Estimation of productive areas of stone pine cone in Portugal with geostatistical tools

Gonçalves A., Ferreira J., Godinho-Ferreira P., Rodrigues A.

in

Carrasquinho I. (ed.), Correia A.C. (ed.), Mutke S. (ed.).
Mediterranean pine nuts from forests and plantations

Zaragoza : CIHEAM
Options Méditerranéennes : Série A. Séminaires Méditerranéens; n. 122

2017
pages 55-61

Gonçalves A., Ferreira J., Godinho-Ferreira P., Rodrigues A. **Estimation of productive areas of stone pine cone in Portugal with geostatistical tools.** In : Carrasquinho I. (ed.), Correia A.C. (ed.), Mutke S. (ed.). *Mediterranean pine nuts from forests and plantations.* Zaragoza : CIHEAM, 2017. p. 55-61 (Options Méditerranéennes : Série A. Séminaires Méditerranéens; n. 122)

Estimation of productive areas of stone pine cone in Portugal with geostatistical tools

A. Gonçalves[1], J. Ferreira[2], P. Godinho-Ferreira[3] and A. Rodrigues[3,*]

[1]Departamento de Geografia e Desenvolvimento Regional / Faculdade de Ciências Sociais e Humanas,
Universidade Nova de Lisboa, Avenida de Berna, 26-C / 1069-061 Lisboa (Portugal)
[2]CICS.NOVA, Centro Interdisciplinar de Ciências Sociais, Universidade Nova de Lisboa,
Avenida de Berna, 26-C / 1069-061 Lisboa (Portugal)
[3]INIAV, Instituto Nacional de Investigação Agrária e Veterinária,
I. P. Quinta do Marquês, Av. República, 2780-152 Oeiras (Portugal)
*e-mail: abel.rodrigues@iniav.pt

Abstract. This work follows previous field studies aiming to define five different stone pine (*Pinus pinea* L.) tree developmental stages, of which three corresponded to cone production stages. Based on collected data of cone production of stone pine stands in 40 georeferenced plots and 330 trees in three production campaigns (2004/05, 2005/06 e 2006/07), geostatistical methods of semivariogram and kriging were used to assess the areas of cone productive classes. The study area is located in the region of Setúbal Peninsula in South-Western Portugal, the denominated Provenance Region V, which is the area with the highest cone production in the country. A simple kriging model with a detrended exponential semivariogram was selected and the results of cone productive classes 1 (299-658 kg.ha^{-1}), 2 (122-298 kg.ha^{-1}), and 3 (lower than 112 kg.ha^{-1}), have a potential extent of 345,056 ha, 189,571 ha, and 148,275 ha, respectively. Considering the actual stone pine area the estimated areas of the three productive classes were 23,196 ha, 29,329 ha and 3559 ha for classes 1, 2 and 3, respectively. Further application of geostatistical methodology can be used to evaluate the potential of expansion of stone pine in this region on areas with other land cover types considering cone and/or kernel productive classes.

Keywords. Kriging – Semivariogram – Stone pine – Cone Productive areas.

Résumé. *Ce travail suit des études visant à définir cinq stades de développement d'arbres de pin parasol (*Pinus pinea *L.), dont trois correspondaient à étapes de la production de cônes. Les données de terrain ont été collectées en 40 placettes géoréférencées avec 330 arbres, pendant trois campagnes de production (2004/05, 2005/06 e 2006/07). L'évaluation des zones de classes productives de cône a été exécuté avec les méthodes géostatistiques de semivariogramme et krigeage. Les placettes ont été installées en peuplements de pin parasol dans la région de la péninsule de Setúbal au sudouest du Portugal, la Région de Provenance V, qui est la productrice plus importante de cônes du pays. Un modèle de krigeage simple, avec une semivariogramme exponentielle et retrait de tendance a été sélectionné, et les résultats des classes productives de cône : 1 (299-658 kg.ha^{-1}), 2 (122-298 kg.ha^{-1}), et 3 (inférieure à 112 kg.ha^{-1}), ont donné des zones d'occupation potentiel de 345 056 ha, 189 571 ha, et 148 275 ha, respectivement. Compte tenu de la surface effective de pin parasol, les zones estimées des trois classes productives sont 23 196 ha, 29 329 ha et 3559 ha pour les classes 1, 2 et 3 respectivement. L'application de la méthodologie géostatistique peut être également utilisée pour évaluer le potentiel d'expansion du pin parasol dans des zones avec d'autres types d'occupation du sol dans cette région, envisageant les classes productives de cône et /ou de pignon.*

