivariogram model constructed previously. Prediction was made following the kriging method, so called in recognition of the pioneering work of D. Krige.

Kriging estimators are among the BLUEs (best linear unbiased estimators), and are obtained in a similar way to linear regression estimators. Their two main properties are:

- prediction from these estimators is unbiased; and
- the error variance is minimized (the precision is going to be maximized).

Ordinary kriging estimators for the mean value of the attribute  $z^*(\mathbf{u})$  at an unsampled location  $\mathbf{u}$  is written as a linear combination of the values of the same attribute in the  $n$  sampled points  $\{z(\mathbf{u}_a), a=1,2,\dots,n\}$ :

$$z^*(\mathbf{u}) = \sum_{a=1}^{n(\mathbf{u})} \lambda_a(\mathbf{u}) z(\mathbf{u}_a) \quad (2)$$

where  $\lambda_a(\mathbf{u})$  is the assigned weight for data  $z(\mathbf{u}_a)$ .

From this formulation, to find the best estimator we have to determine the optimum value for the weights  $\lambda_a(\mathbf{u})$ . In order to get non-bias conditions for the estimator  $z^*(\mathbf{u})$ , we have to impose the constraint that sum of the weights  $\lambda_a(\mathbf{u})$  should be equal to 1 (universality condition):

$$\sum_{a=1}^{n(\mathbf{u})} \lambda_a(\mathbf{u}) = 1 \quad (3)$$

In order to verify the property of minimum error variance, prediction error is defined as  $R(\mathbf{u}) = Z^*(\mathbf{u}) - Z(\mathbf{u})$ , and its variance equals zero:

$$\text{Var}\{R(\mathbf{u})\} \rightarrow 0 \quad (4)$$

Variance error defined in Equation 4 is only related to spatial covariance, which can be estimated from the variogram. In order to minimize Equation 4 we use the Lagrange technique, under the universality condition imposed by Equation 3. We get the following equation system, whose solution gives us the optimum weights:

$$\begin{aligned} \sum_{b=1}^{n(u)} \lambda_a \gamma(u_a - u_b) + \mu &= \gamma(u_a - u) \quad a = 1, \dots, n(u) \\ \sum_{a=1}^{n(u)} \lambda_a(u) &= 1 \end{aligned} \quad (5)$$

where  $\gamma(u_a - u_b)$  is the semivariance among points  $u_a$  and  $u_b$ , where the attribute has been measured;  $\gamma(u_a - u)$  is the semivariance among  $u_a$  and  $u$ , unsampled location, and  $\mu$  is the value for the Lagrange multiplier.

### *Conditional simulation*

Every prediction has an associated estimation error that should be quantified. In the geostatistical framework the main tool for quantifying the kriging estimation errors has been the kriging variance. However, it has been shown that this measure of spatial uncertainty is not reliable, since its value depends only on the spatial arrangement of the sampled plots and not on the actual values of the studied attribute.

Simulation methods are preferred, then, as primary tools for quantifying the spatial uncertainty of the estimation. Several methods have been proposed and used in practice. In this chapter we have used sequential Gaussian simulation (SGS). The steps that have been followed during the simulation can be summarized as follows.

1. The original data should be transformed to normal-score data. This transformation is usually achieved through the normal-score transformation.
2. A grid of the desired resolution is defined over the studied area.
3. Each node of the grid is visited (only once) in a random sequence. At each grid-node the parameters of the random variable (mean and standard deviation) are determined by simple kriging.
4. Draw randomly a simulated datum from the distribution specified in the previous step and add this datum to the data set.
5. Proceed to the next grid-node (chosen randomly) and repeat this process until all the nodes are simulated.
6. Repeat steps 3 to 5 until the desired number of simulations is achieved. In this study we realized 50 simulations.
7. The final step of the simulation is the back-transformation of the normal data to their original values, as described in Goovaerts (1997).

### *Simulation post-processing*

There are several ways to summarize the results of the conditional simulations, the most important being the measurement of uncertainty. In the present study we used the standard deviation of the simulated data as an overall measurement of uncertainty. Other approximations include the interquartile range or the entropy of the simulated distribution (Goovaerts, 1997).

We also constructed a map showing the probability for the mean cone production to be larger than 200 kg/ha. This production can be considered as the minimum stand production for cone collection to be profitable.

## Results and Discussion

### Structural analysis

In Fig. 12.2 we present the histogram of the original values over the 123 sampled plots. The directional experimental variograms of the normal-transformed values are presented in Fig. 12.3. The variogram shows the existence of spatial correlation for the mean cone production per hectare, up to a distance of about 2000 m (range of spatial correlation).

