STATISTICS IN TRANSITION new series and *SURVEY METHODOLOGY*
Joint Issue: Small Area Estimation 2014
Vol. 17, No. 1, pp. 105–132

DEVELOPMENT OF SMALL AREA ESTIMATION IN OFFICIAL STATISTICS[1]

Jan Kordos[2]

ABSTRACT

The author begins with a general assessment of the mission of the National Statistics Institutes (NSIs), main producers of official statistics, which are obliged to deliver high quality statistical information on the state and evolution of the population, the economy, the society and the environment. These statistical results must be based on scientific principles and methods. They must be made available to the public, politics, economy and research for decision-making and information purposes.
Next, before discussing general issues of small area estimation (SAE) in official statistics, the author reminds: the methods of sampling surveys, data collection, estimation procedures, and data quality assessment used for official statistics. Statistical information is published in different breakdowns with stable or even decreasing budget while being legally bound to control the response burden.
Special attention is paid, from a practitioner point of view, to synthetic development of small area estimation in official statistics, beginning with international seminars and conferences devoted to SAE procedures and methods (starting with the Canadian symposium, 1985, and the Warsaw conference, 1992, to the Poznan conference, Poland, 2014), and some international projects (EURAREA, SAMPLE, BIAS, AMELI, ESSnet). Next, some aspects of development of SAE in official statistics are discussed. At the end some conclusions regarding quality of SAE procedures are considered.

Key words: small area estimation, official statistics, sampling survey, direct estimation, indirect estimation, empirical Bayes estimator; hierarchical Bayes estimator; data quality.

1. Introduction

National Statistics Institutes (NSIs) are the most important statistical information providers for official statistics. Their mission is to produce high

[1] This is an updated and extended version of the first part of the paper entitled "Small Area Estimation in Official Statistics and Statistical Thinking" presented at the International Conference on Small Area Estimation 2014, held in Poznan, Poland, 3-5 September 2014.

[2] Central Statistical Office of Poland and Warsaw Management Academy.

quality statistical information on the state and evolution of the population, the economy, the society and the environment. These statistical results must be based on scientific principles and methods. They must be made available to the public, politics, economy and research for decision-making and information purposes. One important challenge that NSIs have to face is the growing users' demand with stable or even decreasing budget while being legally bound to control the response burden. The use of more and more efficient statistical methods is a way to take up this challenge. To collect, estimate, process and publish statistical information NSIs use different methods and procedures, but special emphasis is paid to sampling surveys, taking into account basic needs, cost and respondent burden. For this reason, issues connected with sampling surveys in official statistics from a practitioner point of view are considered first here, using different approaches, methods, and variety of data, mainly sampling data, censuses and registers (Brakel & Bethlehem, 2008; Little, 2004. 2012).

2. Sampling surveys in official statistics and issues of SAE methods

First, the author would like to remind that the purpose of sampling surveys is to obtain statistical information about a finite population by: a) selecting a probability sample from this population, b) obtaining or measuring the required information about the units in this sample, and c) estimating finite population parameters such as means, totals, ratios, etc., and assessing their variances (Brakel & Bethlehem, 2008). The statistical inference in this setting can be: (i) *design-ba*sed, (ii) *model-assisted* or (iii) *model-based*. In the design-based and model-assisted approach, the statistical inference is based on the stochastic structure induced by the sampling design. Parameter and variance estimators are derived under the concept of repeatedly drawing samples from a finite population according to the same sampling design, while statistical modelling plays a minor role. This is the traditional approach of survey sampling theory, followed by authors like Hansen et al. (1953), Kish (1965), Cochran (1977), Yates (1981) and Särndal et al. (1992).

In the model-based context, the probability structure of the sampling design plays a less pronounced role, since the inference is based on the probability structure of an assumed statistical model. This is the position taken by authors like Gosh and Meeden (1997), Gosh &Rao (1994), Rao (1999), Valliant et al. (2000), Rao (2003), Pfeffermann (2002, 2013) and Jang & Lahiri (2006). An overview of the different modes of inference in survey sampling is given by Little (2004).

Design-based and model-assisted estimators refer to a class of estimators that expand or weight the observations in the sample with the so-called survey weights. Survey weights are derived from the sampling design and available auxiliary information about the target population. Functions of the expanded observations in the sample are used as (approximately) design-unbiased estimators for the unknown population parameters of interest. The associate inferences are based on the probability distribution induced by the sampling design with the population values held fixed.

A well-known design-based estimator is the π-estimator or Horvitz-Thompson estimator, developed by Narain (1951), and Horvitz and Thompson (1952) for unequal probability sampling from finite populations without replacement. The observations are weighted with the inverse of the inclusion probability, also called design-weights. This estimator is design-unbiased, since the expectation of the estimator with respect to the probability distribution induced by the sampling design is equal to the true but unknown population value.

The precision of the Horvitz-Thompson estimator can be improved by making advantage of available auxiliary information about the target population (Wywiał, 2000). In the model-assisted approach developed by Särndal et al. (1992) this estimator is derived from a linear regression model that specifies the relationship between the values of a certain target parameter and a set of auxiliary variables for which the totals in the finite target population are known. Based on the assumed relationship between the target variable and the auxiliary variables, a generalized regression estimator can be derived of which most well-known estimators are special cases. After this estimator is derived, it is judged by its design-based properties, such as design expectation and design variance.

Most NSIs surveys are designed to provide statistically reliable estimates at national or high-level geographies. So when statistics are required for more detailed geographical areas or small subgroups of the population, the sample sizes just are not big enough to make reliable estimates. Increasing the size of samples would be prohibitively expensive – instead, estimation methods have been developed that combine data from administrative, census and survey sources to produce estimates for small areas or domains. There are many statistical techniques covered by small area estimation, a frequently used approach is a model-based one, where local area outcomes are estimated from the regression between survey data and auxiliary data from census and administrative data sources.

The great importance of SAE stems from the fact that many new programs, such as fund allocation for needed areas, new educational or health programs and environmental planning rely heavily on these estimates. SAE techniques are also used in many countries to test and adjust the counts obtained from censuses that use administrative records and for post-enumeration surveys after the population censuses for quality assessment. SAE is researched and applied so broadly because of its usefulness to researchers who wish to learn about the research carried out in SAE and to practitioners who might be interested in applying the new methods.

The problem of SAE is twofold. First, the fundamental question is how to produce reliable estimates of characteristics of interest (means, counts, quantiles, etc.) for small areas or domains, based on very small samples taken from these areas. The second related question is how to assess the estimation error. Budget and other constraints usually prevent the allocation of sufficiently large samples to each of the small areas. Also, it is often the case that domains of interest are only specified after the survey has already been designed and carried out. Having only a small sample (and possibly an empty sample) in a given area, the only possible solution to the estimation problem is to *borrow information* from other related data sets.

As it has been mentioned, and from theoretical point of view, SAE methods can be divided broadly into "*design-based*" and "*model-based*" methods. The latter methods use either the frequentist approach or the full Bayesian methodology, and in some cases combine the two, known in the SAE literature as "*empirical Bayes".* Design-based methods often use a model for the construction of the estimators (known as "*model assisted*"), but the bias, variance and other properties of the estimators are evaluated under the randomization (design-based) distribution. The randomization distribution of an estimator is the distribution over all possible samples that could be selected from the target population of interest under the sampling design used to select the sample, with the population measurements considered as fixed values (parameters). Model-based methods, on the other hand, usually conditioned on the selected sample, and the inference is with respect to the underlying model. A common feature to design- and model-based SAE is the use of auxiliary covariate information, as obtained from large surveys and/or administrative records such as censuses and registers. Some estimators only require knowledge of the covariates for the sampled units and the true area means of these covariates. Other estimators require knowledge of the covariates for every unit in the population. The use of auxiliary information for SAE is vital because with the small sample sizes often encountered in practice, even the most elaborated model can be of little help if it does not involve a set of covariates with good predictive power for the small area quantities of interest.