Mots-clés. *Krigeage – Semivariogramme – Pin parasol – Zones productives de cône.*

I – Introduction

The area of stone pine (*Pinus pinea* L.) in Portugal has increased 46% since 1995, corresponding currently to about 6% (about 176,000 ha) of the total forest area (ICNF, 2015) and ranking as the fifth species in occupied area. The main source of income for stone pine producers is kernel production for food industry, but environmental benefits in terms of sand dune fixation, soil protection and improvement of degraded ecosystems should also be envisaged. The cone production in Portugal is ca. 65 millions of cones corresponding to about 600 to 700 tons of kernel, mostly aimed at exportation.

This work follows a study that defined five tree developmental stages, of which three corresponded to cone production stages (Carrasquinho *et al*., 2010), in stone pine stands installed in the Provenance Region V located in South-western Portugal (Cardoso and Lobo, 2001). A complementary work of evaluation of the areas of these productive classes is thereby required. Modeling cone production as a function of stone pine tree biometric variables showed the relationships among these variables (e.g., Calama and Montero, 2005; Freire, 2009; Rodrigues *et al*., 2014) and the need of evaluation of their spatial variability. Geostatistical tools, such as the ones implemented in Geographical Information Systems (e.g., Geostatistical Analyst tool ArcGis 10.2.2). Geographical Information Systems (GIS) for quantifying spatial variation of clustered data variables, can be used in forest applications. Within this context, this work intended to quantify areas for three cone productive classes, in Provenance Region V, through a geostatistical interpolation of cone production data obtained in field plots. That information should allow for a detailed assessment of the productive potential of the regions abridged by the study.

II – Materials and methods

1. Data collection in the plots

Cone production was obtained from 40 circular georreferenced plots with 330 trees, installed in 2004 and 2005, in eight counties on Provenance Region V, with data production collected in the 2004/05, 2005/06 and 2006/07 campaigns. The data used in this study were the cone production per plot, obtained from the sum of the productions available for each tree in the three campaigns, as it was not possible to obtain production data for the three campaigns from some trees. The productive classes were defined with average productions per ha in the ranges of: lower than 112 kg.ha^{-1} for class 1, 122-298 kg.ha^{-1} for class 2, and 299-658 kg.ha^{-1} for class 3, respectively. The eight counties where plots were located in Chamusca, Coruche, Vendas Novas, Montemor-o-Novo, Setúbal, Alcácer do Sal, Grândola and Santiago do Cacém, respectively (Fig. 1). These counties corresponded to a total area of 682,902 ha. To obtain digital information concerning the area of distribution of stone pine in Portugal, Land Use and Land Cover Map of Portugal mainland 2007 (COS 2007) (DGT, 2007) was used. This digital chart is based on the interpretation of orthorectified aerial photos and also in a multi series satellite imagery enabling a better identification of the vegetation phenology and of the soil occupation. The minimal cartographic unit is 1 ha with a minimum 20 m distance between lines; consequently the grain of analysis, a pixel of 20m width, was used.

2. Data treatment

The productive area estimations were performed using the geostatistic extension implemented in the Geostatistical Tool in the ArcGis package 10.2.2 (Esri Inc.) Geostatistics are based on the concepts of kriging and semivariogam. Kriging is the ultimate spatial linear interpolating model of a spatial random variable $Z(s)$: s D R^2 of the general form (Schabenberger and Pierce, 2002; Gonçalves, 2015):

$$Z(s) = \mu(s) + \delta(s) \qquad \text{(Equation 1)}$$

where $\mu(s)$ is the mean of the random field and is its variance. The variance is evaluated by the semivariogram $\gamma(h)$, h being the Euclidean space between two points, which in its general form is given by:

$$\gamma(h) = ½ \, Var\,[\,Z(s) – Z\,(s+h)] \qquad \text{(Equation 2)}$$

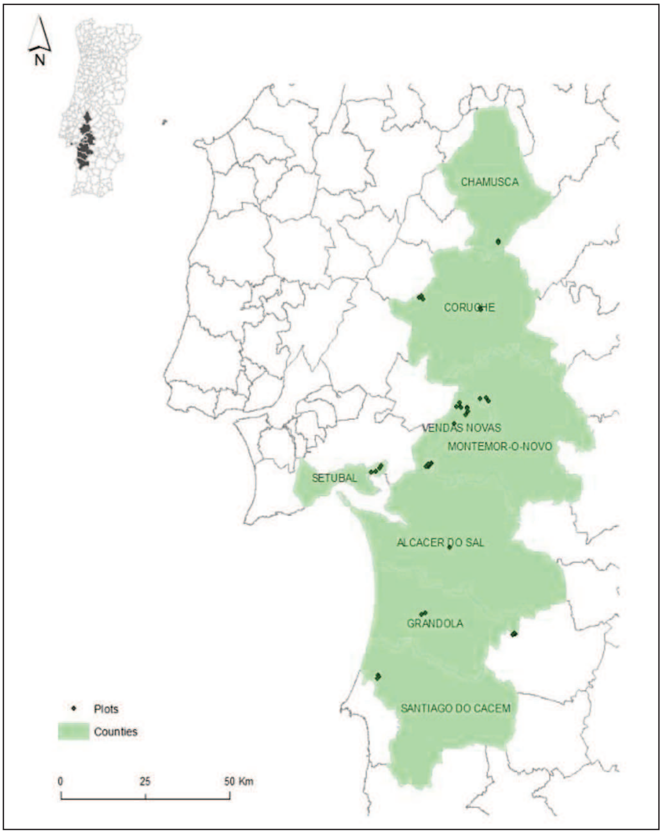

Fig. 1. Counties of stone pine locations.

The semivariogram is thereby representative of the spatial correlation of values of a variable $Z(s)$ in the points s and (s+h) and can be modeled by several kinds of equations such as spherical, exponential or Gaussian. The common applications of the semivariogram rest on assumptions of intrinsic or second order stationarity and isotropy. A theoretical configuration of the semivariogram is shown in Fig. 2. It can be noticed that the semivariogram converges monotonically to an asymptote (sill) at a given lag distance (pratical range) representative of the distance from the first point.

In real conditions a discontinuity at the origin, the so-called nugget effect, can exist, due for example to measurement error and the convergence occurs on a partial sill (Fig. 3).

In this study we considered the geostatistical interpolation by simple, ordinary and universal krigings. General equation 1 allows to define simple kriging if the mean of equation $\mu(s)$ is known, ordinary kriging if the mean is unknown, or universal kriging if the mean is given by an expression like:

$$\mu(s) = x'(s)\beta \qquad \text{(Equation 3)}$$

with β unknown. The kriging predictors of $Z(s)$ with a general form $p(Z\ s_o)$ are estimated as a general sum:

$$p(Z; s_o) = \lambda' Z(s) \qquad \text{(Equation 4)}$$

where λ is the vector of kriging weights, representative of the contribution of each of the measured points to the estimated spatial variation. The predictors are calculated in order to minimize parameters error parameters such as the mean prediction error:

$$\frac{Z(s_o) - \sum_{i=1}^{n} \lambda(s_t) Z(s_t)}{n} \qquad \text{(Equation 5)}$$

where n is the number of plots. Other statistics for kriging error evaluation are the mean prediction error, the average standard error, the root mean square error and the standardized root mean square (Johnson *et al.*, 2001). A good spatial model requires that the mean prediction error values should be close to zero, root mean square values as lower as possible, average standard errors close to and root mean square and standardized errors close to 1. Detrending and kernel treatment of semivariograms were also applied in this study. Indeed experimental semivariograms can follow a curve pattern distinct from the Fig. 2 and Fig. 3 above, for example as smooth concave curves that approach the origin with decreasing gradients or that increase sharply after reaching the sill. These patterns are indicative of a global trend or rift which is imposed on the short range variation in the spatial variable. The global trend is an overriding process that affects the measurements on a deterministic way. The superficial trend can be represented by an equation (e.g., a polynomial) removed to pay attention to the stochastic spatial structure submitted to kriging processes and replaced before final predictions are made. Experimental semivariograms can also present some high frequency white noise fluctuations, which can be smoothed through convolution with the so-called kernel functions (e.g., exponential, Gaussian or spherical).