The practical absence of nugget variance in the experimental variogram is very important for the present study. The absence of nugget variance implies that cone production is very continuous at a larger scale of variation and that factors controlling the production have a large range of spatial correlation.

Detection for anisotropic spatial continuity in the production showed that the main axes of anisotropy laid on the  $0^\circ$  and  $90^\circ$  direction (the former is the direction of maximum spatial continuity). Since the directional experimental variogram exhibits significant differences, it is preferable to model spatial variance using directional variogram models. These models are the following:

$$\gamma(h_0) = 1.028 \times \text{SPHER}_{2000}(h_0) + 0.756 \times \text{EXP}_{7000}(h_0) \quad (6)$$

$$\gamma(h_{90}) = 1.028 \times \text{SPHER}_{1000}(h_{90}) \quad (7)$$

where SPHER and EXP are the spherical and the exponential variogram models (the subscript indicates the range of the variogram model), while  $h_0$  and  $h_{90}$  are the distances following  $0^\circ$  and  $90^\circ$  direction.

### Prediction and simulation of the mean cone production

A kriged map for the mean cone production can be seen in Fig. 12.4. There are four nuclei with high production surrounded by large areas where the production seems to be very low. The kriged map for the mean cone production is in general accordance with the practical experience of forest rangers. Besides being a useful tool for forest management, the kriged map represents a certain scale of variation of the phenomenon under study and its use should be restricted to that specific scale of variation. The cone production from stone pine seems not to be influenced

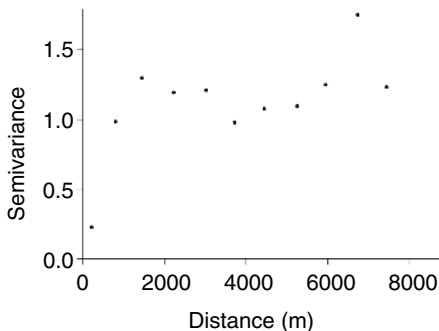


Fig. 12.2. Experimental variogram of mean cone production.

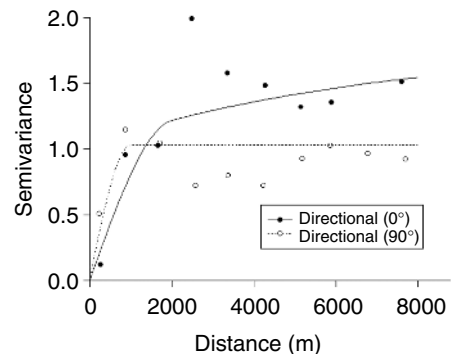
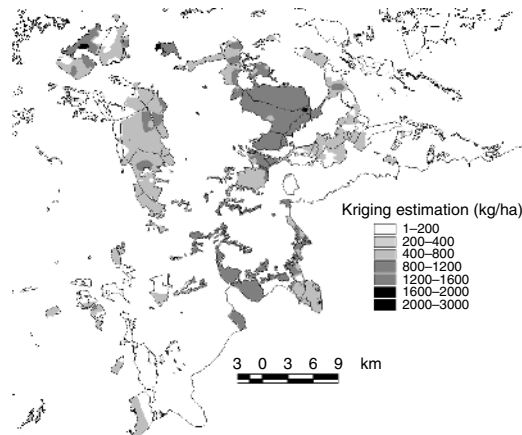


Fig. 12.3. Directional experimental variograms and models.



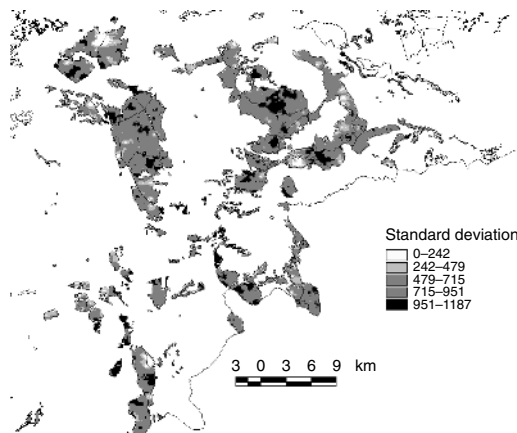
**Fig. 12.4.** Kriged map of mean cone production.

by factors with a small range of spatial variation (less than 1000–1500 m) since the nugget effect of the experimental variogram has been shown to be null.

