ANNEX 7 CUIDEMOS MEXICO PoA – SAMPLING PLAN

1. Introduction

This document details the sampling plan (the plan) for the "CUIDEMOS Mexico (Campana De Uso Intelegente De Energia Mexico) – Smart Use of Energy Mexico" Programme of Activities (PoA). The plan seeks to obtain unbiased and reliable estimates of the mean value of parameters used in the calculations of greenhouse gas emission reductions. 'Unbiased' in this case indicates that the sampling will not systematically underestimate or overestimate the mean value determined.

The CME for the Cuidemos Mexico PoA will implement a sampling regime for monitoring a national roll out of the Cuidemos Mexico PoA. In order to ascertain the appropriate statistical sampling methodology for the PoA, the plan design was outsourced to the Statistical Department of the University of Melbourne. The sampling approach that is described in the plan is based on the sampling plan designed by the University of Melbourne.

In order to obtain unbiased and reliable estimates of the mean value of parameters used in the calculation of greenhouse gas emission reduction for the CUIDEMOS Mexico PoA, the plan utilises the "Standard for Sampling and Surveys for CDM Project Activities and Programme of activities (Version 02.0)

This document will replace (ANNEX 8 CUIDEMOS MEXICO PoA – Sample group calculation, selection and maintenance).

2. Sampling Design

2.1. Objective and Reliability Requirements

The sampling objective is to determine the:

- Mean value of the operating hours of CFLs for each monitoring period during the crediting period with a 95/10 Confidence /Precision
- Proportion of operating CFLs for each monitoring period during the crediting period with a 95/10 confidence / precision

in compliance with "Standard for Sampling and Surveys for CDM Project Activities and Programme of activities (Version 02.0).

2.2. Target Population and Sampling Frame

The target population is households that can participate in the PoA within the geographic boundary of Mexico and complies with requirements of the project (e.g. exchanged up to 4 incandescent bulbs at a project distribution point). A list of households that participates in the PoA will be used as a sampling frame.

Households will be used as the unit for average operating hours calculations. This would involve the averaging of CFLs within households, which ensures that each household contributes equally to the overall mean, even in cases when there are only data available from at least two CFLs for a given household.

2.3. Sampling Method

2.3.1. Sampling Method - Project Sample Group (PSG)

A Project Sample Group (PSG) will be established in order to monitor a representative sample of all participating households in the PoA.

The samples for PSG will be randomly selected and selected across all combined CPAs under the PoA by applying 95/10 confidence /precision for sample size calculation in accordance with the footnote 13 of paragraph 19 of EB 65, Annex 2.

The chosen sampling method for the PSG is a national[1] stratified random sample, with stratification by municipality. This involves selecting households randomly within municipalities. The number of households to be randomly selected from each municipality will be proportional to the targeted distribution of CFLs to participating municipalities based on population. Selecting households in this manner will result in the strata being proportionally representative of the population. The households will then be selected randomly from within the municipalities.

The number of PSG houses in each participating municipality will be calculated using the following steps:

1. Taking the population in each municipality divided by the total population in all of the participating municipalities multiplied by the initial sample size (initially 220[2]).
2. Rounding these figures to the nearest whole number.
3. Sorting the list of municipalities from largest to smallest based on the values obtained in step 2.
4. Selecting municipalities from the above list until a total of 220 houses is achieved (starting from the largest to the smallest)

The result of this process will be a list of 220 households spread across municipalities across Mexico.

Note that certain municipalities have been excluded from the potential sample group. This is because there are no distribution points in these municipalities and therefore it is very unlikely that households from those municipalities will participate in the project. This ensures the number of households selected in each municipality will proportionally represent the expected distribution of CFLs across Mexico.

The method for selecting households to participate in the PSG from the municipalities identified above is systematic and random: a list of recruited households is generated and households are randomly selected from this group. Households will be recruited from distribution points that participated in the exchange of light bulbs. The following outlines the procedure:

---

[1] Appendix A rationalises the choice of national estimate
[2] Refer to Section 3.1 for sample size calculation

The CME will be recruiting households for PSG via distribution points. Firstly, a list of households that are interested in participating in the PSG will be generated. Households that are participating in the PSG will then be randomly selected from this list. The CME will explain to the randomly selected households the requirements for participating in the PSG. An agreement will be signed with each household thereby allowing CME to install monitoring equipment metering 4 light bulbs in the house.