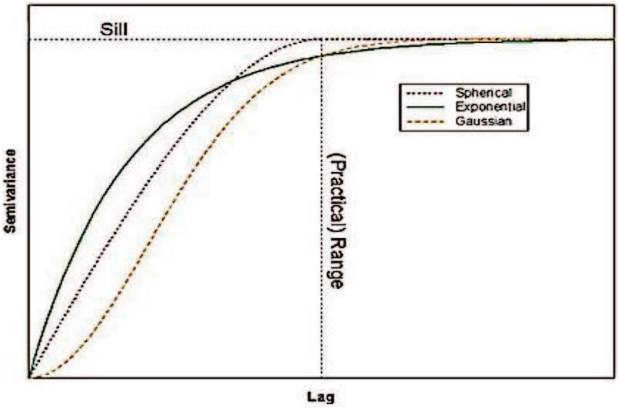

Fig. 2. Theoretical configuration of a semivariogram (adapted from Johnston, 2001).

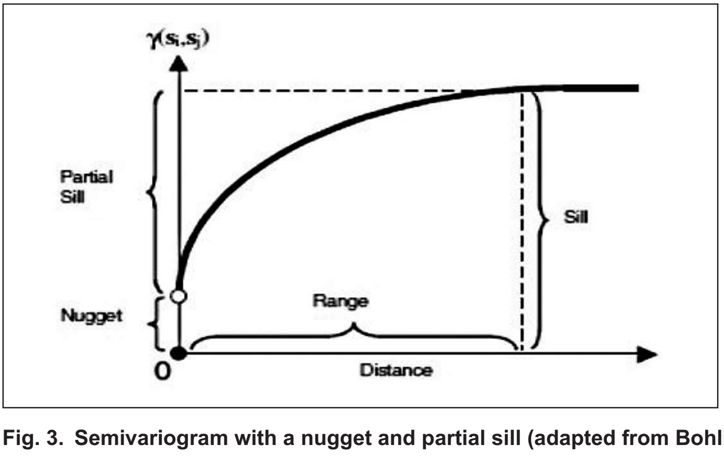

Fig. 3. Semivariogram with a nugget and partial sill (adapted from Bohlig, 2001).

Following the mentioned geostatistical principles, 24 kinds of kriging were applied to evaluated areas for cone production, differing about the type of kriging (simple, ordinary or universal) semivariogram (spherical, exponential or Gaussian), kernel and trend removal.

III – Results and discussion

The spatial model chosen, which optimized the four error criteria aforementioned, was a simple kriging with an exponential semivariogram, a lag size of 11.86m, a second order polynomial trend function and an exponential kernel function (Fig. 4). The exponential function provided a convenient

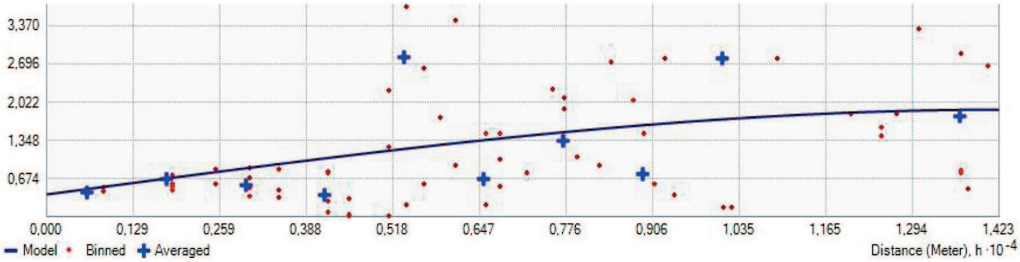

Fig. 4. Semivariogram chosen for the modeling of spatial variability of cone production.