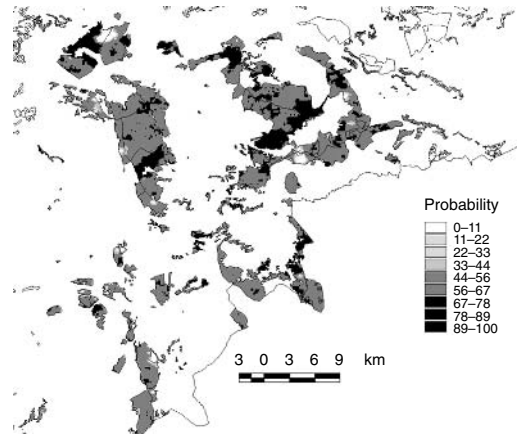
In Figs 12.5 and 12.6 we present the results of the simulations. The standard deviation of the simulated data, shown in Fig. 12.5, may be used as an indication of uncertainty, while the probability map, Fig. 12.6, should be used in order to guide the future campaigns for collection of cones. This map shows the probability of a stand producing more than 200 kg/ha of cones.

## Conclusions

The mean cone production from stone pine stands has been shown to be a variable exhibiting spatial correlation up to a distance of 2000 m. Furthermore, it has been shown that the production is very continuous at a large scale of variation (absence of nugget variance). This result may be used for the optimization of the sampling design in future sampling campaigns. On the other hand, this result has some direct theoretical implications: cone production is not influenced by micro-environmental conditions. Intuitively,



**Fig. 12.5.** Map of standard deviation for mean cone production.



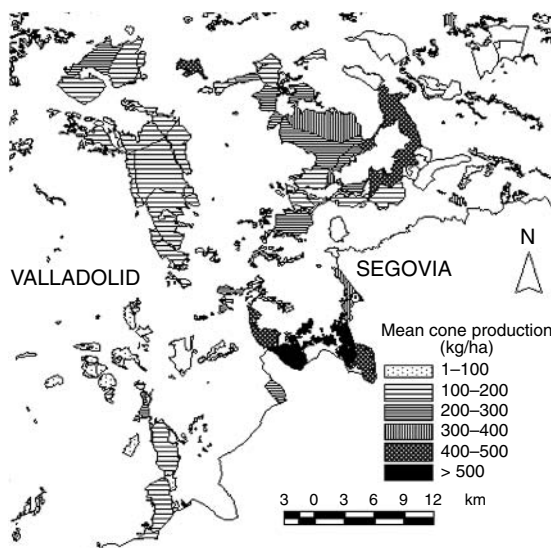
**Fig. 12.6.** Probability map for cone production >200 kg/ha.

the mean stand production is influenced by small-scale climatic factors such as precipitation and temperature.

Regarding the spatial prediction of the mean stand production, we believe that, at this stage of the modelling process, the geostatistical approach is satisfactory, since other predictive models require expensive measurements to be realized (single-tree models) or they have a very low predictive power (whole-stand models).

We compared the results obtained with those from the work by Gordo Alonso *et al.* (2000). The authors used data from post-crop visual estimation made by forest rangers every year in all the public forestlands. Data used are the mean value from a series of 34 years. Despite the roughness of the data, the work is a good approach to a geographical classification of cone productivity for stone pine stands in Valladolid (Fig. 12.7).

In comparing both studies, we find the main differences in the total amount of the cone crop per hectare, ranging in Gordo's work from 0 to 600 kg/ha and in our work



**Fig. 12.7.** Mean annual cone production for public forestlands (from Gordo Alonso *et al.*, 2000).

from 0 to 2500 kg. The reason for this difference is that kriging estimation is more accurate, detecting differences at a small spatial scale, while mean production has been calculated for a whole forest. Apart from this, in our work, the complete crop was collected in fully stocked stands, while in real harvesting work, there are many areas in the forestlands where cones are not collected (young stand areas, small productive areas, isolated trees). Despite differences in total crop amount, we find similarities in locating the most productive areas in the region.

The probability map (Fig. 12.6), showing the probability of a crop larger than 200 kg/ha, could be a very useful tool for designing cone collection campaigns.

Stone pine area in Valladolid province is about 30,000 ha. Valladolid is one of the provinces in Spain where cone collection and the pinyon market have been a traditional activity. Since ancient times, cone production has been included as a main objective in forestland management (Romero y Gilsanz, 1884; Olazabal, 1917). Nowadays, the pinyon and transformed products industry invoices an annual amount of 12 million euros, generating nearly 2000 jobs, which gives an idea of the importance of this industry in the province (Gordo Alonso, 1998). On account of this, any attempt to improve the ability for predicting cone production should be considered positive.

