STATISTICS IN TRANSITION new series, Summer 2014
Vol. 15, No. 3, pp. 467—476

REPORT

International Conference on Small Area Estimation (SAE 2014)

The international conference on Small Area Estimation (SAE 2014) was held from 3rd to 5th September 2014 in Poznan, and was devoted to the methodology of small area statistics, which, following arrangements made by the European Working Group on Small Area Estimation, was organized by the Department of Statistics at the Poznan University of Economics (PUE). The conference was co-organized by the Central Statistical Office (CSO) in Warsaw and the Statistical Office in Poznan. Professor Janusz Witkowski, President of CSO and Professor Marian Gorynia, Rector of the Poznan University of Economics took the Honorary Patronage over the conference. The conference was preceded by a special workshop devoted to using R in SAE conducted by Li-Chun Zhang from University of Southampton and Statistics Norway.

The conference was partially financed by R Revolution Analytics with the support of the National Bank of Poland granted under the program of economic education.

The Chairperson of the Organizing Committee of the Conference SAE 2014 was Marcin Szymkowiak from the Department of Statistics, PUE. Other members of the committee included: Wojciech Adamczewski and Katarzyna Cichońska from the Central Statistical Office in Warsaw, Tomasz Józefowski, Tomasz Klimanek and Jacek Kowalewski from the Statistical Office in Poznan. The Programme Committee of the conference was headed by Professor Domingo Morales (Universidad Miguel Hernández de Elche). Other members of the programme committee included Professors: Ray Chambers (University of Wollongong), Grażyna Dehnel (University of Economics in Poznan), Elżbieta Gołata (University of Economics in Poznan), Malay Gosh (University of Florida), Jan Kordos (CSO), Partha Lahiri (University of Maryland), Risto Lehtonen (University of Helsinki), Isabela Molina (Universidad Carlos III de Madrid), Ralf Münnich (University of Trier), Jan Paradysz (University of Economics in Poznan), Danny Pfeffermann (Hebrew University of Jerusalem), J.N.K. Rao (Carleton University) and Li-Chun Zhang (University of Southampton). The Steering Board of the SAE 2014 Conference was chaired by Domingo Morales and included Ray Chambers, Elżbieta Gołata, Partha Lahiri and Danny Pfeffermann.

The idea behind the SAE 2014 conference was to provide a platform for the exchange of ideas and experiences between statisticians, scientists and experts from universities, statistical institutes, research centers as well as other governmental agencies, local government and private companies involved in developing and applying the methodology of regional surveys, in particular, small area estimation. The Poznan conference was another one in the series of conferences (Jyväskylä 2005, Piza 2007, Elche 2009, Trier 2011), intended to combine theoretical considerations and practical applications of SAE in public statistics.

The SAE 2014 conference was focused on applications of SAE in censuses, model-based estimation and its evaluation, the use of spatio-temporal models, robust methods, non-response, issues in sample selection, poverty estimation, teaching SAE and its applications in public statistics. The conference featured a discussion panel and a specialist workshop devoted to the theory and practice of indirect estimation methodology.

The SAE 2014 conference was attended by 140 researchers and practitioners from 22 countries (Albania 2, Australia 6, China 1, Czech Republic 1, Finland 3, Spain 9, The Netherlands 3, Israel 1, Japan 2, Canada 2, Kuwait 1, Lithuania 3, Luxembourg 1, Malta 1, Germany 9, Norway 1, New Zealand 1, Poland 64, Thailand 1, Turkey 1, USA 15, United Kingdom 6, Italy 6). The participants represented both domestic and foreign research centers. The SAE 2014 conference was attended by scientists from 46 universities from around the world, representatives of statistical offices from nearly 30 countries, as well as representatives of international scientific organizations, the World Bank, the Central Statistical Office and statistical offices from Poland.

The conference brought together some eminent experts in this field of statistics including: J.N.K. Rao, Malay Ghosh, Ray Chambers, Li-Chun Zhang, Partha Lahiri, Danny Pfeffermann, Risto Lehtonen, Ralf Münnich, Domingo Morales, Graham Kalton (Westat), Wayne Fuller (Iowa State University) and Isabel Molina.