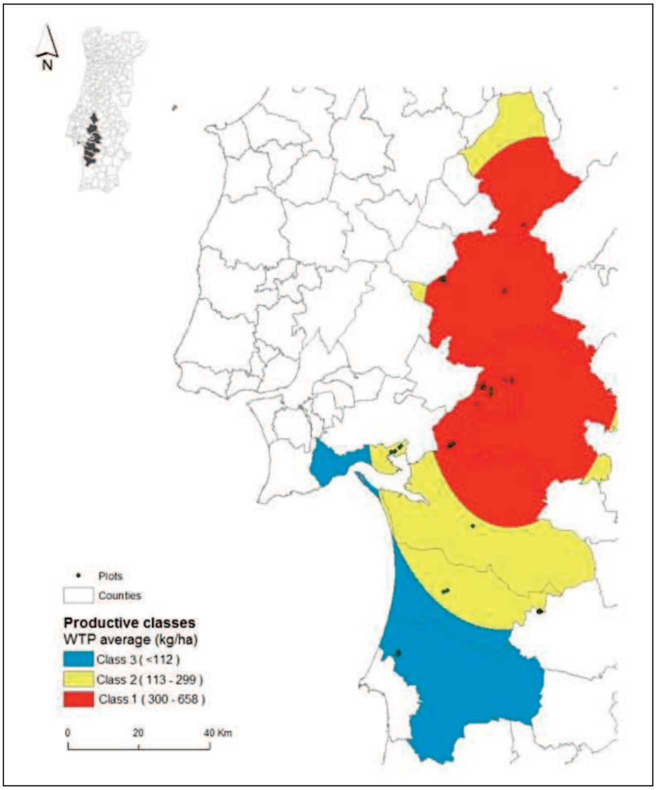

Fig. 5. Kriging results for the three cone productive classes.

smoothing of point autocorrelation. The values of mean prediction error root mean square error, average standard error and standardized root mean square were -0.61 kg, 66.57 kg, 2.3 kg and 0.94 respectively. In the total 682,902 ha counties' area, the results from kriging of cone productive classes 1 (299-658 kg.ha^{-1}), 2 (122-298 kg.ha^{-1}), and 3 (lower than 112 kg.ha^{-1}), gave potential occupation areas of 345,056 ha, 189,571 ha, and 148,275 ha, respectively (Fig. 5).

The magnitude of these areas is indicative of the potential of this region for stone pine cultivation. After intercepting these kriging results with the actual pine stone areas in the 8 counties, using COS 2007, the areas for productive classes 1, 2, and 3 were 23,196, 29,329 and 3558 ha, respectively (Fig. 6). The red (dark) area concerning to higher production (1), corresponded to transect of Chamusca, Coruche, Vendas Novas Novas, Montemor o Novo and Alcácer do Sal. The blue (grey) area concerning the lower productive class (3) was located mainly in Santiago do Cacém County. The yellow (brighyt) area concerning productive class 2 was located in Grândola, Northern Alcácer and Southern Chamusca. The digital chart COS 2007 was therefore a valuable tool to overlap and compare the theoretical kriging results with the actual soil occupation.