However, we believe that the geostatistical model presented in this study should be considered as a temporary prediction model, until other more sophisticated modelling tools, including site, stand and single-tree variables, as well as temporal variability, become available.

## Acknowledgements

The authors wish to thank Javier Gordo Alonso, from the Servicio Territorial de Medio Ambiente, Valladolid, for technical support and data supply. We are also grateful to Angel Bachiller, for his work in installing the plots, and to Sven Mutke, for his valuable comments on the manuscript. The research was partially supported by a grant to R. Calama from the Consejería de Educación, Comunidad de Madrid, in the context of the INIA project SC-99-017.

## References

- Cañadas, M.N. (2000) *Pinus pinea* L. en el Sistema Central (valles del Tiétar y del Alberche): desarrollo de un modelo de crecimiento y producción de piña. PhD thesis, ETSI Montes, Universidad Politécnica de Madrid, Spain.
- Cappelli, M. (1958) Note preliminari sulla produzione individuale di stroboli in *Pinus pinea* L. *L'Italia Forestale e Montana* 13(5), 181–203.
- Castellani, C. (1989) La produzione legnosa e del frutto e la durata economico delle pinete coetanee di pino domestico (*Pinus pinea* L.) in un complesso assestato a prevalente funzione produttiva in Italia. *Annali ISAF* 12, 161–221.
- García Güemes, C. (1999) Modelo de simulación selvícola para *Pinus pinea* L. en la provincia de Valladolid. PhD thesis, ETSI Montes, Universidad Politécnica de Madrid, Spain.
- Goovaerts, P. (1997) *Geostatistics for Natural Resources Evaluation*. Oxford University Press, New York.
- Gordo Alonso, F.J. (1998) Programa de Mejora Genética de *Pinus pinea* L. en Castilla y León. *Montes* 52, 71–84.
- Gordo Alonso, F.J., Mutke, S. and Gil, L. (2000) La producción de piña de *Pinus pinea* L. en los montes públicos de Valladolid. In: de Castilla y León, J. (ed.) *Primer Simposio sobre el pino piñonero (Pinus pinea L.)*, Valladolid, pp. 269–277.

- Hock, B.K., Payn, T.W. and Shirley, J.W. (1993) Using a geographical information system and geostatistics to estimate site index of *Pinus radiata* for Kaingaroa Forest, New Zealand. *New Zealand Journal of Forestry Science* 23, 264–277.
- Holmgren, P. and Thuresson, T. (1997) Applying objectively estimated and spatially continuous forest parameters in tactical planning to obtain dynamic treatment units. *Forest Science* 43, 317–326.
- Isaaks, E.H. and Srivastava, R.M. (1989). *An Introduction to Applied Geostatistics*. Oxford University Press, New York.
- Mandallaz, D. (2000) Estimation of the spatial covariance in universal kriging: application to forest inventory. *Environmental and Ecological Statistics* 7, 263–284.
- Montero, G., Candela, J.A., Ruiz-Peinado, R., Gutiérrez, M.F., Pavón, J., Bachiller, A., Ortega, C. and Cañellas, I. (2000) Density influence in cone and wood production in *Pinus pinea* L. forests in the south of Huelva province. In: CIFOR-INIA (ed.) *Mediterranean Silviculture with Emphasis in Quercus suber, Pinus pinea and Eucalyptus sp.* IUFRO Working Group 1.05.14, Seville, pp. 129–142.
- Nanos, N., Tadesse, W., Montero, G., Gil, L. and Alía, R. (2001) Spatial stochastic modelling of resin yield from pine stands. *Canadian Journal of Forest Research* 31, 1140–1147.
- Olazabal, D. (1917) Datos sobre el problema silvopastoral en España. *Revista de Montes* 41.
- Pozzera, G. (1959) Raporti fra produzione di stroboli in *Pinus pinea* L. de andamento stagionale. *L'Italia Forestale e Montana* 14(5), 196–206.
- Romero y Gilsanz, F. (1884) De los pinares en las arenas sueltas de la provincia de Valladolid. *Revista de Montes* 9, 220–228.
- Samra, J.S., Gill, H.S. and Bhatia V.K. (1989) Spatial stochastic modelling of growth and forest resource evaluation. *Forest Science* 35, 663–676.
- Stein, A., Van Groenigen, J.W., Jeger, M.J. and Hoosbeek, M.R. (1998) Space-time statistics for environmental and agricultural related phenomena. *Environmental and Ecological Statistics* 5, 155–172.
- Ximénez de Embún, J. (1960) *Arenas movedizas y su fijación (Dunas continentales)*. Ministerio de Agricultura, Madrid, Spain.