Federica Piersimoni[1], Paolo Postiglione[2]

Spatial Sampling for Agricultural Data

Abstract

The particular nature of geo-referenced data has influenced the definition of the models for the analysis of spatial data, but surprisingly, we do not find a similar research effort for the issue of spatial sampling. This phase is crucial and influences in a remarkable way the following step of data analysis. In particular, our interest is mainly focused on spatially distributed agricultural data, which are extremely important as a support for both policy makers and market stakeholders [1].
The importance of sampling spatial units is recently acknowledged in various practical problems. In most cases, spatial units are defined across a geographical domain partitioned into a number of predetermined regularly or irregularly shaped locations. In standard sampling theory, spatial units have been traditionally represented as a mosaic of areas in which individual primary units are essentially viewed as identical members of the same population [4]. When population units are geographically distributed, classical random sampling strategies may be inefficient designs. In fact, nearby locations tend to have more similar values for measured attributes than distant ones. The choice of neighboring locations adds less additional information about the target area. So, it is clear that the definition of sampling schemes for spatial units can not be reasonably treated unless taking into account the concept of spatial dependence.
Sampling agricultural methods in agriculture can be based on a list frame or on spatial frame [3]. Each method has advantages and drawbacks: in general, list frame surveys are cheaper because sampled farms provide in one interview a large amount of information on crop area and yields, livestock, inputs or socio-economic variables [2]. On the other hand, spatial frame samples are better protected against non-sampling errors due to missing or overlapping units in the frame. The units of a spatial frame can be points, transects (i.e. lines of a certain length) or portion of territory, often named segments.
In this paper, our goal is to present the state of art about the methods that can be used in spatial agricultural surveys. In particular, the first aim is to review some spatial sampling designs in the presence of spatial dependence and some spatial frame-based surveys that are used for agricultural data (for example, the *June Area Survey* (JAS), the *Land Use/Cover Area frame statistical Survey* (LUCAS), the Italian *AGRIT* program, the French TER-UTI). We motivate the need for *ad hoc* spatial techniques to achieve higher levels of efficiency.
Furthermore, we suggest some ideas about the use of an algorithm based on distances in which the statistical units are, one from each other, as distant as possible. The interpretation of this statement in terms of variability means that the selection method that we propose will probably produce a sampling distribution of the estimates which is more representative of the population because searching for distance among the sample units, means to increase the variability inside the sample, and thus to decrease the variance between samples. The performances of the proposed method are evaluated on real data.

1. *ISTAT, Agricultural Statistics Service, (speaker)*, Italy, piersimo@istat.it
2. *University of Chieti-Pescara, Department of Economic Studies,* Italy, postigli@unich.it

References

[1] Benedetti R., Piersimoni F., Postiglione P. (2012) *Statistics in agricultural production*. In: N. Balakrishnan (Ed). Methods *and Applications of Statistics in the Social and Behavioral Sciences*, John Wiley & Sons, 357-373.
[2] Cotter J., Davies C., Nealon J., Roberts R. (2010) Area frame design for agricultural surveys. In: R. Benedetti, M. Bee, G. Espa, F. Piersimoni (Eds). *Agricultural Survey Methods*, John Wiley & Sons, 169-192.
[3] FAO (1996) *Multiple frame agricultural surveys. v. 1: Current surveys based on area and list sampling methods*, Statistical Development Series n.7.
[4] Haining R.P. (2003) *Spatial data analysis: theory and practice*, Cambridge University Press.

1. *ISTAT, Agricultural Statistics Service, (speaker)*, Italy, piersimo@istat.it
2. *University of Chieti-Pescara, Department of Economic Studies,* Italy, postigli@unich.it