It is now generally accepted that the indirect estimates should be based on explicit models that provide links to related areas through the use of supplementary data such as census counts or administrative records. See, for example, Ghosh and Rao (1994), Rao (1999), Rao (2003), Pfeffermann (2002, 2013), and Jiang and Lahiri (2006) for more discussion on model-based small area methods.

Thus, the model-based estimates are obtained to improve the direct design-based estimates in terms of precision and reliability, *i.e.*, smaller coefficients of variation (CVs). Supplementary data are vital for improving quality of small area statistics. These data are used to construct predictor variables for use in a statistical model that can be used to predict the estimate of interest for small areas. The effectiveness of small area estimation depends initially on the availability of good predictor variables that are uniformly measured over the total area. It next depends on the choice of a good prediction model. Effective use of small area estimation methods further depends on a careful, thorough evaluation of the quality of the model. Finally, when small area estimates are produced, they should be accompanied by valid measures of their precision. Now, there is a wide range of different, often complex models that can be used, depending on the nature of the measurement of the small area estimates and on the auxiliary data available. One key distinction in model construction is between situation where the auxiliary data are available for the individual units in the population and those where they are available at aggregate level for each small area. In the former case, the data can be used in unit level models. Another feature involved in the choice of a model is whether the model borrows strength across sectional or over time, or both. There are also now a number of different approaches, such as empirical best linear prediction (EBLUP), empirical Bayes (EB) and hierarchical Bayes (HB) that can be used to estimate the models and the variability of the model dependent small area estimates (Choudry et al., 1989, 2012; Data. 2009; Datta et al.,1999, 2012; Gosh & Meeden, 1997, 1999; Kubacki, 2004; Lehtonen et al., 2003, 2005, 2009; Molina et al., 2009; Moura et al., 2002; Pfeffermann,1999, 2013; Pfeffermann & Tiller, 2006; Pratesi &Salvati, 2008; Rao, 2003, 2011; You & Dick, 2004). Moreover, complex procedures that would have been extremely difficult to apply a few years ago can now be implemented fairly straightforwardly, taking advantage of the continuing increases in computing power and the latest developments in software.

Thus, there are two broad classifications for small area models: ***area level models*** and ***unit level models***:

- ***Area level models*** that relate small area means and totals to area-specific auxiliary variables,
- ***Unit level models*** that relate the unit values of the dependent variable to unit specific auxiliary variables.

Among the ***area level models,*** the ***Fay-Herriot model*** (Fay and Herriot, 1979) *is a basic and widely used area level model in practice to obtain reliable model-based estimates for small areas.* The Fay-Herriot model basically has two components, namely, a sampling model for the direct estimates and a linking model for the parameters of interest. The sampling model involves the direct survey estimate and the corresponding sampling variance. The Fay- Herriot model assumes that the sampling variance is known in the model. Typically, a smoothed estimator of the sampling variance is obtained and then treated as known in the model. Wang and Fuller (2003), You and Chapman (2006), Gonzalez-Manteiga, et al. (2010), considered the situation where the sampling variances are unknown and modelled separately by direct estimators.

The linking model relates the parameter of interest to a regression model with area-specific random effects. In the Fay-Herriot model, the area random effects are usually assumed to be independent and identically distributed normal random variables to capture geographically unstructured variations among areas. However, in some small area applications, particularly in public health estimation problems, geographical variation of a disease is a subject of interest, and estimation of overall spatial pattern of risk and borrowing strength across regions to reduce variances of final estimates are both important. Thus, it may be more reasonable to construct spatial models on the area-specific random effects to capture the spatial dependence among them. The spatial models are generally used in health related small area estimation, and various spatial models have been proposed for small area estimation [(*e.g.* Ghosh et al., 1999; Moura et al., 2002; Pratesi and Salvati (2008), Singh et al., (1994) and Molina et al., (2009)]. Best et al., (2005) provided a comprehensive review on spatial models for disease mapping. Rao (2003) also discussed several spatial small area models.

The unit model originates with Battese, Harter and Fuller (1988). They used the nested error regression model to estimate county crop areas using sample survey data in conjunction with satellite information. Prasad and Rao (1990, 1999) were first to include the survey weights in the unit level model: they labelled their estimator as a pseudo-EBLUP estimator of the small area mean. Prasad and Rao (1999) also provided based expressions for the MSE of their estimator when it included the estimated variance components. You and Rao (2002) proposed an estimator of β that ensures self-benchmarking of the small area estimates to the corresponding direct estimator.

Thus, an *indirect estimator* uses values of the variable of interest from a domain and/or time period other than the domain and time period of interest. Three types of indirect estimators can be identified:

- A *domain indirect estimator* uses values of the variable of interest from another domain but not from another time period.
- A *time indirect estimator* uses values of the variable of interest from another time period but not from another domain.
- An estimator that is both *domain and time indirect* uses values of the variable of interest from another domain and another time period.

Individual level models work in two stages using regression modelling. Firstly, the survey data are used to predict the probability of the characteristic of interest based on the attributes of the individuals in the survey (such as gender, age and marital status). The aggregate levels of a cross tabulation of these individual characteristics for each local area are obtained, usually from the census, and the coefficients from the regression model are applied to those small area covariate values so as to calculate the expected value of the target outcome variable conditional on the area’s characteristics. The steps are relatively straightforward:

- Ensure that the predictor variables are available in both survey data and for small areas
- Fit a regression model to survey data to predict the probability of chosen outcome
- Use Wald tests to consider dropping non-significant variables.
- Extract parameter estimates and apply to small area data.

In short, SAE is a collection of different methods:

- *Synthetic methods* (often implicit assumptions on the nature of relationship; a simple case: rates at NUTS2 level = rates at NUTS4 level).
- *Composite methods* (linear combination of synthetic methods and direct estimators to balance bias and variance).
- *Estimator based on linear mixed models* (EBLUP, EB, HB).
- *Non-linear models* (e.g. logistic models for binary responses) with spatial and/or temporal correlation structures among random effects.
- *Semi-parametric models.*

Potentially more serious, with respect to *accuracy and quality*, are ***non-sampling errors*** **s**uch as ***coverage errors, measurement errors*** and ***response bias***. Most censuses miss some people, or count some people twice, and it has been repeatedly shown that those miscounted are generally not typical of the population as a whole. Census or sample survey estimates may therefore be biased against certain subgroups of the population. If these subgroups tend to be geographically clustered, this can have a serious impact on estimates for some small areas. Response bias arises if many respondents systematically misunderstand a census or a survey question or are unable or unwilling to give correct answer. Both small area and large area estimates would be affected by such errors (Bethlehem, 1988; Bethlehem et al. 1985; Brackstone, 1999, Eurostat, 2007; Holt et al, 1991; Kalton, 2002; Kalton & Kasprzy, 1986; Kordos, 2005; Longford, 2005; Rao, 2011; Trewin, 2002).

3. Use of administrative data in official statistics

NSIs around the world are coming under increasing pressure to improve the efficiency of the statistical production process, and particularly to make savings in costs and staff resources. At the same time, there are growing political demands to reduce the burden placed on the respondents to statistical surveys. Given these pressures, statisticians are increasingly being forced to consider alternatives to the traditional survey approach as a way of gathering data. Perhaps the most obvious answer is to see if usable data already exist elsewhere. Many nonstatistical organisations collect data in various forms, and although these data are rarely direct substitutes for those collected via statistical surveys, they often offer possibilities, sometimes through the combination of multiple sources, to replace, fully or partially, direct statistical data collection. The degree of the use of administrative sources in the statistical production process varies considerably from country to country, from those that have developed fully functioning register-based statistical systems, to those that are just starting to consider this approach. A significant contribution in this field is publication issued in 2011 by the United Nations Economic Commission for Europe[3], entitled "***Using Administrative and Secondary Sources for Official Statistic, A Handbook of Principles and Practices***". These trends make model-based procedures more and more attractive and relevant for NSIs to apply in the production of official statistics (Chambers et al., 2006).

