Indonesia

Organisation

Badan Pusat Statistik (BPS-Statistics Indonesia)

SAE motivation

Research initiatives and the production of small area statistics in Indonesia have been reported since 2003 (Suryahadi et al., 2003; Notodiputro, K.A. and Kurnia, A., 2007). The 2005-2009 BPS Strategic Plan for Statistical Development defined “the development of an efficient and low-cost methodology, which allows for the creation of small area and local specific statistics data” as one of the main activities to support government decentralization (BPS, 2006).

More recently, in 2019, Statistics Indonesia started a research project on small area estimation (SAE) methods for producing required SDG indicators and other statistics by domains of interest and/or by metropolitan statistical areas (MSA).

An analysis of data availability indicated that not all BPS surveys can be employed to produce SDGs indicators, or the survey estimates do not meet the SDG disaggregation requirements. In addition, local governments require support of BPS data for enabling a decentralised planning environment but there is a lack of data at regency/city level.

Therefore, SAE methods are considered as an alternative for producing reliable official statistics for the world’s largest island country (around 6,000 islands are inhabited) with sizable population (270.2 million people in 2020), integrating census and survey data to overcome the challenge of limited resources for statistics activities.

Indicators in the scope of the study

Strategic Indicators: poverty rate, unemployment rate, Gini coefficient, demographic indicators

Indicators by Metropolitan Statistical Area (MSA): unemployment rate, poverty rate, owned mobile phone/hp, Internet usage, morbidity rate, health complaint, average years of schooling, adjusted per capita expenditure, Human Development Index (HDI).

Indicators for Metropolitan Statistical Area (MSA):

Unemployment rates by subdistrict

Percentage of households with non-electric lighting by subdistrict

Percentage of households without a source of clean drinking water by subdistrict

Percentage of poor people by subdistrict

Average years of schooling for girls by subdistrict

Mortality rate by subdistrict

SDG Indicators at district (regency/city) level:

- Proportion of children (age <18 years) living below the national poverty line or living below twice national poverty line;
- Percentage of infants under 6 months of age who are exclusively breastfed;
- Proportion women aged 15-49 years old who gave birth in health facilities;
- Participation rate in organized learning (one year before the official primary entry age), by sex;
- Completion rate or out of school (primary education, lower secondary education, upper secondary education);
- Percentage of households with access to improved and sustainable sanitation;
- Proportion of children under 5 years of age whose births have been registered with a civil authority, by age;
- Proportion of population using safely managed sanitation services, including a hand-washing facility with soap and water;
- Proportion of women aged 20-24 years who got married under the age of 15 and 18.

SAE work within the organisation

BPS established a strategic research programme comprising 5 steps: a workshop on SAE fundamentals, literature review, data exploration (comparing data from other countries with Indonesian data), experimental SAE with Indonesian data for chosen indicators and evaluation.
The programme has a plan with well-defined phases until 2024, when BPS expects to implement SAE methods for official statistics. The research is carried out by BPS employees from several office units, forming an SAE ad-hoc team comprised of SAE technical coordinators, SAE experts (with PhD /academic degrees), data experts (on the substantive area) and interns. In addition, SAE team liaises with researchers and lecturers from Politeknik Statistika STIS - the official college of BPS, and from other universities in the country.

**Collaboration with other stakeholders**

BPS has undertaken a collaboration with universities, government institutions and international organizations for developing research in SAE modelling. Bogor IPB University and Politeknik Statistik (Polstat STIS) are two universities that provide full technical support for BPS in the forms of workshops and SAE training.

BPS has a collaboration with Ministry of Development Planning/Bappenas in a national project for delineating metropolitan areas in several regions. In this project, SAE BPS Team is expected to produce several strategic indicators for Metropolitan Statistical Area.

SAE development in BPS has also received support from UNICEF in the last couple years. UNICEF’s support has been in such activities as workshops and ground check to the field. The collaboration between BPS and UNICEF in the SAE development has been focused on producing SDGs indicators.

There is also another collaboration between BPS and other institutions especially in finding possible auxiliary information for SAE modelling from other data sources. These auxiliary variables could be taken from administrative/registration data and big data.

**Input data**

The Indonesian National Socioeconomic Survey (SUSENAS) is the source of direct estimates and auxiliary variables with measurement errors whereas the Village Potential Statistics (PODES) provides auxiliary data without measurement errors.

SUSENAS is a repeated sample survey, conducted every year, with big sample size and target population comprising all Indonesian districts.

