# Using Small Area Models to Estimate the Total Area Occupied by Olive Trees

Militino, A.F., Ugarte, M.D. and Goicoa, T.
Departamento de Estadística e Investigación Operativa
Universidad Pública de Navarra
Campus de Arrosadía, 31006 Pamplona, Spain

## Abstract

This work aims at estimating the total area occupied by olive trees in a region called Comarca IV, located in a central region of Navarra, Spain, using as auxiliary information, classified data provided by satellite images. Traditionally, small area linear mixed models have been used for similar purposes using regular quadrats (also called segments) as sampling units, and assuming that these are fully included in the study domain. When it does not happen, the sampling units are of different size, and there exists an extra variability that can be very different within areas. In this case it is advisable to include weights into the model. In this work, we propose a weighted unit level linear mixed model where both, the variance components and the coefficients of the model are estimated using these weights. We also compare the performance of the weighted mixed model to other models already proposed in the literature using the aforementioned real data.

## Introduction

There is an increasing demand in local and central Governments in knowing precise estimates in domains where the size of the samples is small or even zero. These domains are called small areas. Traditionally, the sample sizes are chosen to provide reliable estimates for large geographical regions or aggregates of small areas. However, the statistical methods used for large domains can rarely be applied to small ones. Then, the problem of small area estimation is twofold. First, the fundamental question of producing reliable estimates of characteristics of interest and second, the assessment of the estimation error. When the sample in a given area is very small, a solution to the estimation problem is *to borrow strength* from related areas by means of auxiliary information.

Different model-based methods to accomplish small area issues have been proposed in the literature (Rao, 2003). Battese, Harter and Fuller (1988), popularized the use of linear mixed models in agricultural small area problems. They gave a prediction of the mean hectares of soybeans and corn per segment in 12 counties of Iowa with 36 segments, using as auxiliary information the classified corn and soybean hectares provided by satellite images. The authors consider a simple random sampling plan and segments of 250 hectares entirely included in the study domain. A common approach to account for other sampling plans by using sampling

weights has been done by Prasad and Rao (1999) and You and Rao (2002) who develop design-consistent small area estimation models. In these models the variance components are estimated from a unit-level model but the authors do not incorporate weights into the estimation process.

The aim of this paper is to propose a unit level linear mixed model and an area level linear mixed model where both, the variance components and the coefficients of the model are estimated using sampling weights. The models performance is illustrated by estimating the total area of olive trees in a region called Comarca IV, located in the middle of Navarra, Spain. Traditionally, small area linear mixed models have been used for this purpose based on the common definition of regular quadrats (also called segments) as sampling units, and assuming that these are fully included in the study domain. However, one important feature of this sample is that the square segments are very small, only of 4 hectares, but often, not completely included in the very irregular study domain. The size of sampled segments was limited by the precision of satellite images and could not be reduced.

References

Battese, G. E., Harter, R. M. and Fuller, W. A. (1988). "An Error-Components Model for Prediction of Country Crop Areas Using Survey and Satellite Data". Journal of the American Statistical Association, Vol.83, pp. 28-36.

Prasad, N.G.N. and Rao, J.N.K. (1999). "On Robust Small Area Estimation Using a Simple Random Effects Model". Survey Methodology, Vol. 25, pp. 67-72.

Rao, J.N.K. (2003). Small Area Estimation. Wiley Series in Survey Methodology.

You, Y. and Rao J.N.K. (2002). "A pseudo-empirical best linear prediction approach to small-area estimation using survey weights". Canadian Journal of Statistics, Vol 30, pp. 431-439.