Administrative datasets are typically very large, covering samples of individuals and time periods not normally financially or logistically achievable through survey or even census methodologies. Alongside cost savings, the scope of administrative data is often cited as its main advantage for research purposes, though coverage is recognized to be imperfect. The lack of control the researcher has during the data collection stage and how this affects its quality, and therefore what can be done with the data, are the main problems for administrative data. More general concern has also been voiced about the lack of well-established theory and methodologies to guide the use of administrative data in social science research.

Potential auxiliary data should be evaluated for their relationship to the variable(s) of interest, both theoretically and statistically as well as the accuracy and reliability with which they have been collected. The theoretical relationship should emanate from tested social or economic theories. A careful examination should be made to understand any major differences between the auxiliary data and the variables of interest.

Consideration should be given to the purpose for which the data were initially collected, how it was processed and edited, what conceptual definitions were used and what the scope of the auxiliary data holdings is. This will allow appropriate auxiliary information to be chosen to improve the model, and in explaining to users what factors are driving the small area estimates and help pinpoint potential sources of error.

Although auxiliary information was originally used in the design and estimation procedure of a survey to decrease the sampling variance of estimators, nowadays it is an important tool to decrease the bias due to selective non-response. Estimators using auxiliary information are generally more robust against selective non-response than estimators that do not use auxiliary information (Bethlehem, 1988; Särndal et al., 1987, 2005; Thomsen et al., 1998).

Common concern around the use of detailed administrative data at the small area level includes risks around confidentiality, anonymity and disclosure and this may lead to data controllers refusing to release the data or making it available within very controlled environments. An important consideration therefore for the release or publication of administrative data at individual or aggregate small area level is that the identity of individuals is protected. The assessment of disclosure risk is a complex process. Generally, the more detail the data has and the higher

[3] http://www.unece.org/fileadmin/DAM/stats/publications/Using_Administrative_Sources_Final_for_web.pdf.

the proportion of the population of interest that is captured in the data, the higher the risk.

There are various ways in which extracts of administrative data can be linked with other data sources to create more comprehensive and powerful datasets for analysis (both in terms of cases and variables).

Examples of available administrative data: i) population, ii) building and dwellings, iii) taxes, iv) business registers. Uses of administrative data are especially useful for:

1) improving survey results (sampling frame for persons and business surveys; auxiliary variables for calibration);
2) reducing the respondent burden: directly (some questions are skipped); indirectly (gain efficiency the estimators).

Administrative data holds great research potential for SAE (and other) research in all national contexts although the research availability and use of such data varies significantly between countries. There is also a problem how to use available BIG Data or other approaches for SAE in official statistics.

4. The international conferences and research projects towards application of SAE methods in official statistics

There have been different kind of conferences, seminars, and research projects devoted to exchange of ideas, experiences and achievements related to application of SAE in official statistics. First, some international conferences and next selected international research projects devoted to applications of SAE methods are briefly presented.

4.1. The International Conferences to apply SAE methods in official statistics

The results of the first attempts of applications of SAE methods in official statistics were presented at the symposium held ***in Ottawa in 1985 and published in Platek et al. (1987)***.

This publication had significant impact on academic statisticians and research statisticians in NSIs, and specially on countries in transition in Central and Eastern Europe, which organized international conferences held *in Poland in 1992 (Warsaw Conference in 1992: Kalton et al., 1993), and Latvia in 1999 (Riga conference: Riga, 1999).*

Starting from 2005, a new series of SAE Conferences have taken place in: *Finland, (Jyvaskyla, 2005); Italy, (Pisa, 2007); Spain, (Elche, 2009); Germany, (Trier, 2011); Thailand, (Bangkok, 2013); Poland, (Poznan, 2014), Chile, (Santiago de Chile, 2015).*

SAE conferences are aimed at providing a platform for discussion and exchange of ideas about current developments in small area estimation research in different fields. The conferences address - in a good balance with theoretical and methodological development in small area estimation and related fields, and in practical application - of SAE methods, including their potential uses in various research areas in official statistics. The need to regulate and promote the continuity of SAE conferences required the creation a working group with an acronym: EWORSAE – the European Working Group on Small Area Estimation[4] – aimed to build and maintain a network of researchers and statisticians to foster collaborative work and to increase cooperation between Statistical Offices and the research community within the field of SAE and related topics. Although the working group is basically European, it is open to all people worldwide working in small area estimation[5].

4.2. The International Projects for SAE implementations in official statistics

Before presenting some international projects for SAE methods applications in official statistics, it seems reasonable to begin with a program started in the USA over 20 years ago.

SAIPE – an acronym for ***Small Area Income and Poverty Estimates***[6]. The U.S. Census Bureau's program started at the beginning of 1990s and has provided annual estimates of income and poverty statistics for all states, counties, and school districts. The main objective of this program is to provide estimates of income and poverty for the administration of federal programs and the allocation of federal funds to local jurisdictions. In addition to these federal programs, state and local programs use the income and poverty estimates for distributing funds and managing programs. SAIPE revises and improves methodology as time and resources allow. The details of the methodology differ slightly from year to year. The most significant change was between 2004 and 2005, when SAIPE began using data from the *American Community Survey*, rather than from the *Annual Social and Economic Supplement to the Current Population Survey.*

Some impact on applications of SAE procedures in official statistics has had the following international projects sponsored by the European Union:

EURAREA; SAMPLE; BIAS; AMELI; ESSnet

4.2.1. The EURAREA project investigated methods for small area estimation and their application in official statistics. It was funded by Eurostat under the Fifth Framework (FP5) Programme of the European Union and was carried out by a consortium of NSIs, universities and research consultancies from across the European Union (United Kingdom, Spain, Italy, Sweden, Norway, Finland and Poland). The project was co-ordinated by the UK Office for National Statistics. It ran from January 2001 until June 2004 and was signed off by

[4] On initiative of Spanish Statisticians.
[5] http://sae.wzr.pl/.
[6] http://beta.census.gov/did/www/saipe/about/index.html.

Eurostat in February 2005. The aim of this project was to evaluate the effectiveness of standard estimation techniques for small areas (synthetic estimators, GREGs and composite estimators). The studies carried out until 2004 were based on sampling designs with equal selection probabilities. In order to undertake this project it was necessary to study the existing theory as well as to develop new theories that make it easier to obtain estimation techniques and their mean squared error when other sampling plans are used that are more similar to those applied in official statistics in the real world. Finally, all the theory developed has been implemented in a SAS IT application whose use has been widely documented so that any user is able to apply the programme to his/her own data. The links below provide further information about the project, its aims, objectives and conclusions[7].

The research outputs from the project are available in the download section: these include the final project reference volume and macro language programs written in SAS. The project reference volume contains reviews of existing theory in small area estimation, an assessment of the "standard" estimators and the results of the innovative work undertaken within the project. The program codes for the procedures investigated are provided so that the results can be implemented by other NSIs and statisticians. The program code has been written in SAS macro language or SAS macros or routines that can be called in SAS. Some results are also presented in EURAREA (2004), Heady et al. (2001, 2004) and Chambers et al. (2006).

