Spatial Network Sampling in Small Area Estimation

Roberto Benedetti University "G. d'Annunzio" of Chieti-Pescara, Italy benedett@unich.it

Stefano Marchetti University of Pisa, Italy stefano.marchetti@ec.unipi.it

Federica Piersimoni Istat, Italian National Institute of Statistics, Italy piersimo@istat.it

Monica Pratesi* University of Pisa, Italy m.pratesi@ec.unipi.it

In the last decades, the GIS and GPS technologies have been developing significantly, making affordable to add and maintain information regarding the exact or estimated geographical position of each population unit in the frame. The spatial distribution of a population represents an important tool to design samples. In fact, the information arising from the network of the contiguities between population units has been used in many spatial designs. Spatial frames represent a source of relevant auxiliary information at the design stage also to prepare a more effective analysis of the collected data. The leading idea of this contribution is to propose a spatial sampling design that envisions the analysis for unplanned geographical domains or, in other terms, makes more efficient the application of small area estimation methods. The proposed spatially balanced sampling design guarantees a more even allocation of the sample inside regions and areas, increasing the sample size in unplanned domains and reducing the variability of the expected sampling ratio in comparison to that obtained under simple random sampling.

The contribution presents some algorithms to select samples under spatial constraints (distant units have higher probability of selection than close units). Some examples on real and simulated data are presented where the RMSE of the EBLUP estimates applied to samples selected by these network methods are compared with those obtained by using the Generalized Random Tessellation Stratified (GRTS) method. The proposed algorithm is computationally intensive but seems to be a feasible solution even dealing with large finite network populations frames.

Key Words: Spatially balanced samples, Small Area Estimation, EBLUP, Correlated Poisson Sampling, Pivotal Method, Generalized Random Tessellation Stratified design.