Software packages

To simplify the application of existing SAE methods, available software packages can help. A large range of SAE methods is developed in several software packages such that the choice depends on personal preferences or available licenses. Open software that enable the application of SAE methods are R, Python and the software package PovMap by the World Bank.

- R
- Stata
- SAS
- Python
- Other

References

R

The R project as a free software environment enables everyone to contribute to it. Thus, new methodology is often fast implemented in R. This is also the case for small area estimation methods. For beginners in R, the introduction to small area estimation techniques of the Asian Development Bank provides, among many online tutorials, an introduction to data management using R (Chapter III).

There is a large number of packages that offer the estimation of small area estimates. A first overview is available on the CRAN TASK View for Official Statistics. Besides the packages for standard linear mixed models (nlme and lme4), it proposes six packages for small area estimation: sae, emdi, rsae, hbsae, JoSAE, BayesSAE. The figure below shows the number of downloads by months for these six packages to give an indication of their usage. For all packages, it can be seen that the number of downloads increased over time.

Some packages provide a wide range of different methods, while others are more specialized on a specific method. The following table gives an overview of some R packages for SAE but it has no claim to completeness. The names of the packages serve as a link to the CRAN page while the model type columns gives a first indication which model types are available in the package. For applying the available functions, the vignettes and papers about the packages are useful. These usually contain examples and explanations on how to use the implemented functions.

Overview of R packages for small area estimation

<table>
<thead>
<tr>
<th>Package</th>
<th>Model type</th>
<th>Vignettes /Paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>sae</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

- Basic direct and indirect estimators in sae
- R package sae: Methodology
- sae: An R Package for Small Area Estimation
### R Packages for Small Area Estimation

<table>
<thead>
<tr>
<th>Package</th>
<th>Description</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>emdi</strong></td>
<td>The R Package <em>emdi</em> for Estimating and Mapping Regionally Disaggregated Indicators</td>
<td>A Framework for Producing Small Area Estimates Based on Area-Level Models in R</td>
</tr>
<tr>
<td>povmap</td>
<td>Povmap - extension to the 'emdi' package (The R package 'povmap' supports small area estimation of means and poverty headcount rates. It adds several new features to the 'emdi' package (see &quot;The R Package emdi for Estimating and Mapping Regionally Disaggregated Indicators&quot; by Kreutzmann et al. (2019) <a href="">doi:10.18637/jss.v091.i07</a>). These include new options for incorporating survey weights, ex-post benchmarking of estimates, two additional transformations, several new convenient functions to assist with reporting results, and a wrapper function to facilitate access from 'Stata'.)</td>
<td>povmap: Extension to the emdi package for small area estimation</td>
</tr>
<tr>
<td>rsaep</td>
<td>Robust Small Area Estimation: a Vignette</td>
<td></td>
</tr>
<tr>
<td>hbsae</td>
<td></td>
<td></td>
</tr>
<tr>
<td>JoSAE</td>
<td>Vignette of the JoSAE package</td>
<td></td>
</tr>
<tr>
<td>BayesSAE</td>
<td></td>
<td></td>
</tr>
<tr>
<td>saery</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mme</td>
<td>mme: a package for small area estimation with multinomial mixed models</td>
<td></td>
</tr>
</tbody>
</table>
### Stata

Especially in economic research, Stata is a widely used software tool. In recent years, some commands for small area estimation were provided. The World Bank developed Stata functions for unit-level small area estimation which are especially useful for poverty mapping. The FHSAE module translates the area-level models from the R package sae and the fayherriot command supports the estimation of area-level models with various transformations.

#### Overview of Stata packages for small area estimation

<table>
<thead>
<tr>
<th>Package</th>
<th>Model type</th>
<th>Vignettes/Paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>sae</td>
<td>X</td>
<td>• sae: A Stata Package for Unit Level Small Area Estimation</td>
</tr>
<tr>
<td>FHSAE</td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>fayherriot</td>
<td>X</td>
<td>• The fayherriot command for estimating small-area estimators</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• A practical guide for the computation of domain-level estimates with the Socio-Economic Panel (and other household surveys)</td>
</tr>
</tbody>
</table>
SAS

SAS computer programs were developed by the EURAREA project for small area estimation. Since these are not further developed, Statistics Canada built a prototype for their small area estimation system (Hidiroglou et al. 2019).

The work of Mukhopadhyay and McDowell (2011) describes procedures to fit the basic unit-level and area-level models. In the following, the procedures will be referred to.

Overview of SAS packages for small area estimation

<table>
<thead>
<tr>
<th>Package</th>
<th>Model type</th>
<th>Vignettes/Paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>StatCan prototype</td>
<td>X</td>
<td>Area</td>
</tr>
<tr>
<td>Mukhopadhyay and McDowell</td>
<td>X</td>
<td>Area</td>
</tr>
</tbody>
</table>

- Development of a small area estimation system at Statistics Canada
- Small Area Estimation for Survey Data Analysis Using SAS® Software

Python

Besides R, Python is another programming language for which modules are provided that can be accessed without any license. In 2021, the python package samplics was released that implements a set of sampling techniques for complex survey designs including small area estimation.

Overview of Python packages for small area estimation

<table>
<thead>
<tr>
<th>Package</th>
<th>Model type</th>
<th>Vignettes/Paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>samplics</td>
<td>X</td>
<td>Area level model: empirical best linear unbiased predictor (EBLUP)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Unit level model: empirical best linear unbiased Prediction (EBLUP)</td>
</tr>
</tbody>
</table>

Other

One of the first opportunities to conduct poverty mapping based on the unit-level approach by Elbers, Lanjouw and Lanjouw (ELL or World Bank method) is the software PovMap developed by the World Bank.

Overview of further packages for small area estimation

<table>
<thead>
<tr>
<th>Package</th>
<th>Model type</th>
<th>Vignettes/Paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>PovMap</td>
<td>X</td>
<td>Using PovMap2: A user’s guide</td>
</tr>
</tbody>
</table>
References