4.2.2. SAMPLE: Small Area Methods for Poverty and Living Condition Estimates

The Project was supported by the European Commission (FP7-SSH-2007-1). The aim of SAMPLE project was to identify and develop new indicators and models for inequality and poverty with attention to social exclusion and deprivation, as well as to develop, implement models, measures and procedures for *small area estimation* of the traditional and new indicators and models[8]. This goal was achieved with the help of the local administrative databases. Local government agencies often had huge amount of administrative data to monitor some of the actions which witness situations of social exclusion and deprivation (social security claims for unemployment and eligibility for benefits from any of the programs Social Security administers) of households and citizens. SAMPLE utilised widely used indicators on monetary and non-monetary poverty. Moreover, in collaboration with stakeholders working with the poor, the project developed new poverty indicators that meet local needs. The results of the SAMPLE project will help local authorities and stakeholders to plan and implement their poverty-reduction policies. In fact, more than two thirds of

[7] http://www.ons.gov.uk/ons/guide-method/method-quality/general-methodology/spatial-analysis-and-modelling/eurarea/index.html.

[8] http://www.bing.com/search?ei=UTF-8&pc=AV01&q=http%3A%2F%2Fwww.sample-project.eu%2F&FROM=AVASDF.

stakeholders surveyed said that these local indicators would prove very useful in the planning of social policies. The final goal of the project is to provide a dashboard of reliable indicators of poverty and deprivation defined at NUTS3, NUTS4 level, useful for Local Government Agencies. In the project, the EU-SILC[9] sample will be enlarged at NUTS4[10].

The Project was coordinated by Prof. Monica Pratesi, Italy. Consortium of the project: Prof. Achille Lemmi, Italy; Dr Nikos Tzavidis, UK; Dr Isabel Molina, Spain ; Prof. Domingo Morales, Spain; Prof. Tomasz Panek, Poland; Dr Paolo Prosperini, Dr Claudio Rognini, Dr Moreno Toigo, Italy.

4.2.3. BIAS Project

The BIAS project is an acronym for "*Bayesian methods for combining multiple Individual and Aggregate data Sources in observational studies*" project. The first edition of the project named BIAS I was funded between April 2005 and June 2008 under the first phase of node commissioning. The second edition named BIAS II was funded by the second commissioning phase from July 2008 to June 2011. Description of BIAS I and BIAS II projects is based on information from the web page: www.bias-project.org.uk.

The aims of the project were: a) to develop a set of statistical frameworks for combining data from multiple sources, b) to improve the capacity of social science methods to handle the intricacies of observational data. In this project Bayesian hierarchical models are used as the basic building blocks for these developments. These offer a natural tool for linking together many different sub-models and data sources. The BIAS I research programme consisted of three methodological components: a) multiple bias modelling for observational studies, b) combining individual and aggregate level data, c) small area estimation.

The last one was especially devoted to small area estimation methodology and was being carried out in collaboration with ONS. The basic methodological problem was to estimate the value of a given indicator (e.g. income, crime rate, unemployment) for every small area, using data on the indicator from individual-level surveys in a partial sample of areas, plus relevant area-level covariates available for all areas from census and administrative sources, for example.

4.2.4. The AMELI Project

The project AMELI (*Advanced Methodology for Laeken Indicators*) was a trial to satisfy expectations of the need for effective, high-quality, robust, timely and reliable statistics and indicators related to the social cohesion. It started in April 2008 and ended in March 2011. The main target of the project was to review the state-of-the-art of the existing indicators monitoring the multidimensional phenomena of poverty and social exclusion - the Laeken indicators including their relation to social cohesion. Special emphasis was put on

[9] EU-SILC – an acronym for: European Union Statistics on Income and Living Conditions.
[10] NUTS – an acronym for: Nomenclature of Units for Territorial Statistics.

methodological aspects of indicators and especially on their impact on policy making. This included quality aspects as well as mathematical and statistical properties within a framework of a complex survey in the context of practical needs and peculiarities. The official website of the project is: http://ameli:surveystatistics:net/.

The coordinator of the Project was Prof. Ralf Münnich, University of Trier, Germany. Consortium consisted of: Federal Statistical Office of Germany, Swiss Federal Statistical Office, Statistics Austria, Statistics Finland, University of Helsinki, Vienna University of Technology, Statistical Office of Slovenia, Statistics Estonia.

4.2.5. The ESSnet Project for SAE

The project lasted 27 months (December 2009 to March 2012). The coordinator of the Project was: Stefano Falorsi, ISTAT, Italy. Co-partners: INSEE, France; DESTATIS, Germany; CBS, Netherlands; SSB, Norway; GUS, Poland; INE, Spain; ONS, United Kingdom; SFSO, Switzerland.

The general objective of the project was to develop a framework enabling the production of small area estimates for ESS social surveys.

The specific objectives were to: a) complete (thestate of the art level) the EURAREA project, b) update the documents available on small area estimation, c) describe the current application in UE NSIs and non-UE NSIs, d) create a common knowledge on application of small area estimation methods; e) review and develop suitable criteria to assess the quality of SAE methods for the choice of proper model and the evaluation of MSE; f) make available software tools for SAE to the ESS; g) foster knowledge transfer by the development of case studies and associated recommendations on representative problems in small area estimation in the ESS; h) provide practical guidelines in ESS social surveys context; i) transfer knowledge and know-how to non-participating NSIs and disseminate results.

Results of the project (lessons learned):

1) The work done and the outcomes produced by the project are strategic for increasing the capability within ESS to produce official statistics by SAE techniques.
2) The upload of all outcomes within the EU-cross-portal is very useful for disseminating scientific and applicative results of the ESSnet.
3) It should be useful to try to develop the regular exchange of information about SAE methods and applications among NSIs giving impulse to the use of forum within the website.
4) The course was very useful for involving the non-participating NSIs and transferring the results of the project within the ESS. It was also useful in order to map the real needs of non-participating countries.
5) The different presentations in scientific workshops and conferences were important to disseminate the knowledge of the outcomes of the project.

6) The survey on the use of methods and available tools within the NSIs of ESS and other NSIs has been very useful to map capabilities and application needs. This survey should be updated regularly and published into the website.

The results of the ESSnet project for SAE are strategic for increasing the capability to produce official statistics.

5. Discussion

Small area estimation methods have been developing significantly over the last 30 years and used partially in official statistics. Before small area estimates can be considered fully credible, carefully conducted evaluation studies are needed to check on the adequacy of the model being used. Sometimes model-dependent small area estimators turn out to be of superior quality to sample-based estimators, and this may make them seem attractive.

SAE techniques are becoming a matter of great interest for a variety of people, including statisticians, researchers and other university experts, and institutions, as NSIs, research institutes, governmental bodies, local authorities and private enterprises dealing with research methodology, empirical research and statistics production for regional areas and other population subgroups.

SAE methodologies have become a widely used method across various disciplines as a result of growing policy makers and researchers' demand for spatially detailed information alongside advances in small area data availability and computing power. Currently, despite the potential of these approaches and the growing demands placed upon them, there is little agreement within the academic and policy community as to which method(s) work best, whether different approaches are best suited to different local contexts, how best methods can be implemented and how best results can be validated. Experts from across each of these methodological strands and across a range of academic disciplines are included in the network so as to enable not only improvements in each separate approach but also overall methodological progress through the cross-pollination of ideas and skills.

Accuracy is generally considered to be a key measure of quality. Total survey error is a conceptual framework describing errors that can occur in a sample survey and the error properties. It may be used as a tool in the design of the survey, working with accuracy, other quality characteristics, and costs. Accuracy is often measured by the mean squared error (MSE) of the estimator. Error sources are considered one by one to estimate the uncertainty and also to obtain some indication of the importance of that source. The errors arise from: sampling, frame coverage, measurement, non-response, data processing, and model assumptions.