The CME has designed the monitoring equipment in such a way that its installation and operation is not intrusive to the daily activities of the household. At the same time the household won't have access to the equipment or the possibility of manipulating it.

In the case that a household decides to cease participation in the sampling group, the CME will retrieve the monitoring equipment from the house, and recruit a replacement house using the method described above.

2.3.2. Sampling Method - Project Cross Check Sample Group (PCCG)

A sample of non-metered households will be surveyed at least annually to estimate the number of CFLs installed that are still in operation out of the total of four. The sampling method for PCCG will be a simple random sample. The sampling frame for the PCCG will be defined in the following ways. As it is expected that CFLs installed in the same period will have similar operating life, CPAs will be grouped according to distribution date.

A separate sample will be taken for each group (hereafter referred as "block") and the sampling frame will be the list of all households for a given block of CPA/s. Specifically, all CPA/s where distribution occurred within a three-month period will be combined for the purposes of this estimation and a sample will be taken from the set of all non-metered households in that block of CPA/s.

3. Sample Size

3.1. Sample size for PSG

A Project Sample Group (PSG) will be established at the PoA level. The desired precision of 10% for a 95% confidence interval is the basis for selection of the sample size for a national sample.

The samples for PSG will be randomly selected and selected across all combined CPAs under the PoA by applying 95/10 confidence /precision for sample size calculation in accordance with the footnote 13 of paragraph 19 of EB 65, Annex 2.

The purpose of establishing the PSG is to monitor a representative sample of all participating households in the PoA. The following statistical formula has been used to calculate the sample size for the PSG.

$$n > \left(\frac{Z * s}{e * \bar{x}}\right)^2$$

where,

n= Minimum sample size
z =Z Value, 1.96 for 95% confidence level
e= Precision, 10%
Standard deviation (s) = 2.330
Average daily operating hours ($\bar{x}$) = 3.234

The standard deviation (s) and average daily operating hours ($\bar{x}$) is based on the result from CUIDEMOS Mexico – CPA 1, Puebla data from 1 December 2009 to 30 November 2010. To allow for potential higher variance at the national level, increased contingency is built into the recommended sample size through an inflated standard deviation. Standard Deviation of 1.977 was estimated for the CUIDEMOS Mexico – CPA 1, Puebla, which was inflated to 2.33 to allow for potential higher variance at the national level and therefore achieve the required precision.

$$n > (\frac{1.96 * 2.33}{0.1 * 3.234})^2 = 199.4$$

Therefore, Minimum sample size (n) = 200 households

A further 10% oversampling has been applied to account for monitoring metering failure or losses incurred in the data collection process, resulting in a total sample size of 220 households. Based on the monitoring results, the CME may choose to increase or decrease the initial sample size to meet the required precision. This would be assessed with reference to the desired precision of 10% for a 95% confidence interval. If additional households were found to be required they would be determined using the same stratified random sample approach outlined above.

3.1.1. Incorporating the effect of meter failure

Note that it is inevitable that at some times, some meters will fail. When this occurs, the CME will work to repair the meter. During the time that the meter is not working or under repair, data will not be available. In this case, only days for in which there were meters functioning would be included in the calculation of the mean operating hours for each CFL. These will then be averaged across households to give an overall household average operating hours per CFL. However, in order to ensure all households included are statistically representative there ought to be a lower limit on the number of metered days acceptable for that CFL to be included.

All of the monitoring days within the time period will be used, in order to determine a lower limit on the number of acceptable days where monitored data is available, a minimum sample size for the number of monitoring days required will be calculated. Based on the most recent and accurate "CUIDEMOS Mexico – CPA 1, Puebla" data for Puebla from 1 December 2009 to 30 November 2010, the average and standard deviation for the individual 240 CFLs were:

Average daily operating hours ($\bar{x}$) = 3.234
Standard deviation (s) = 2.867

Since this sample is likely to be a large fraction of the total days measured, it is appropriate to use a finite population correction. An appropriate formula in this instance is therefore:

$$n > \frac{1}{(0.1\bar{x}/1.96s)^2 + 1/N}$$

Where N is the number of metered days. For 365 days, this results in a sample size of 166 days. Therefore, any CFLs for which there are at least 166 annual measurements will be included in the overall calculations. A minimum of two CFLs ought to satisfy this for each household, in order for that household mean to be included in the overall sample.