The conference featured two plenary sessions and special lectures given by Professors J.N.K. Rao and Malay Gosh. During the lecture entitled **Inferential Issues in Model-Based Small Area Estimation: Some New Developments** Professor J.N.K. Rao discussed developments in the theory and applications of small area estimation which have taken place especially over the past 15 years in response to a growing demand for reliable small area statistics. In particular, the lecture addressed some recent important developments concerning area level and unit level models, mainly addressing issues related to assumed models. During his talk Professor Rao specifically focused on the bootstrap methods for mean squared error (MSE) estimation and confidence interval construction. Other subjects raised during the talk included recent work on robust estimation of small area means, informative sampling, new developments in model selection and

checking, methods for the estimation of complex parameters such as small area poverty measures and the role of 'big data' in small area estimation.

Professor Malay Ghosh's lecture, entitled **Benchmarked Empirical Bayes Estimators for Multiplicative Area Level Models,** was devoted to empirical Bayes and benchmarked empirical Bayes estimators of positive small area means under multiplicative models. In his presentation, he discussed the transformed Fay-Herriot model as a multiplicative model for estimating positive small area means and a weighted Kullback-Leibler divergence as a loss function. Professor Malay Ghosh demonstrated that the resulting Bayes estimator is the posterior mean and that the corresponding benchmarked Bayes and empirical Bayes estimators retain the positivity constraint. The prediction errors of the suggested empirical Bayes estimators were investigated asymptotically, and their second-order unbiased estimators were provided. In addition, bootstrapped estimators of these prediction errors were also provided. The performance of the considered procedures was investigated by the author through simulation and in an empirical study.

The SAE 2014 conference included ten invited sessions organized by top specialists:

- SAE: robust and nonparametric methods (Professor Ray Chambers, University of Wollongong),
- Small Area Methods for Repeated Survey (Professor Partha Lahiri, University of Maryland),
- SAE in poverty mapping (Professor Isabela Molina, Universidad Carlos III de Madrid),
- SAE models: selection and checking (Professor Danny Pfeffermann, Hebrew University of Jerusalem),
- SAE in official statistics (Professor Jan Kordos, CSO),
- Teaching SAE (Professor Risto Lehtonen, University of Helsinki),
- SAE applications (Professor Ralf Münnich, University of Trier),
- Benchmarking, design issues and nonresponse in SAE (Professor Stefano Falorsi, ISTAT),
- Population Census and SAE (Professor Li-Chun Zhang, University of Southampton),
- Other topics related to SAE (Professor Domingo Morales, Universidad Miguel Hernández de Elche).

During the session **SAE: robust and nonparametric methods** organized by Ray Chambers, and chaired by Graham Kalton, four papers were presented:

- Raymond Chambers − Two Recent Developments in Robust and Semiparametric Small Area Estimation,
- Beate Weidenhammer, Nikos Tzavidis, Timo Schmid, Nicola Salvati − Domain Prediction for Counts using Microsimulation via Quantiles,

- Payam Mokhtarian − On Outlier Robust Small Area Prediction of the Empirical Distribution Function,
- Forough Karlberg − Small Area Prediction for Skewed Data in the Presence of Zeroes.

In particular, Ray Chambers presented new developments in modeling for small area estimation including the spatial extension of recently published results on robust bias correction when asymmetric unit level and area level outliers in the survey data are used to predict a small area mean. Professor Chambers specifically focused on the extension of M-quantile modelling for small area estimation for count data rather than realizations of continuously distributed variables. Three other presentations were devoted to different aspects of robust and nonparametric methods in small area estimation and included the problem of estimation when outliers occur, the problem of estimation for asymmetric distribution with zeroes and using quantiles for the purpose of prediction of counts in the context of microsimulation.

The invited session on **Small Area Methods for Repeated Survey,** organized by Partha Lahiri and chaired by Wayne Fuller, consisted of four presentations:

- Partha Lahiri − An Overview of Small Area Estimation with Repeated Survey Data,
- Jan A. van den Brakel, Sabine Krieg − Small area estimation with state-space common factor models for rotating panels,
- Enrico Fabrizi, Maria Rosaria Ferrante, Carlo Trivisano − Estimation of value added for firms cross-classified by region, industry and size using repeated survey data,
- Carolina Franco, William R. Bell − Alternative Approaches to Borrowing Information Over Time in Small Area Estimation with Application to Data from the Census Bureau's American Community Survey.