Therefore, indirect estimators are constructed that borrow strength from related areas, increasing the effective sample size and with it the estimation precision. These indirect estimators are based on either explicit or implicit models providing a link between the small area in question and related areas through ancillary information. These auxiliary variables can be miscellaneous, cross-sectional as well as across time, for example information from neighbouring or next higher populations, data from a previous census or administrative records. Due to the growing demand for reliable small area statistics, small area estimation is becoming an important field in survey sampling.

Weighting is a statistical technique commonly used and applied in practice to compensate for nonresponse and coverage error. It is also used to make weighted sample estimates conform to known population external totals. In recent years a lot of theoretical work has been done in the area of weighting and there has been a rise in the use of these methods in many statistical surveys conducted by NSIs around the world.

In the last decade, calibration has been used to reduce both sampling error and nonresponse bias in surveys. In the presence of auxiliary variables with known population totals or with known values on the originally sampled units, the calibration procedure generates final weights for observations that, when applied to those auxiliary variables, yield their population totals or unbiased estimates of these totals, respectively. Unfortunately, in practice availability of such of auxiliary variables is rather not often.

The move to a more overt modelling approach means that government agencies need to recruit and train statisticians who are adept in modelling methods, as well as being familiar with survey sampling design. Survey sampling needs to be considered a part of mainstream statistics, in which Bayesian models that incorporate complex design features play a central role. A Bayesian philosophy would improve statistical output, and provide a common philosophy for statisticians and researchers in substantive disciplines such as economics and demography. A strong research program within government statistical agencies, including cooperative ties with statistics departments in academic institutions, would also foster examination and development of the viewpoints (Lehtonen et al. 2002, Lehtonen and Sarndal 2009).

5.1. Results of international conferences and projects

It is difficult to assess the impact of the international conferences and different projects on application of SAE methods in official statistics. General conclusion is that development and results of SAE methods in official statistics obtained so far from these conferences and the international projects have been mostly academic. Several projects aimed at development of SAE methodology such as *EURAREA, SAMPLE, BIAS, AMELI*, etc. are either completed or still ongoing at a country level. Next to these methodologies-oriented projects, quite few projects focused on estimating variables for social surveys undertaken by some NSIs. What is more, methodological know-how and techniques in SAE differ in NSIs. Some of NSIs have a great deal to offer in terms of expertise, links with academic experts and experience of implementation of these techniques while some others are just at the empirical stage of practice. The first projects aiming at the development of SAE methodology did not highlight differences between European NSIs in the way they introduce the SAE methodology into the process of producing statistics. Only *ESSnet project for SAE* provides an overview of applications to social statistics for many European and some non-European NSIs. Furthermore, this project describes the research of the NSIs concerning SAE, which eventually will lead to a greater number of applications.

As it has been already stressed, the application of model-based estimation procedures in official statistics is limited. Several factors have been mentioned for the slow adoption of these methods. One is the fact that many NSIs are rather reserved in the application of model-based estimation procedures and generally rely on the more traditional design-based or model-assisted procedures for producing their official statistics. NSIs need to play safe in the production of official statistics and therefore do not want to relay on model assumptions, particularly if they are not verifiable (Chambers et al, 2006; Brakel & Betheyem, 2008; Eurarea, 2004; Little, 2004, 2012).

The availability of small area data has improved dramatically since the 1990s yet many spatial variables of interest – income, fear of crime, health-related behaviours, and so the list goes on – remain impossible to access at small area geographies in many national contexts. Within this context SAE methodologies have become increasingly demanded, increasingly used and increasingly refined. Yet the methodological landscape around SAE remains in need of attention in at least three key ways, according to Whitworth A. (edt)[11]. "Firstly, various alternative SAE methodologies have emerged and it is often unclear to some researchers what these alternative approaches are, how they relate to each other and how they compare in terms of their estimation performance. These methodological approaches can be classified broadly either as spatial microsimulation (which tend to be used by geographers predominantly) or statistical approaches (the use of which is dominated by statisticians). Secondly, despite recent advances in SAE methodologies there remain key methodological challenges and uncertainties to explore (e.g. how exactly each method can be best implemented in relation to weights, constraints, seeding, etc.) as well as innovative methodological advances to be brought together and extend (e.g. any role for agent-based modelling, estimating distributional functions or spatially varying interactions). Thirdly, the different methodological approaches to SAE in large part operate in parallel to one another without a clear understanding of the conceptual and methodological linkages between them. This is particularly true between the statistical and spatial microsimulation approaches and greater understanding of the linkages between methodologies within these two differing approaches could support important contributions to the effectiveness of current SAE best practice".

[11] http://eprints.ncrm.ac .uk/3210/1/sme_whitworth.pdf.

Nevertheless, SAE methods have been used in applications including *employment and unemployment statistics, health, poverty, agriculture, business, demography, census undercount, ecology, and education* (Datta et al.,1999, 2002; Dehnel, 2010; Dick, 1995; Drew et al., 1982; Elazar, 2004; Esteban et al., 2012; Gambino et al., 1998, 2000; Dehnel et al., 2004; Golata, 2004; Hidiroglou et al., 1985, 2007; Kordos, 1994, 2006; Kubacki, 2004; Molina et al., 2010; Paradysz, 1998; Paradysz & Dehnel, 2005; Schaible et al., 1994).

5.2. Differentiation in utilization of SAE methods by NSIs

As it has been stressed, NSIs are facing increasing demand for statistics below the level for which most large scale surveys have been designed. The survey methodologists are turning toward SAE techniques to satisfy the need for reliable estimates for small domains.

However, there are some common characteristics connected with applications of SAE procedures in official statistics. Usually such applications are prepared and implemented in cooperation with academic statisticians or subject-matter specialists and official statisticians. Very often there are still R&D approaches. It is impossible to discuss the differences by countries here, but the author confines himself to some issue connected with R&D in this field and quality aspects of the results. The author has found a number of very interesting publications in the Internet connected with applications of SAE methods in different fields and countries. Some of them include: Statistics Canada[12]; USA – Bureau of Census[13]; U.K- Office for National Statistics (ONS)[14] and Australian Bureau of Statistics[15]. The author would like to add the network, funded by the ESRC's National Centre for Research Methods (NCRM) Programme, which brings together experts in small area estimation techniques from the academic and policy (e.g. Office for National Statistics) communities in the UK and internationally in order to seek innovative ways to advance knowledge and understanding in SAE methodologies[16].

As it has already been stressed, it is impossible to discuss the differentiation of application of SAE procedures in different countries here, but the following issues will be considered: a) *Assessing the quality of small area estimates; b) Communicating quality to users.*

"*A Guide to Small Area Estimation*" published by the Australian Bureau of Statistics[17] has been mainly used here.

---

[12] http://www.bing.com/search?ei=UTF-8&pc=AV01&q=Small+area+estimation+in+Statistics+Canada&FROM=AVASDF&first=71&FORM=PORE.
http://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3604#a3

[13] www.census.gov/hhes/www/saipe/documentation.html.
http://www.census.gov/did/www/saipe/methods/10change.html.

[14] U.K. ONS: http://www.ons.gov.uk/ons/guide-method/method-quality/survey-methodology-bulletin/.

[15] http://www.nss.gov.au/nss/home.NSF/pages/Small+Areas+Estimates?OpenDocu.

[16] http://www.bing.com/search?ei=UTF-8&pc=AV01&q=Evaluations+and+improvements+in+small+area+estimation+methodologies++Adam+Whitworth+%28edt%29%2C+University+of+Sheffield&FROM=AVASDF.

[17] http://www.nss.gov.au/nss/home.NSF/pages/Small+Areas+Estimates?OpenDocu.

5.3. Assessing the quality of small area estimates

Small area estimates are usually obtained by fitting statistical models to survey data and then applying these models to auxiliary information available for the small area population of interest. Often a number of potential or candidate models are considered involving various combinations of the auxiliary variables.

