Verification on appropriate sampling method
for Osram CDM CFL projects

based on

SACHS,L/HEDDERICH,J: Angewandte Statistik – Methodensammlung mit R, 12. Aufl., 2006

and

PAPULA, L: Mathematik für Ingenieure und Naturwissenschaftler, Bd. 3, 2. Aufl., 1997

Background

Osram is planning to implement several CDM projects involving the replacement of conventional incandescent bulbs (also none as GLS bulbs) by energy-efficient compact fluorescent lamps. Each project may involve around 500,000 GLS to be replaced. During the project, Osram needs to monitor the average operating hours of the replaced GLS as well as of the distributed CFLs. Additionally, Osram needs to check whether CLF distributed are still functional. Osram intends to do this by sampling.
The CDM Executive Board generally requires that whenever a sample is used to derive a parameter, "sampling shall be representative of the population" and that "parameter uncertainties" should be "quantified at the 95% confidence level."
Osram aims to have three different sampling groups that are described in detail in the following.
The above listed literature is consulted to check, if the sampling approach proposed by Osram fulfils the criteria of the applied statistical methods and the criteria specified by the CDM EB.

The Osram sampling proposal

The first sample group (the baseline sample group), will consists of a minimum of 100 households that use GLS and that will be randomly drawn from the target population of households that are later offered to exchange the GLS for a CFL (project database).
The second group (the project sample group), will consist of a minimum of 100 households that had used GLS and that have exchanged their GLS to CFL as part of the project. Those households will be also randomly drawn from the target population of households (project database).
In both groups, operating hours will be monitored.
The third group (the cross-check group), will consist of a minimum of 100 households that used GLS and that have exchanged their GLS to CFL as part of the project. Those households will be randomly drawn from the list of all households that have replaced GLS to CFL under the project. In the cross-check group the functionality of the CFLs will be monitored.

Statistical methods applied for the Osram sampling proposal

According to the laws of the inferential statistics, random sampling ensures results that are representative for the total population with a certain random sampling error (confidence interval +/-). Hence, the Osram sampling proposal ensures representative results, if the random sampling error is conservatively accounted for. The CDM EB requires that the random sampling error should be accounted for at the 95% confidence level.

If the number of samples (n) exceeds 30 (n>30)[1] the standardized normal distribution can be applied as approximation for the estimation of the average operating hours. As a result the following formula describes the confidence interval for the average operating hours.

$$\mu_{o,adjust} = \mu_{o,sample} \pm z * \frac{\sigma_{o,sample}}{\sqrt{n}}$$ [2]

Where:

$\mu_{o,adjust}$	confidence interval of the mean of the operating hours
$\mu_{o,sample}$	Mean of the operating hours of the sample
z	confidence level of 95% (z = 1.96)
$\sigma_{o,sample}$	Standard deviation of operating hours per day during the monitoring interval
n	Number of sampled cases

That means, that the average operating hours of all participating households is located within the above described confidence interval at a confidence level of 95 %.

For reasons of conservativeness the above formula is adjusted for the baseline operating hours:

$$\mu_{o,BL,adjust} = \mu_{o,BL,sample} - z * \frac{\sigma_{o,BL,sample}}{\sqrt{n}}$$

And accordingly for the project operating hours:

$$\mu_{o,SC,adjust} = \mu_{o,SC,sample} + z * \frac{\sigma_{o,SC,sample}}{\sqrt{n}}$$

For the sampling results from the <u>functionality check</u> in the cross-check sample group (% of CFL not functioning in the sample), the following procedure has to be conducted to be representative:

$$p_{adjust} = \hat{p} \pm z * \sqrt{\frac{\hat{p} * (1 - \hat{p})}{n}}$$ [3]

Where:

p_{adjust}	confidence interval of the CFLs not functioning
$\hat{p}$	Percentage of the CFLs not functioning in the sample
z	confidence level of 95% (z = 1.96)
n	Number of CLFs sampled

---

[1] PAPULA, L: Mathematik für Ingenieure und Naturwissenschaftler, p. 533

[2] PAPULA, L: Mathematik für Ingenieure und Naturwissenschaftler, p. 515 and SACHS,L/HEDDERICH,J: Angewandte Statistik – Methodensammlung mit R, p. 266

[3] Formula transformed: PAPULA, L: Mathematik für Ingenieure und Naturwissenschaftler, p. 530

For reasons of conservativeness the above formula is adjusted for the cross-check results:

$$p_{adjust} = \hat{p} + z * \sqrt{\frac{\hat{p} * (1 - \hat{p})}{n}}$$

Summary

The Osram sampling proposal ensures representative sampling results, if it is ensured that sampling is conducted randomly and the sampling error is conservatively taken into account at a confidence interval of 95% (as required by the CDM EB) and the above specified formulas are used to adjust the sampling results.