Geographic information in Small Area Estimation. Small area models with spatially correlated random area effects.

Alessandra Petrucci
Dipartimento di Statistica "G. Parenti,
Università degli Studi di Firenze, viale Morgagni, 59 – 50134 Firenze
alex@ds.unifi.it

Monica Pratesi, Nicola Salvati
Dipartimento di Statistica e Matematica Applicata all'Economia,
Università degli Studi di Pisa, via Ridolfi, 10 – 56124 Pisa
m.pratesi@ec.unipi.it salvati@ec.unipi.it

Abstract: Small area indirect estimators are often based on area level random effects models. Under this class of models, when only aggregate specific covariates are available, the Best Linear Unbiased Predictor (BLUP) is obtained under the assumption of uncorrelated random area effects (Fay and Herriot, 1979). The EBLUP takes advantage of the between small area-variation. The evidence is that the EBLUP estimator is significantly better than the sample-size dependent estimators, especially when the between small area-variation is not large relative to the within small area variation (Rao and Choudhry, 1995). This suggests that the location of the small areas may also be relevant in modelling the small area parameters and that further improvement in the EBLUP estimator can be gained by including eventual spatial interaction among random area effects (Petrucci, Salvati, 2004a; Pratesi, Salvati, 2005). Spatially correlated effects can also have a pragmatic role (Cressie, 1991). Ideally all relevant variables are chosen in the model, or proxies for them appear in the regression relation.
These variables - and the dependent variable - often all vary spatially, so the benefit obtained from including spatial dependence is presumed to be considerable. In addition, it should be noted that small area boundaries are generally defined according to administrative criteria without considering the eventual spatial interaction of the variable of interest. As a result, there is no reason to exclude the assumption that the random effects between the neighbouring areas are correlated and that the correlation decays to zero as distance increases.

This work deals an extension of the Fay-Herriot model with spatial correlation between the random small area effects modelled through the Simultaneously Autoregressive (SAR) process (Petrucci, Salvati, 2004a; Pratesi, Salvati, 2005). The best linear unbiased predictor under this model is called Spatial BLUP. Its empirical version (EBLUP) is obtained and an estimator of its MSE is proposed. Relative performances of the Spatial EBLUP are evaluated through a Monte Carlo experiment (Pratesi, Salvati, 2005).

Moreover, in some study it happens that, some small areas are not represented in the sample. This problem can be addressed specifying a nested error unit level regression model with dependent area level random effects (Petrucci, Salvati, 2004b). Allowing area random effects to be spatially correlated, the Empirical Best Linear Unbiased Predictions for the area parameters can be computed, taking into account also the contribution of the random part of the model, for sampled areas as well as out of sample areas (Saei, Chambers, 2005).

The properties of various estimators are evaluated applying the proposed estimator to two environmental case studies.

References

Cressie, N. (1991): Small-Area Prediction of Undercount Using the General Linear Model. *Proceedings of the Statistic Symposium 90: Measurement and Improvement of Data Quality*, Ottawa: Statistics Canada, 93-105.

Fay, R.E., Herriot, R.A. (1979): Estimates of income for small places: an application of James-Stein procedures to census data, *Journal of the American Statistical Association*, *74*, 269–277.

Petrucci Alessandra, Salvati Nicola (2004a): Small Area Estimation Using Spatial Information. The Rathbun Lake Watershed Case Study, *Working Paper n° 2004/02, Dipartimento di Statistica "G. Parenti", Firenze.*

Petrucci, A., Salvati, N. (2004b): Small Area Estimation considering spatially correlated errors: the unit level random effects model, *Working Paper n° 2004/10, Dipartimento di Statistica "G. Parenti", Firenze*

Pratesi, M., Salvati, N. (2005): Small Area Estimation: the EBLUP estimator with autoregressive random area effects, *Report n° 261, Dipartimento di Statistica e Matematica Applicata all'Economia, Pisa.*

Rao, J.N.K., Choudhry, G.H. (1995): Small Area Estimation: Overview and Empirical Study in *Business Survey Method*, Edited by Cox, Binder, Chinnappa, Christianson, Colledge, Kott, John Wiley & Sons, 38, 527-540.

Saei, A, Chambers, R. (2005): Empirical Best Linear Unbiased Prediction for Out of Sample Areas, *Working Paper M05/03, Southampton Statistical Sciences Research Institute, University of Southampton.*