The most reliable of these candidate models is then chosen as the final model, on the basis of:

- plausibility of the model in light of previous studies or accepted wisdom;
- how well the model fits the observed data; and,
- accuracy of the small area estimates predicted from the model.

In light of this, there is a need to examine various quality diagnostics to determine which of the candidate models to use. Having chosen a model, it is then necessary to provide users with an assessment of its quality as well as the quality of the small area estimates produced from it. In doing so, ranges of diagnostics are used to assess the accuracy, validity and consistency of the small area estimates.

These include:

- a bias test that compares the small area predictions with direct estimates;
- testing whether model assumptions are met and that the model is a good fit;
- checking that small area estimates add to published state or national estimates;
- local knowledge and expert advice on the spread of estimates across small areas; and,
- relative root mean squared errors (RMSE) - in modelling these are analogous to sampling errors calculated for survey estimates.

Although these diagnostics are crucial in terms of assessing the relative performance of competing small area models, they have to be supported by good judgement from practitioners and expert advice from users.

5.4. Communication with users on quality of accepted results

From current practice we may draw conclusions that there are problems with users' communication regarding quality of accepted results. There are several propositions to improve this practice, but it is suggested to consider the following Trewin's proposition.

Trewin (1999) encouraged NSIs to make greater use of small area estimation methods to generate statistical output. However, in doing so, he emphasised that:

a) *"the estimates need to be branded differently from other official statistics (the methods and the assumptions should be described in any releases);*

b) *their validity needs to be assessed to provide user confidence;*

c) the underlying models need to be described in terms that users can understand and the validity of the underlying assumptions should be discussed with the key users;

d) their quality should be described in quantitative terms as far as possible; and

e) there should be peer review of the models by an expert as the models are very complex and the choice of methods is considerable."

The author would like to add in this section the Eurostat publication (Eurostat, 2007) devoted to data quality assessment, presenting different methods and tools.

6. Concluding remarks

Small area estimation methods have been developing significantly over the last 30 years and used partially in official statistics. Before small area estimates can be considered fully credible, carefully conducted evaluation studies are needed to check on the adequacy of the model being used. Sometimes model-dependent small area estimators turn out to be of superior quality to sample-based estimators, and this may make them seem attractive.

It seems reasonably to give some recommendations and suggestions compiled from different papers, conferences and projects related to SAE methods:

1. Good auxiliary information related to the variables of interest plays a vital role in model-based estimation. Expanded access to auxiliary data, such as census and administrative data, through coordination and cooperation among federal agencies is needed.
2. Preventive measures at the design stage may reduce the need for indirect estimators significantly.
3. Model selection and checking plays an important role. External evaluations are also desirable whenever possible.
4. Area-level models have wider scope because area-level data are more readily available. But assumption of known sampling variance is restrictive.
5. HB approach is powerful and can handle complex modelling, but caution should be exercised in the choice of priors on model parameters. Practical issues in implementing HB paradigm should be addressed.
6. Model-based estimates of area totals and means are not suitable if the objective is to identify areas with extreme population values or to identify areas that fall below or above some pre-specified level.
7. Suitable benchmarking is desirable.
8. Model-based estimates should be distinguished clearly from direct estimates. Errors in small area estimates may be more transparent to users than errors in large area estimates.

9. Proper criterion for assessing quality of model-based estimates is whether they are sufficiently accurate for the intended uses. Even if they are better than direct estimates, they may not be sufficiently accurate to be acceptable.
10. Overall program should be developed that covers issues related to sample design and data development, organization and dissemination, in addition to those pertaining to methods of estimation for small areas.

Acknowledgements

The author would like to thank his colleagues from LinkedIn: J. Bethlehem (Netherlands), J. Cochran (USA), A. Fuller (USA), D. Hedlin (Sweden), D. Kasprzyk (USA), S. Laaksonen (Finland), L. Lyberg (Sweden), D. Marker (WESTAT), Siu-Ming Tam (Australia), A. G. Turner (USA), D. Trewin (Australia), who have helped him with preparing the session "*Small Area Estimation in Official Statistics and Statistical Thinking*" held at the International Conference on *Small Area Estimation 2014,* in Poznan, Poland, 3-5 September 2014.

The author is also grateful to the anonymous referee whose comments have helped to improve the quality of this paper.

REFERENCES

AUSTRALIAN BUREAU OF STATISTICS, (2006). A Guide to Small Area Estimation – Version 1.1. Internal ABS document. available online at: http://www.nss.gov.au/nss/home.NSF/pages/Small+Areas+Estimates?OpenDocu.

BATTESE, G. E., HARTER, R. M., FULLER, W. A., (1988). "An Error-Components Model for Prediction of County Crop Areas Using Survey and Satellite Data," *Journal of the American Statistical Association*, 83, 28–36.

BEST, N., RICHARDSON, S., THOMSON, A., (2005). A comparison of Bayesian spatial models for disease mapping. *Statistical Methods in Medical Research*, 14, 35–39.

BETHLEHEM, J. G., (1988). Reduction of nonresponse bias through regression estimation. *Journal of Official Statistics*, 4, 251–260.

BETHLEHEM, J. G., KERSTEN, H. M. P., (1985). On the treatment of nonresponse in sample surveys. *Journal of Official Statistics*, 1, 287–300.

BRACHA, CZ., LEDNICKI, B., WIECZORKOWSKI, R., (2003). Estimation of Data from the Polish Labour Force Surveys by poviats (counties) in 1995–2002 (in Polish), Central Statistical Office of Poland, Warsaw.

BRACKSTONE, G. J., (2002). Strategies and Approaches for Small Area Statistics, *Survey Methodology*, 28(2), 117–123.

BRACKSTONE, H., (1999). Managing Data Quality in a Statistical Agency. *Survey Methodology*, 25, 2, 129–149.

BRAKEL, J. A. VAN DEN, BETHLEHEM, J., (2008). Model-Based Estimation for Official Statistics, Statistics Netherlands, Voorburg/Heerlen.

CHAMBERS, R., BRAKEL, J. A. VAN DEN, HEDLIN, D., LEHTONEN, R., ZHANG, LI-CHUN, (2006). Future Challenges of Small Area Estimation. *Statistics in Transition*, 7, 759–769.

CHOUDRY, G. H., RAO, J. N. K., (1989). Small area estimation using models that combine time series and cross sectional data, in: Singh, A.C., Whitridge, P. (Eds.*), Proceedings of Statistics Canada Symposium on Analysis of Data in Time*, pp. 67–74.

CHOUDHRY, G. H., RAO, J. N. K., HIDIROGLOU, M. A., (2012). On sample allocation for efficient domain estimation. *Survey Methodology*, 38, 23–29.

COCHRAN, W. G., (1977). *Sampling Techniques*, 3rd ed. New York: Wiley.

DATTA, G., (2009). Model-based approach to small area estimation. Chapter 32 in Rao C.R. and Pfeffermann D. (Eds.). Handbook of Statistics. Sample Surveys: Inference and Analysis. Vol. 29B. New York: Elsevier. (251-288).

DATTA, G. S., LAHIRI, P., MAITI, T., LU, K. L., (1999). Hierarchical Bayes estimation of unemployment rates for the U.S. states. *Journal of the American Statistical Association*, 94, 1074–1082.

DATTA, G. S., LAHIRI, P., MAITI, T., (2002). Empirical Bayes estimation of median income of four person families by state using time series and cross-sectional data. *Journal of Statistical Planning and Inferenc*e, 102, 83–97.

DEHNEL, G., (2010). *The development of micro-entrepreneurship in Poland in the light of the estimation for small areas*, Publ. University of Economics in Poznan, Poznan (in Polish).