This session provided an overview of different small area estimation methods for repeated surveys. In particular, the main presentation, given by Partha Lahiri, highlighted the fact that repeated surveys not only offer opportunities for improving small area statistics that are usually produced in cross-sectional surveys, but they may also deliver reliable estimates of changes over time, which may be more important than estimating current time. Professor Lahiri also pointed out that repeated surveys could conveniently help statisticians explain the benefits of small area statistics to public policy makers. Three other presentations addressed different aspects of small area estimation methods for repeated surveys and covered such issues as modeling for rotating panels, estimation of value added in business statistics and the problem of borrowing strength over time.

The third invited session on **SAE in poverty mapping**, organized by Isabel Molinaand chaired by Monica Pratesi, was devoted to issues connected with

poverty in the context of small area estimation methodology. This session consisted of four presentations:

- Isabel Molina, J.N.K. Rao − An overview of small area estimation methods for poverty mapping,
- Gauri Datta, Abhyuday Mandal − Small area estimation with uncertain random effects,
- Domingo Morales − Partitioned area-level time models for estimating poverty indicators,
- Roy Van der Weide − Estimation of Normal Mixtures in a Nested Error Model with an Application to Small Area Estimation of Poverty and Inequality.

The main purpose of this session was to show how different techniques offered by small area estimation can be used in the field of poverty. This is especially very important for many institutions which have to conduct more effective and efficient policy at the regional level. During this session the main approaches for small area estimation techniques for poverty mapping were reviewed and their advantages and disadvantages were discussed. In particular, special attention was given to recent variants of the basic methods in the field of poverty mapping and inequality.

The invited session on **SAE models: selection and checking**, organized by Danny Pfeffermann and chaired by Elżbieta Gołata, consisted of four presentations:

- Danny Pfeffermann − Model Selection and Checking for Small Area Estimation, Graham Kalton – discussant,
- Jay Breidt, Daniel Hernandez-Stumpfhauser, Jean D. Opsomer − Variational Approximations for Selecting Hierarchical Models of Circular Data in a Small Area Estimation Application,
- Jiraphan Suntornchost, Partha Lahiri − Variable selection for Linear Mixed Models with Applications in Small Area Estimation,
- Yahia El Horbaty − A Simple Score Test for Random Effects with Application to Small Area Models.

The main aim of this session was to present recent developments in the field of modeling in small area estimation methodology. In his paper, Professor Danny Pfeffermann gave an overview of some methods proposed in the literature for small area model selection and checking, distinguishing between frequentist methods and Bayesian methods. He also discussed some issues related to the theoretical foundation of small area estimation models and in particular, the interpretation and role of the random effects. Three other presentations in this session were devoted to practical aspects of using proper chosen models in different surveys.

The invited session on **SAE in official statistics**, organized by Jan Kordos and chaired by William Bell, featured four presentations:

- Jan Kordos − Small area estimation in official statistics and statistical thinking,
- Danute Krapavickaite, Tomas Rudys − Application of small area estimation methods for Lithuanian Labour force survey data,
- Jan Paradysz, Karolina Paradysz − Indirect estimation of disability on the base of Polish National Census 2011,
- Jan Kubacki, Alina Jędrzejczak − Small area estimation under spatial SAR model.

This session was a response to the growing role of small area estimation methodology in official statistics. In the main presentation, Professor Jan Kordos outlined the general mission of national statistics institutes to produce high quality statistical information on the state and evolution of the population, the economy, the society and the environment. Professor Kordos paid special attention to the so called statistical thinking in the context of small area statistics and Total Quality Management. He also presented selected applications of Small Area Estimation procedures in official statistics in the context of an increasing demand for information. Other presentations in this session were more practical and focused on applications of SAE methodology, with particular emphasis on issues related to labour market and disability.

