

Using complex surveys to estimate the L_1 -median of a functional variable : application to electricity load curves

Mohamed Chaouch^(a) and Camelia GOGA^(b)

^(a) University of Reading, United Kingdom

^(b) Université de Bourgogne, Dijon, France

13 mars 2013

Mean profiles are widely used as indicators of the electricity consumption habits of customers. Currently, in Électricité De France (EDF), class load profiles are estimated using point-wise mean function. Unfortunately, it is well known that the mean is highly sensitive to the presence of outliers, such as one or more consumers with unusually high-levels of consumption. In this paper, we propose an alternative to the mean profile : the L_1 -median profile which is more robust. When dealing with large datasets of functional data (load curves for example), survey sampling approaches are useful for estimating the median profile avoiding storing the whole data. We propose here estimators of the median trajectory using several sampling strategies and estimators. A comparison between them is illustrated by means of a test population. We develop a stratification based on the linearized variable which substantially improves the accuracy of the estimator compared to simple random sampling without replacement. We suggest also an improved estimator that takes into account auxiliary information. Some potential areas for future research are also highlighted.

Key Words : Horvitz-Thompson estimator, k-means algorithm, poststratification, stratified sampling, substitution estimator, variance estimation.