DEHNEL, G., GOLATA, E., KLIMANEK, T., (2004). Consideration on Optimal Design for Small Area Estimation, Statistics in Transition, vol. 6, Nr 5, pp. 725–754.

DEVILLE, J., SÄRNAL, C.-E., (1992). Calibration Estimators in Survey Sampling, *Journal of the American Statistical Association*, 87, 376–382.

DICK, P., (1995). Modelling net undercoverage in the 1991 Canadian census. *Survey Methodology*, 21, 45-54.

DREW, D., SINGH, M. P., CHOUDHRY, G. H., (1982). Evaluation of Small Area Estimation Techniques for the Canadian Labour Force Survey, *Survey Methodology*, 8, 17–47.

ELAZAR, D., (2004), Small Area Estimation of Disability in Australia, *Statistics in Transition*, 6, 5, 667–684.

ESTEBAN, M. D., MORALES, D., PEREZ, A., SANTAMARIA, L., (2012). Small area estimation of poverty proportions under area-level time models. *Computational Statistics and Data Analysis*, 56, 2840–2855.

EURAREA, (2004). Project reference volume, deliverable D7.1.4, *Technical report, EURAREA consortium.*

EUROSTAT, (2007). *Handbook on Data Quality Assessment: Methods and Tools*, Luxembourg.

FAY, R. E., HERRIOT, R. A., (1979). Estimation of income from small places: An application of James-Stein procedures to census data. *Journal of the American Statistical Association*, 74, 269–277.

FULLER, W. A., (1999). Environmental surveys over time. *Journal of the Agricultural, Biological and Environmental Statistics*, 4, 331–345.

FULLER, W. A., (2002). Regression estimation for survey samples. *Survey Methodology*, 28, 5–23.

GAMBINO, J. G., Dick, P., (2000). Small Area Estimation Practice at Statistics Canada, *Statistics in Transition,* 4, 4, 597–610.

GAMBINO, J. G., SINGH, M. P., DUFOUR, J., KENNEDY, B., LINDEYER, J., (1998). *Methodology of the Canadian Labour Force Survey*, Statistics Canada.

GOLATA, E., (2004). Problems of Estimate Unemployment for Small Domains in Poland**,** *Statistics in Transition,* 6, 5. 755–776.

GOSH, M., MEEDEN, G., (1997). Bayesian Methods for Finite Population Sampling . London: Chapman & Hall.

GHOSH, M., NATARAJAN, K., WALTER, L. A., KIM, D. H., (1999). Hierarchical Bayes GLMs for the analysis of spatial data: An application to disease mapping. *Journal of Statistical Planning and Inference*, 75, 305–318.

GHOSH, M., RAO, J. N. K., (1994). Small Area Estimation: An Appraisal, *Statistical Science*, 9, 55–93.

GONZALEZ-MANTEIGA, W., LOMBARDIA, M. J., MOLINA, I., MORALES, D., SANTAMARIA, L., (2010). Small area estimation under Fay-Herriot models with nonparametric estimation of heteroscedasticity. *Statistical Modelling*, 10, 2, 215–239.

HANSEN, M. H., HURWITZ, W. N., MADOW. W. G., (1953). *Sample Survey Methods and Theory*, Vol. I and II. New York: Wiley.

HEADY, P., HENNELL, S., (2001). Enhancing Small Area Estimation Techniques to Meet European Needs, *Statistics in Transition*, 5, 2, 195–203.

HEADY, P., RALPHS, M., (2004). Some Findings of the EURAREA Project – and their Implications for Statistical Policy, *Statistics in Transition*, 6, 5, 641–654.

HIDIROGLOU, M. A., (2014). Small-Area Estimation: Theory and Practice, Section on Survey Research Methods, Statistics Canada.

HIDIROGLOU, M. A., SINGH A., HAMEL M., (2007). Some Thoughts on Small Area Estimation for the Canadian Community Health Survey (CCHS). Internal Statistics Canada document.

HIDIROGLOU, M. A., SÄRNDAL, C. E., (1985). Small Domain Estimation: A Conditional Analysis, *Proceedings of the Social Statistics Section, American Statistical Association*, 147–158.

HIDIROGLOU, M. A., PATAK, Z., (2004). Domain estimation using linear regression. *Survey Methodology*, 30, 67–78.

HOLT, D., ELLIOT, D., (1991). Methods of weighting for unit non-response. *The Statistician*, 40, 333–342.

HORVITZ, D. G., THOMPSON, D. J., (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association*, 47, 663–685.

JIANG, J., LAHIRI, P., (2006). Mixed model prediction and small area estimation. *Test,* 15, 1–96.

KALTON, G., (2002). Models in the Practice of Survey Sampling (Revisited), *Journal of Official Statistics,* 18, 129–154.

KALTON, G., KASPRZYK, D., (1986) The treatment of missing data. *Survey Methodology*, 12, 1–16.

KALTON, G., KORDOS, J., PLATEK, R., (1993). *Small Area Statistics and Survey Designs,* Vol. I: Invited Papers; Vol. II: Contributed Papers and Panel Discussion. Central Statistical Office, Warsaw.

KISH, L., (1965). *Survey Sampling*. New York: Wiley.

KORDOS, J., (1994). Small Area Statistics in Poland (Historical Review). *Statistics in Transition*, 1, 6,783–796.

KORDOS, J., (2005). Some Aspects of Small Area Statistics and Data Quality, *Statistics in Transition*, 7, 1, 63–83.

KORDOS, J., (2006). Impact of different factors on research in small area estimation in Poland**,** „*Statistics in Transition”*, 7, 4, 863–879.

KORDOS, J., PARADYSZ, J., (2000). Some Experiments in Small Area Estimation in Poland, *Statistics in Transition,* 4, 4, 679–697.

KUBACKI, J., (2004). Application of the Hierarchical Bayes Estimation to the Polish Labour Force Survey, *Statistics in Transition,* 6, 5, 785–796.

LEHTONEN, R., SÄRNDAL, C. E., VEIJANEN, A., (2003). The effect of model choice in estimation for domains, including small domains. *Survey Methodology*, 29, 33–44.

LEHTONEN, R., SÄRNDAL, C. E., VEIJANEN, A., (2005). Does the model matter? Comparing model-assisted and model-dependent estimators of class frequencies for domains. *Statistics in Transition*, **7**, 649–673.

LEHTONEN, R., VEIJANEN, A., (2009). Design-based methods of estimation for domains and small areas. Chapter 31 in Rao C.R. and Pfeffermann D. (Eds.). Handbook of Statistics. Sample Surveys: Inference and Analysis. Vol. 29B. New York: Elsevier, 219–249.

LEHTONEN, R., SÄRNDAL, C. E., VEIJANEN, A., (2009). Model calibration and generalized regression estimation for domains and small areas. Invited paper, SAE2009 Conference on Small area estimation, 29 June – 1 July, 2009, Elche, Spain.

LITTLE, R. J. A., (2004). To Model or Not to Model? Competing Modes of Inference for Finite Population Sampling, *Journal of the Royal Statistical Association*, 99, 546–556.

LITTLE, R. J. A., (2012). Calibrated Bayes, an Alternative Inferential Paradigm for Official Statistics *Journal of Official Statistics*, 28, 3, 309–334.

LONGFORD, N., (2005). *Missing Data and Small-Area Estimation: Modern Analytical Equipment for the Survey Statistician*, Springer.

LUNDSTRÖM, S., SÄRNDAL, C. E., (1999). Calibration as a standard method for treatment of nonresponse. *Journal of Official Statistics*, 15, 305–327.

MARHUENDA, Y. MOLINA, I., MORALES, D., (2013). Small area estimation with spatio-temporal Fay-Herriot models. *Computational Statistics and Data Analysis*, 58, 308–325.