The deliberate 10% oversampling of households is designed to allow for some households not satisfying this minimum threshold while still satisfying the precision requirements. Therefore, the maximum number of samples that the CME could exclude for this reason from the calculation is number of samples equivalent to oversamples.

In some instances, it may be desirable to provide estimates for periods shorter than one year. In this instance, the same formula could also be used, with N replaced by the relevant number of days in the sampling period. For instance, if the sampling period was 183 days or six months, the minimum number of allowable days should be 114.

3.2. Sample Size for PCCG Group

In order to constitute the sample size for PCCG group, CPAs will be grouped according to distribution date. Each block of CPA/s may consist of one or more CPAs. Survey will be done for each block of CPA/s whether the block contains a single CPA or more. A separate sample will be taken for each of these blocks. Specifically, all CPAs where distribution occurred within a three-month period will be combined for the purposes of this estimation and a sample will be taken randomly from the set of all non-metered households in that block of CPA/s. If no group of CPA could be formed or a single CPA distribution occur in three months time then a separate PCCG survey will be carried out for that CPA. Sample size for the PCCG survey is calculated as below.

The desired precision of 10% for a 95% confidence interval is the basis for selection of the sample size for PCCG group. This is achieved by assuming the following parameters to calculate the sample size:

Average number of CFLs operating per household $(\bar{x}) = 3.6$
Standard deviation for number of CFLs operating per household $(s) = 1.6$

The required precision can therefore be achieved using the following equation:

$n > (\frac{1.96s}{0.10\bar{x}})^2 = 96.0$, giving a minimum sample size of 97 households.

This means that each block of CPA/s will have a minimum sample size of 97 non-metered households.

The CME may adjust this sample size if required. If additional households were found to be required they would be determined using the same sample approach outlined above.

4. Consequences for individual CPAs

4.1. Application of PSG results to individual CPAs

The initial sampling frame will be for all existing CPAs at the time of sampling, which will be a subset of the entire project targeted over the course of the PoA. However, all future CPAs will be based around similar distribution points, target the same population and will have similar usage patterns. Therefore there is no need to conduct additional sampling for future CPAs.

It should be noted that only the sampling that occurs once; the monitoring is ongoing throughout the life of the PoA. The representative sample of CFLs will be monitored for each monitoring period throughout the life of the PoA. The CME will record the "power" of all CFLs using the nameplate data during the distribution of the CFLs. This means that the power rating of the CFLs that will be distributed in the future CPAs will be also recorded and this information will be utilized for emission reduction calculation. Similarly, "Operating hours" data will be obtained by monitoring representative samples of CFL using monitoring equipment.

It can be argued that factors such as socio-economic status, building type, climate and elevation may cause some variability in the operating hours of CFLs, however, analysis of CUIDEMOS Mexico – CPA 1, Puebla data suggested that differences in daylight hours are the largest source of variability, accounting for 79% of the variability in operating hours. This result is countered by the fact that over the course of a year, the average daylight hours is highly consistent. These results are detailed in Appendix A.

As long as the operating hours estimates used for all CPAs in the PoA are based on the corresponding months from the sample, the obtained results can be applied to the future CPAs. Therefore, the CME will utilize the results obtained from the national sample for all future CPAs. The decision to stratify the sample across municipalities ensures that the sample accurately represents the target population in terms of socio-economic status, building type, climate and elevation.

Note that the proposed sample is national and the sampling frame also includes Puebla region where CPA 1 is currently under operation. As demonstrated in Appendix A, the national sample will be representative of individual regions therefore the mean value of the operating hours data obtained from the national sample will be considered as the mean value of the operating hours for CPA 1 and will be applied in the emission reduction calculation of CPA1. All future estimates for CPA 1 will be based on the national sample results.

5. Data

5.1. Field Measurements

The variables to be measured are as follows:

- Mean value of the operating hours of CFLs for each monitoring period during the crediting period
- Proportion of operating CFLs for each monitoring period during the crediting period

The length of each monitoring period will be decided by the CME (within the UN guideline of a maximum period of 1 year), with surveys of crosscheck lamps to occur at least annually.

