# RANDOM-FOREST MODEL-ASSISTED ESTIMATION

**C. Goga[1]**

[1] University of Franche-Comté, France
e-mail: mailto:camelia.goga@univ-fcomte.fr

## Abstract

Abstract: Nowadays, surveys face more and more complex data sets with a large number of variables. These new data sets raise many challenges and traditional parametric methods for estimating interest parameters such as totals, ratios or quantiles may prove inefficient. In this work, we propose a new class of model-assisted estimators based on random forests. Under certain regularity conditions on the study variable, the random forest as well as the sampling design, the proposed model-assisted estimator is shown to be asymptotically design unbiased and consistent for the population total. A consistent variance estimator is proposed and the asymptotic distribution of the random-forest model-assisted estimator is obtained allowing to build confidence intervals. A new variance-estimator based on cross-validation technique is suggested. Simulation illustrate that the proposed estimator is efficient and can outperform state-of-the-art estimators, especially in complex and high-dimension settings.