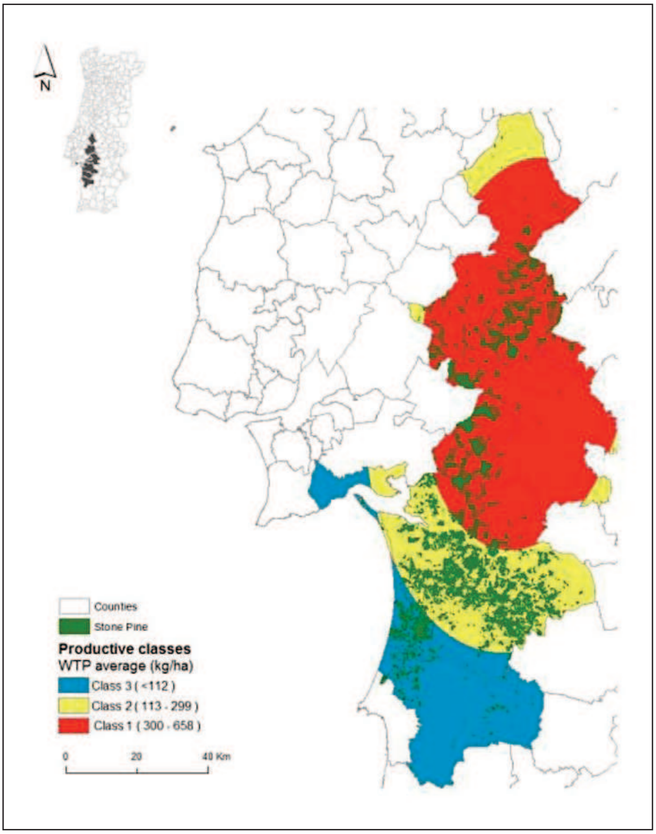

Fig. 6. **Actual predicted areas of distribution the three cone productive classes.**

IV – Conclusions

This study allowed a prediction of the areas of three productive classes of stone pine in the Provenance Region V, respectively 23,196, 29,329 and 3558 ha, for cone productive classes 1, 2 and 3, located in a continuous longitudinal transect from Chamusca to Alcácer do Sal. The results also showed the potential of geostatistical tools, integrated with GIS, for evaluating of forest biometrical and environmental variables, such as soil and topography. In particular, for stone pine, two interesting field studies that can be envisaged with these tools are the variability of kernel production with biometrical, soil and climate variables and also the potential expansion of stone pine to areas actually occupied with maritime pine or grassland. With these studies an improved validation of the geostatistical predictions should also be achieved.

References

Bohling G., 2005. *Introduction to Geostatistics and Variogram Analysis*. http://people.ku.edu/~gbohling/cpe940/Variograms.pdf (accessed in July, 2015).

Calama R. and Montero G., 2005. Cone and seed production from stone pine (*Pinus pinea* L.) stands in Central Range (Spain). *Eur. J. Forest Res.,* 126(1), pp. 23-35.

Cardoso M. and Lobo P., 2001. Delimitação de pisos bioclimáticos e regiões de proveniência de Pinheiro manso em Portugal, usando sistemas de informação geográfica. *Silva Lusitana*, 9(1), pp. 93-108.

Carrasquinho I., Freire J., Rodrigues A. and Tomé M., 2010. Selection of *Pinus pinea* L. plus tree candidates for cone production. *Ann. For. Sci.*, 67, p. 814. DOI:10.1051/forest /2010050.

DGT – Direcção Geral doTerritório, 2007. *COS 2007 – Carta de Ocupação do Solo 2007*.

Freire J., 2009. *Modelação do crescimento e da produção de pinha no pinheiro manso*. PhD Thesis, Instituto Superior de Agronomia, Universidade Técnica de Lisboa.

Gonçalves A.P.S., 2015. *Metodologias de Sistemas de Informação Geográfica em sistemas de produção florestal : uma aplicação geoestatística*. MSc Thesis, Faculdade de Ciências Sociais e Humanas, Universidade Nova de Lisboa.

ICNF, 2015. *Inventário Florestal Nacional (Portugal Continental)*. Lisboa. http://www.icnf.pt/portal/florestas/ifn/ifn6.

Johnston K., Ver Hoef, J.M, Krivoruchko K. and Lucas N., 2001. *Using ArcGis Geostatistical Analyst*. Redlands, USA, 300 p. *http://dusk2.geo.orst.edu/gis/geostat_analyst.pdf.*

Rodrigues A., Silva G., Casquilho M., Freire J., and Tomé M., 2014. Linear Mixed Modelling of Cone Production for Stone Pine in Portugal. *Silva Lusitana*, 22(1), pp. 1-27.

Schabenberger O. and Pierce J.F., 2002. *Contemporary statistical models for the plant and soil science*s. Taylor and Francis, Boca Baton, USA, 738 p.