The PSG will initially consist of 220 households selected using a stratified random sampling process. Equipment will be installed in the selected households to monitor the operating hours of their light bulbs (four bulbs will be monitored per household). The operating hours data for each light bulb will be collected via the monitoring equipment installed in the PSG

households. The equipment will feed monitoring information back to a centralised database in real-time. The household data will then be averaged over the monitoring period to determine the average operating hours of CFLs distributed.

The PCCG will initially consist of at least 97 households selected using a random sampling process as described above. The PCCG survey will check the continuing operation of CFLs. This will ensure that a minimum of 97 households will be surveyed for each block of CPA/s.

The mean value of operating hours data obtained at the PoA level and the proportion of operating CFLs data obtained from the block of CPA/s will be used to calculate the CERs generated from the CPAs within the PoA.

5.2. Quality Assurance /Quality Control

The CME will implement the data record keeping system as described in section E.7.2 of the SSC-PoA-DD. In order to achieve good quality data, the CME will outsource the monitoring and data collection to an experienced and qualified third party. This will also rule out any conflict of interest of those involved in the data collection and analysis. The CME will oversee the entire data collection and analysis process.

5.3. Analysis

The mean value of operating hours data and proportion of operating lamps data will be collected for each monitoring period, and will be used to calculate emission reductions for that portion of the crediting period.

The Project Sample Group (PSG) households will be monitored to obtain the mean value of operating hours. The mean value of the operating hours will be directly extrapolated to all households involved in the PoA.

The proportion of operating CFLs for each block of CPA/s will be surveyed at least annually. The results obtained from the PCCG survey will be extrapolated to all households involved in that block of CPA/s.

6. Implementation

The CME will outsource the PSG and PCCG data collection to an experienced and qualified third party.

7. Monitoring Equipment

The monitoring equipment will record the operating hours and/or electricity consumption of CFLs belonging to the PSG group. Monitoring equipment will be spot checked to ensure ongoing functionality and accurate calibration. If irregularities are recorded with equipment, this will be flagged immediately by the monitoring system and corrective actions will be implemented to repair or re-calibrate metering equipment. Calibration of the equipment will be conducted by the CME at least once in three years or as required.

Appendix A

This section of the plan has been extracted from the sampling plan designed by the University of Melbourne for the CUIDEMOS Mexico PoA.

A. Sources of variability in daily operating hours

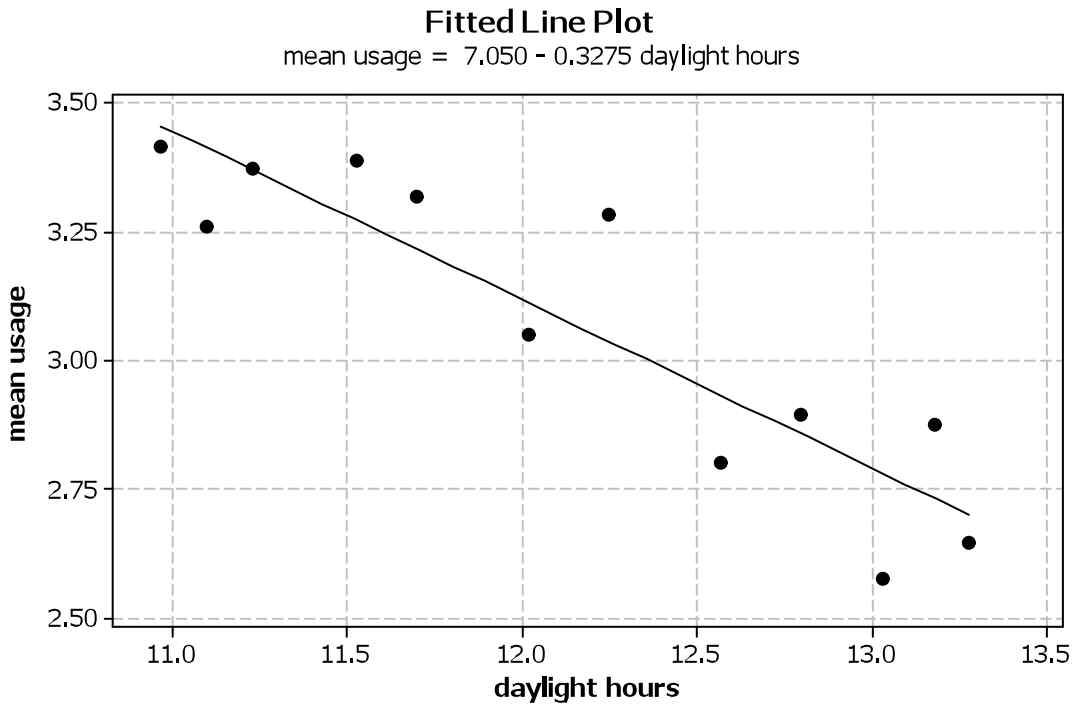
It is reasonable to imagine that the amount of daylight in a region, would directly affect the average operating hours. This has been investigated in the next two subsections.