The invited session on **Teaching SAE**, organized by Risto Lehtonen and chaired by Gauri Datta, was devoted to different aspects of teaching small area estimation methods and consisted of four presentations:

- Risto Lehtonen − Experiences and challenges in teaching small area estimation,
- Jan Pablo Burgard, Ralf Münnich − SAE teaching using simulations,
- Elżbieta Gołata, Tomasz Klimanek − Challenges faced by academics and the NSI in SAE education,
- Esther Lopez Vizcaino Lombardía Cortiña, M. José, Domingo Morales − mme: An R package for small area estimation with multinomial mixed models.

This session was a response to the problems and issues related to the basic and fundamental question of how to teach small area estimation methodology at universities and within statistical offices. In the main presentation, Professor Risto Lehtonen argued that SAE teaching should be treated as one of the main components of the ‘ecosystem’, which consists of scientific conferences devoted to SAE, textbooks related to SAE, SAE chapters in edited books and hundreds of journal articles, active research groups, large-scale international research projects and programs, geo-coded and spatio-temporal databases, ’big data’ sources and a variety of software tools for computing and graphical illustration. In this context

some selected aspects of teaching SAE including problems, challenges and experiences were discussed in detail. Three other presentations in this session looked at different aspects of teaching SAE at universities and, in particular, raised issues related to using simulations while teaching SAE, using selected R packages in this field and challenges faced by the system of education in terms of the needs of both academics and statistical offices.

The next invited session on **SAE applications** was organized by Ralf Münnich and chaired by Roy van der Weide from the World Bank. This session consisted of four presentations:

- Ralf Münnich − Small area applications: some remarks from a design-based view,
- Ugarte, MD, Adín, A., Goicoa, T., Militino, A.F., López-Abente, G. − Space-time analysis of young people brain cancer mortality in Spanish provinces,
- Rebecca C. Steorts − Constrained Smooth Bayesian Estimation,
- William R. Bell, Mark Seiss − A Modeling Approach to Estimating the Mean Squared Error of Synthetic Small Area Estimators.

This session was mainly devoted to different SAE applications using real data. From one point of view, there are only few National Statistical Institutes which use the SAE methodology in the production of statistical data. The reason is the difficulty of using model-based techniques in the production of small area estimates. On the other hand, statistical offices are increasingly responsible for delivering estimates at a lower level of spatial aggregation. This calls for applications using real data and taking into account practical situations which are faced by statistical offices.

In the main talk, Professor Ralf Münnich highlighted the impact of sampling designs on small area estimation methods. He also presented real applications of using small area estimation methods in the context of household and business data. In addition to sampling designs, Professor Münnich also considered methods of benchmarking in order to provide coherent results between design-based and model-based estimates. Three other presentations in this session were devoted to different SAE applications and included: analysis of young people brain cancer mortality in Spanish provinces, analysis using data coming from U.S. Census's Small Area Income and Poverty Estimates program and application of the modeling approach to a real application involving synthetic estimation of correct enumerations in the 2010 U.S. census using data from a post-enumeration follow-up survey.

The next invited session on **Benchmarking, design issues and nonresponse in SAE** was organized by Stefano Falorsi and chaired by Michel Hidiroglou. This session consisted of four presentations:

- Andrea Fasulo, Michele D'Alo', Lorenzo Di Biagio, Stefano Falorsi, Fabrizio Solari − Benchmark constraints for space and time unit level EBLUP estimators,
- Li-Chun Zhang, Alison Whitworth − Benchmarked synthetic small area estimation,
- Serena Arima, Gauri S. Datta, Brunero Liseo − Multivariate Fay-Herriot model with structural measurement error,
- Janusz Wywiał − On sampling design proportional to function of auxiliary variable order statistics.

Some very important topics were raised during this session, which related to the negative impact of nonresponse in the process of estimation, benchmarking and design issues. In the main presentation delivered by Fasulo *et al*., the authors focused on small area estimators based on unit level nonparametric mixed models with area random effects. They also considered the benchmark problem for SAE estimates, which was consistently extended to the case of space and time benchmark constraints. The presenters demonstrated practical applications of the issues raised in their presentation by reviewing two empirical studies and presenting their conclusions. The three other presentations were devoted to the problem of benchmarking, which is very crucial in production of statistical information as estimates for lower level of aggregations should add up to estimates at higher level, the problem of modeling using multivariate Fay-Herriot approach with structural measurement error and issues to do with basic properties of sampling strategies based on the sampling designs dependent on quintiles.