PODES collects data of specific variables for every village in Indonesia and is the source of the main auxiliary variables for SAE.

**Building the SAE model/ Model Building**

Due to multiple characteristics of target indicators and auxiliary variable, the SAE project roadmap indicates a variety of SAE methods to be tested/evaluated: SAE models with measurement errors in covariates, non-normal SAE models, hierarchical Bayesian models, and etc. In the early phase of SAE development, a basic SAE model without measurement error in auxiliary variables was developed. Further problem emerged as auxiliary information would be very limited if the SAE modelling only relies on auxiliary variables with no measurement error from certain sources, which are not available annually. Further development, then, focuses on the SAE development with measurement errors in auxiliary variables, which can be obtained from survey activities or indicators that have already been produced officially.

Target indicators are in the form of counts, proportions, indexes, means and totals. Besides that, many indicators produced directly from the survey results could provide a very different pattern of errors among regions. So, it is necessary to explore different SAE methods.

**Benchmarking/data validation**

There is need to validate the results of SAE model in order to provide accurate estimates. In terms of benchmarking the SAE results, model-based estimates at lower level of administrative areas should be compatible with indicators officially published at a higher level of administrative areas. The aggregation of SAE model-based estimates for all lower administrative areas will equal a higher-level estimate officially produced by BPS.

To validate the results of SAE model, the current practice implemented by BPS is to conduct an evaluation process. In this ground check, in-depth interview or a focus group discussion is conducted to examine the SAE estimation results with key informants (district/local government, experts, academia, and local BPS officers) in order to get their insights about the ranking and the level of estimates.

**Update of SAE methods**

BPS has been continuously improving SAE models. The development is needed to account for a number of challenges such as administrative areas with no data sample, non-normal data distribution and the possible use of auxiliary information from survey data (auxiliary variables with sampling error).

Some SAE models that have already been implemented by BPS include:

- **SAE Cluster** for some small areas that have no samples,
- **SAE Benchmarking**;
- **SAE with Log transformation and SAE with measurement error for non-normal data**;
- **SAE Hierarchical Bayesian -Beta-distribution for non-normal data, especially when target the indicator is a proportion**;
- **SAE-Non Parametric for nonlinear data**.

The SAE implementation challenge is how to update models to produce estimates for succeeding periods. The auxiliary variable selected for the SAE models should be regularly available from BPS statistical activities or from administrative records.
Use of SAE estimates

SAE estimates are produced for two main purposes. First, to support government policies and these are not available for public domain. Indicators produced by SAE modelling to support the development of Metropolitan Statistical Area (MSA) is of one example.

Second, small area estimates are also being developed to be published by BPS. In the near future, SAE methods are expected to provide a number of indicators to support the monitoring and evaluation of SDGs implementation at district level.

Capacity building for SAE production

Capacity building on SAE modelling is very crucial for the development of SAE BPS team and other human resources in a statistical office. It will be very useful for the SAE implementation widely. Sharing knowledge and capacity building cover topics including theory and concept of SAE modelling, the use of R-software (or other software) for SAE modelling and implementation.

Future work on SAE

The future work on SAE will focus on several areas. First, BPS will develop the standard procedure in SAE modelling. Second, the office will improve the method for model validation to be able to disseminate the results of SAE modelling for official statistics. Third, BPS will explore SAE spatial models as previous work indicates the effect of spatiality. Next, small area estimates will be published as official statistics, and the use of SAE methods will be implemented intensively for other strategic indicators. In addition, R-software based SAE packages will be developed for various SAE models to facilitate the practical implementation of SAE models.

Challenges

- Model building complexity
  - The model building procedure is laborious: besides technical expertise, it requires knowledge about the target indicator and auxiliary variables.
  - Model building faces numerous challenges such as: the lack of auxiliary variables and auxiliary variables with measurement error, lack of sample units in small areas or domains, data skewness of target variable (in case of unit level models), reliability of estimates is not straightforward to calculate, there may be problems with model fitting quality plus the need to validate the model, and corresponding estimates, with local government officials and experts.
- SAE models are tailor made solutions – one model to each target indicator.
- SAE team members are usually very busy on their main tasks/functions, not fully dedicated to SAE project. They are from several BPS units, making it more difficult to coordinate a joint work.

Sources/References


Also: Information provided by Ms. Haoyi Chen from meetings with BPS SAE staff and BPS presentations.