A.1 Average operating hours for Puebla

Daily operating hours data are available from Puebla for December 2009 to November 2010 (inclusive). It is possible to consider the relationship between daylight hours and mean CFL operating hours for this state and period of time. We expect to see an increase in average operating hours for a decrease in natural daylight. The results of this are given below, along with the standard deviation for the operating hours.

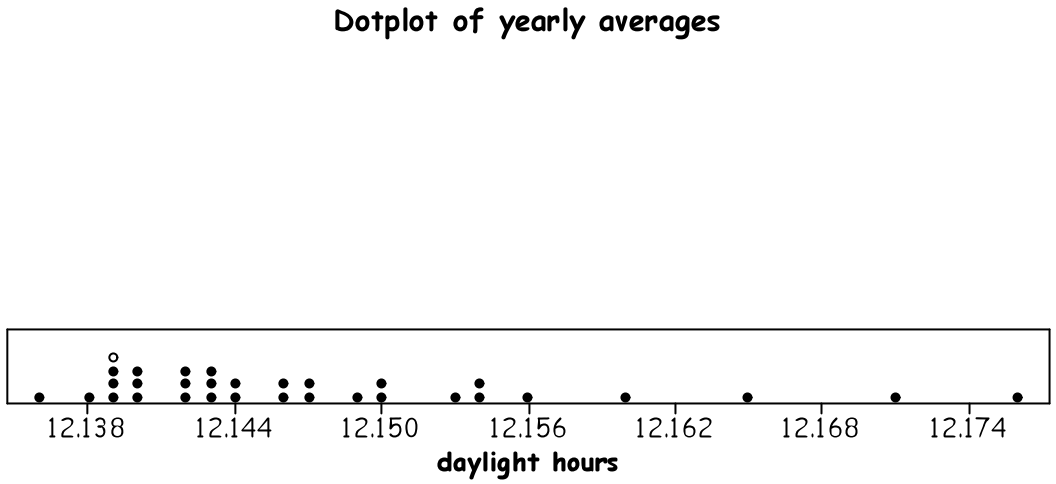
month	average operating hours	SD operating hours	daylight hours
December 2009	3.416	3.284	10.97
January 2010	3.261	3.245	11.10
February 2010	3.388	3.155	11.53
March 2010	3.052	3.112	12.02
April 2010	2.803	2.761	12.57
May 2010	2.575	2.635	13.03
June 2010	2.648	2.850	13.28
July 2010	2.874	2.919	13.18
August 2010	2.896	2.875	12.80
September 2010	3.285	3.150	12.25
October 2010	3.318	2.997	11.70
November 2010	3.372	3.093	11.23

The linear relationship between daylight hours and average operating hours according to these data is shown in the plot below. This demonstrates a strong relationship between daylight hours and operating hours with an adjusted R^2 of 79%, suggesting that 79% of the variability in mean operating hours can be explained by the variability in daylight hours.



A.2 Yearly averages of daylight hours

While the differences in latitude affect the number of hours of daylight throughout the year, it is expected that these effects would be averaged out over the course of the year. The following plot gives the yearly averages for the 32 states.



As this demonstrates, the yearly averages are much more consistent between states (the range is only 2.4 minutes in total in the year) than the monthly averages, which vary from 10 hours to 14.3 hours. Puebla is indicated with an open circle; it is on the lower end of the range of average daylight hours.

Ultimately, it is not possible to be certain how daily operating hours will vary between states and regions without collecting data across the country. However, these results suggest that, while daylight hours contribute substantially to variation in operating hours on a monthly basis, this averages out over the course of a year. For these reasons, it is most likely not necessary to perform CPA-specific sampling. By assuming relative uniformity across the country, this national estimate will provide a more precise estimate for each individual CPA, compared with estimates based on smaller sub-samples within each CPA region.