MARKER, D. A., (2001). Producing small area estimates from national surveys: Methods for minimizing use of indirect estimators. *Survey Methodology*, 27, 183–188.

MARSHALL, A., (updated by V. Higgins, 2013). *Small area estimation using key UK surveys – An Introductory guide.* UK Data Service, University of Essex and University of Manchester.

MOLINA, I., RAO, J. N. K., (2010). Small area estimation of poverty indicators**,** *The Canadian Journal of Statistics*, 38, 3, 369–385.

MOLINA, I., SALVATI, N., PRATESI, M., (2009). Bootstrap for estimating the MSE of the Spatial EBLUP. *Computational Statistics*, 24, 441–458.

MOURA, F. A. S., MIGON, H. S., (2002). Bayesian spatial models for small area proportions. *Statistical Modelling*, 2, 3, 183–201.

NARAIN, R., (1951). On sampling without replacement with varying probabilities. *Journal of the Indian Society of Agricultural Statistics*, 3, 169–174.

OPSOMER, J. D., CLAESKENS, G., RANALLI, M. G., KAUERMANN, G., BREIDT, F. J., (2008). Nonparametric small area estimation using penalized spline regression. *Journal of the Royal Statistical Society*, B70, 265–286.

PARADYSZ, J., (1998). Small Area Statistics in Poland - First Experiences and Application Possibilities, *Statistics in Transition*, 3, 5, 1003–1015.

PARADYSZ, J., DEHNEL, G., (2005). Attempts to Estimate Basic Information for Small Business in Poland, SAE2005 Conference, Jyväskylä, Finland, 28–31 August 2005.

PLATEK, R., RAO, J. N. K. SÄRNDAL, C. E., SINGH, M. P. (Eds.), (1987). *Small Area Statistics*, John Wiley & Sons, New York.

PFEFFERMANN, D., (1999). Small area estimation – big developments. Proceedings of IASS Conference on Small Area Estimation, Riga. Latvian Council of Sciences, 129–145.

PFEFFERMANN, D., (2002). Small area estimation – new developments and directions. *International Statistical Review* 70, 125–143.

PFEFFERMANN, D., (2013). New important developments in small area estimation, *Statistical Science*, 28, 1, 40–68.

PFEFFERMANN, D., BURCK, L., (1990). Robust small area estimation combining time series and cross-sectional data. *Survey Methodology*, 16, 217–237.

PFEFFERMANN, D., Tiller, R., (2006). Small Area Estimation with State Space Models Subject to Benchmark Constraints. *Journal of the American Statistical Association*, 101, 1387–1897

PRASAD, N. G. N., RAO, J. N. K., (1990). "The Estimation of Mean Squared Error of Small-Area Estimators", *Journal of the American Statistical Association*, 85, 163–171.

PRASAD, N. G. N., RAO, J. N. K, (1999). On robust small area estimation using a simple random effects model. *Survey Methodology*, 25, 67–72.

PRATESI, M., SALVATI, N., (2008). Small area estimation: the EBLUP estimator based on spatially correlated random area effects. *Statistical Methods and Applications*, 17, 113–141.

RAO, J. N. K., (1999). Some recent advances in model-based small area estimation, *Survey Methodology*, 25, 2, 175–186.

RAO, J. N. K., (2003). *Small Area Estimation*, John Wiley & Sons, New Jersey.

RAO, J. N. K., (2011). Impact of Frequentist and Bayesian Methods on Survey Sampling Practice: A Selective Appraisal, *Statistical Science*, 26, 2, 240–256.

RAO, J. N. K., YU, M., (1994). Small area estimation by combining time series and cross sectional data. *Canadian Journal of Statistics*, 22, 511–528.

RIGA, (1999). *Small Area Estimation–Conference Proceedings*, Riga, Latvia, August 1999.

SÄRNDAL, C. E., SWENSSON, B., WRETMAN, J., (1992). *Model Assisted Survey Sampling*. New York: Springer-Verlag.

SÄRNDAL, C.-E., LUNDSTRÖM, S., (2005). References, in Estimation in Surveys with Nonresponse, John Wiley & Sons.

SCHAIBLE, W. A., (1978). Choosing Weights for Composite Estimators for Small Area Statistics, *Proceedings of the Section on Survey Research Methods, American Statistical Association*, 741–746.

SCHAIBLE, W. L., CASADY, R. J., (1994). The Development, Application, and Evaluation of Small Area Estimators. *Statistics in Transition*, 1, 6, 727–746.

SINGH, M. P., GAMBINO, J., MANTEL, H. J., (1994). Issues and Strategies for Small Area Data, *Survey Methodology*, 20, 3–22.

SINGH, B., SHUKLA, G., KUNDU, D., (2005). Spatio-temporal models in small area estimation. *Survey Methodology, 31*, 183–195.

SZARKOWSKI, A., WITKOWSKI, J., (1994). The Polish Labour Force Survey, *Statistics in Transition*, 1, 4, 467–483.

THOMSEN, I., HOLMOY, A. M. K., (1998). Combining data from surveys and administrative record systems. The Norwegian experience, *Inter. Satist. Rev*., 66, 201–221.

TREWIN, D., (1999). Small Area Statistics Conference, *Survey Statistician*, 41, 8–9.

TREWIN, D., (2002). The importance of a Quality Culture, *Survey Methodology*, 28, 2, 125–133.

UN, (2011). Using Administrative and Secondary Sources for Official Statistics – A Handbook of Principles and Practices, United Nations Commission for Europe.

VALLIANT, R., DORFMAN, A. H., ROYALL, R. M., (2000). *Finite Population Sampling and Inference, A Prediction Approach*. New York: Wiley.

WANG, J., FULLER, W. A., (2003). The mean square error of small area predictors constructed with estimated area variances. *Journal of the American Statistical Association*, 98, 463, 716–723.

WANG, J., FULLER, W. A., QU, Y., (2008). Small area estimation under a restriction. *Survey Methodology*, 34, 1, 29–36.

WOODRUFF, R. S., (1966). Use of a Regression Technique to Produce Area Breakdowns of the Monthly National Estimates of Retail Trade, *Journal of the American Statistical Association*, 61, 496–504.

WYWIAŁ, J., (2000). On precision of Horvitz–Thompson strategies, *Statistics in Transition*, 4. 5, 779–798.

YATES, F., (1981). *Sampling methods for censuses and surveys.* 4th Ed., Charles Griffin and Co., London, U.K.

YOU, Y., (2008). An integrated modelling approach to unemployment rate estimation for sub-provincial areas of Canada. *Survey Methodology*, 34, 19–27.

YOU, Y., CHAPMAN, B., (2006). Small area estimation using area level models and estimated sampling variances. *Survey Methodology*, 32, 97–103.

YOU, Y., RAO, J. N. K., (2002). A pseudo empirical best linear unbiased prediction approach to small area estimation using survey weights. *The Canadian Journal of Statistics*, 30, 431–439.

YOU, Y., RAO, J. N. K., GAMBINO, J. G., (2003). Model-based unemployment rate estimation for the Canadian Labour Force Survey: A hierarchical Bayes approach, *Survey Methodology*, 29, 25–32.

YOU, Y, DICK, P., (2004). Hierarchical Bayes Small Area Inference to the 2001 Census Undercoverage Estimation. Proceedings of the ASA Section on Government Statistics, 1836–1840.

YOU, Y., ZHOU, Q. M., (2011). Hierarchical Bayes small area estimation under a spatial model with application to health survey data, *Survey Methodology*, 37, 1, 25–37.

WHITWORTH, A., (edt). Evaluations and improvements in small area estimation methodologies, University of Sheffield: http://eprints.ncrm.ac .uk/3210/1/sme_whitworth.pdf.