The invited session on **Population Census and SAE,** organized by Li-Chun Zhang and chaired by Stephen Haslett, included four presentations:

- Li-Chun Zhang − Census and SAE: Population size estimation,
- Ralf Münnich − Small area estimation in the German Census 2011,
- Paul Williamson, Karyn Morrissey, Ferran Espuny-Pujol − Survey reweighting as a means to SAE,
- Angela Luna-Hernandez, Li-Chun Zhang − Multivariate Generalized Structure Preserving Estimation.

The main aim of organizing this session was to present how small area estimation methodology can be used in the field of modern censuses, in which data are often collected using the mixed approach. In the main presentation delivered by Li-Chun Zhang, the author focused on the topic of census or census-like population size estimation. The presentation reviewed common traditional direct estimation methods, as well as some new developments in the treatment and modeling of enumeration coverage errors. Prof. Li-Chun Zhang also discussed some perceived challenges to the indirect estimation of local population size as well as the question of how design-based and model-based estimation can be used in the context of modern censuses. In the next talk, Ralf Münnich discussed the use of small area estimation methodology in the German census

2011. Some aspects of survey reweighting and SPREE estimation were also discussed as part of this session.

The last invited session, entitled **Other topics related to SAE**, which was organized by Domingo Morales and chaired by María Dolores Ugarte, included four talks:

- Wayne Fuller, Andreea Erciulescu − Small Area Prediction under Alternative Model Specifications,
- Domingo Morales, Miguel Boubeta, María José Lombardía − Empirical best prediction in Poisson mixed models,
- M.A. Hidiroglou, Victor Estevao − A comparison of small area and direct estimators via simulation,
- Monica Pratesi, Fosca Giannotti, Caterina Giusti, Stefano Marchetti, Dino Pedreschi, Nicola Salvati − Area level SAE models with measurement errors in covariates: an application to sample surveys and Big data sources.

This session mainly covered issues related to small area estimation methodology and not discussed in detail in the others. During this session special attention was paid to issues concerning modeling in the field of SAE. That was the topic of the main presentation given by Wayne Fuller and Andreea Erciulescu. In their talk, they discussed the construction of small area predictors and estimation of the prediction mean squared error, given different types of auxiliary information and for different population models and illustrated the problem with a study of soil erosion. The three other presentations also dealt with modeling and addressed such topics as the use of Poisson mixed models, area level SAE models with measurement errors in covariates and the comparison between SAE and direct estimators using a simulation approach.

There were also six sessions of contributed papers organized during the SAE 2014 conference, which covered different issues related to small area estimation methodology. Some of the topics covered included modeling in SAE, SAE applications, poverty mapping, SAE in business statistics and the use of Big Data in the context of SAE. In total, 72 talks and 10 poster presentations were delivered during the conference.

One of the highlights of the SAE 2014 conference was the panel discussion organized and chaired by Professor Elżbieta Gołata from the Department of Statistics at the University of Economics in Poznan, which addressed the newest achievements in SAE both in the theoretical and practical field. The panel brought together nearly all of the organizers and chairs of the invited sessions, who gave an account of the most important issues raised within the ten invited sessions. The panel constituted a scientific summary of the whole conference and was a great opportunity to review recent developments both in the field of theoretical and practical use of SAE.

Detailed information about the SAE 2014 conference is available on the conference website at www.sae2014.ue.poznan.pl.

The plan to organize future SAE conferences is an expression of the growing role of small area estimation methodology in the modern statistical world.

It is worth noting that the next SAE 2016 conference in the series of conferences under the auspices of the EWORSAE group is planned to be held in the Netherlands.

Moreover, the First Latin American International Statistical Institute Satellite Meeting on Small Area Estimation will be held on August 3-5, 2015, in Santiago, Chile. More details about this conference can be found at www.encuestasuc.cl/sae2015.

Prepared by:
Marcin Szymkowiak
Poznan University of Economics