SAE motivation

ISTAT has a long tradition in the development and application of small area estimation methods. The office was a member of the EURAREA Consortium (Enhancing Small Area Estimation Techniques to meet European needs), a research programme funded by Eurostat from 2000-2004, and has been taking part in several other international projects over the last 20 years (e.g. ESSnet on Small Area Estimation and Horizon 2020 - MAKSWELL).

Motivation for the use of small area estimation methods is the need of "local governments for accurate information for small geographical areas" or specific domains. Since 2002, ISTAT has been applying small area estimation methods to produce estimates of employment and unemployment rates for local labour market areas (D'Alò, 2008). From 2004, a unit-level EBLUP estimator with spatially autocorrelated random area effects has been introduced (D'Alò et al., 2012; D'Alò et al., 2017). ISTAT publishes these small area official statistics annually and has recently made available the estimates for the year 2019, updating the historical series starting from 2006.

In addition, ISTAT developed experimental small area estimates of a selection of labour market variables for cities and functional urban areas (FUA) based on a unit-level multivariate model. These are published as Italian sub-national statistics within the scope of an agreement between ISTAT and Eurostat. A city usually consists of a single municipality and a FUA is composed of a group of economically integrated municipalities based on the OECD-EC city definition.

Indicators in the scope of the study for experimental statistics by city and FUA

The main indicators produced are:

- Economically Active Population, total, and by sex
- Economically Active Population aged 20-64, total, and by sex
- Persons Unemployed, total, and by sex
- Persons Employed aged 20-64, total, and by sex

Input data

The Italian Labour Force Survey (LFS) referring to year 2018 is the source of direct estimates whereas the Integrated System of Registers, in particular the Base Register of Individuals (RBI) and the Thematic Labour Registry (RTL), the Italian Ministry of Finance and the Revenue Agency provided auxiliary data (demographic, employment, social security and income information).

The LFS is the main source of information on the Italian labour market and produces monthly, quarterly, and yearly estimates of employment, unemployment, and inactivity rates for different geographical areas, however cities and FUAs constitute unplanned domains.

Building the SAE model/ Model Building

The small area estimation method is based on a multivariate model implemented in the R package MIND (Multivariate model-based INference for Domains), developed by ISTAT. The method is a multivariate version of the unit level linear mixed model that can incorporate two or more random effects and a multivariate qualitative variable as dependent variable in the model (D'Alò et al., 2021).

The multivariate response variable is a vector composed by three dichotomous variables representing the mutually exclusive, and exhaustive, categories of labour market status (employed, unemployed and out of the labour force). The domains of interest are specified combining the geographical areas (cities and FUAs) with age and sex groups, as required according to the target indicators. Information about model selection procedures, model predictors, as well as evaluation of model-based estimates, are reported by D'Alò et al. (2021), and the data are disseminated by Eurostat.

Benchmarking/data validation

Traditional quality assessment of model-based estimates was carried out, e.g. comparing final model results to the direct estimates. Since the reported small area estimation project was developed in the scope of the Eurostat City Statistics initiative, procedures for ensuring data quality defined in the Methodological Manual on City Statistics (Eurostat, 2017) were also implemented.

In addition, guidelines for good practice on small area estimation are followed:

- ESSNet on Small Area Estimation - Report on Workpackage 6 - Guidelines
- Guidelines on small area estimation for city statistics and other functional geographies (Eurostat, 2019)
- From start to finish : a framework for the production of small area official statistics (Tzavidis et al., 2018)
Update of SAE methods and of small area estimates

Labour market official and experimental small area estimates are published annually, hence ISTAT is frequently engaged in model checking and validation activities. Also, other small area estimates (listed below) are currently being produced, requiring continuous development of methods and practices.

Other indicators related to SDGs that are produced for unplanned survey domains (14 metropolitan cities):

- Percentage of people aged 14 and over who have used the internet in the last 12 months to relate for private use with the Public Administration or with public service providers (Goals 9 and 17)
- People aged 15 and over who leave home to go to work by bicycle or on foot out of the total number of employed people (Goal 11)
- Household overcrowding (Goal 11)

In addition, the office has plans for enhancing the use of SAE methods and is carrying out research in the following topics/models:

- Multivariate unit level models (MIND) that consider: spatial and time correlation among area and marginal effect, correlation among variables; as well as models that incorporate temporally, or spatially, correlated area and marginal effects;
- Methods that, taking into account sampling design and benchmarking needs, would deliver a set of weights for each domain of interest that can be used to produce the estimates of all target indicators based on a survey (reducing the disadvantage of current SAE tailor-made methods);
- Small area estimation of unemployment using Latent Markov Models (Bertarelli et al., 2018);
- Area level models with distributional hypotheses that are appropriate to the chosen target parameter that may reduce the bias of estimates;
- Development of design-based measures of MSE for assessing a model-based estimator on a design-based perspective.

SAE work within the organisation

There is a dedicated team in ISTAT comprised of 4 methodologists that liaises with other office units through working groups. The joint work allows specification of the user needs, the required level of disaggregation for small area statistics, the target indicators, the available auxiliary information; and provides an environment for validation of the model results.

Future work on SAE

ISTAT plans to establish a task force to define, develop and propose new small area estimates for unplanned domains of some SDG indicators derived from the main surveys.

- Goal 1 - End poverty in all its forms everywhere
  - Relative and Absolute Poverty indicators from EU-SILC and the Household Budget Survey (Laken indicators and others) for metropolitan cities and provinces.
  - Different estimators have already been tested (EBP - emdi package), area and unit level EBLUP, unit level spatial EBLUP - MIND package.
  - The challenge is how to choose the best method and the benchmarking approach.

- Goals 9 and 17 - Build resilient infrastructure, promote sustainable industrialization and foster innovation; Strengthen the means of implementation and revitalize the Global Partnership for Sustainable Development
  - ICT indicators based on Multipurpose survey on everyday life aspect for metropolitan cities and provinces.
  - Specific small area estimators can be applied like the design-based projection estimators, besides the classic model-based small area estimators (D’Alò et al., 2018)

- Goal 5 - Achieve gender equality and empower all women and girls
  - Health indicators based on European Health Interview Survey for Italian health districts/regions.
  - Direct and model-based estimates for indicators related to violence against women.

Challenges

ISTAT team indicates the decision on the best method/model to be applied for a group of indicators (besides model diagnostics and the assessment of estimates) as the challenging aspects related to small area estimation projects.

In addition, they highlighted the availability of auxiliary information for model building, the integration of survey and administrative data, the quality assessment of the estimates and the selection of an appropriate method/model as important elements for successfully using SAE for official data production.

Sources/References


Information provided by Dr Michele D’Alò - Head of the Initiative “Design of small area estimation strategies for sample surveys” and Dr Stefano Falorsi - Head of the Division “Process design and support to the system of statistical registers”.