SAMPLING MANUAL

A GUIDE TO SAMPLING UNDER THE CDM WITH SPECIAL FOCUS TO POAS

FIRST EDITION

Published by/ Copyright © 2012
KfW Bankengruppe
Palmengartenstraße 5-9
60325 Frankfurt am Main
Germany
Fax +49 69 7431-4775
www.kfw.de/carbonfund

Contact.
Martin Schröder, KfW Bankengruppe

Authors.
Sven Feige, Marc André Marr, Perspectives GmbH, Hamburg, Germany

First printing April 2012
Frankfurt am Main

Important Notice - Disclaimer

The information contained in this document is for the purposes of exemplification only and is provided solely to assist any prospective entity involved in CDM activities with regard to its own individual analysis, evaluation and documentation. It does not purport to be comprehensive, it cannot substitute individual analysis and documentation. Any reader of this document is recommended to seek its own individual technical, financial, legal and other advice as it deems necessary for such purpose.

Neither KfW nor any of its directors, officers, employees, advisors or agents makes any representation or warranty or gives any undertaking of any kind, express or implied, as to the actuality, adequacy, accuracy, reliability or completeness of any information contained in this document, or assumes any undertaking to supplement any such information as further information becomes available or in light of changing circumstances. No liability of any kind whatsoever is assumed by KfW or any of its directors, officers, employees, advisors or agents in relation to any information contained in this document.

The elaboration of the Sampling Manual was financed by the KfW-managed “PoA Support Centre”, which contributes to the expanded use and implementation of the PoA approach and which has been initiated and funded by the German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety (BMU).

Photos:
Cover B. Feige

Table of contents

List of figures

List of tables

List of abbreviations

AMS	Approved methodology for small-scale CDM project activities
BMU	German Federal Ministry for the Environment, Nature Conservation and Nuclear Safety
CDM	Clean Development Mechanism
CME	Coordinating/Managing Entity
CER	Certified Emission Reduction
CMP	Conference of the Parties serving as the meeting of the Parties to the Kyoto Protocol
CFL	Compact fluorescent lamp
CO_2	Carbon dioxide
COP/MOP	Conference of the Parties serving as the meeting of the Parties to the Kyoto Protocol
CPA	Component project activity
CPA-DD	Component project activity design document
DNA	Designated National Authority
DOE	Designated Operational Entity
EB	CDM Executive Board
GHG	Greenhouse gas
ILC	Incandescent light lamp
IPCC	Intergovernmental Panel on Climate Change
kWh	Kilowatt hours
LED	Light-emitting diode
LDC	Least Developed Countries
LSC	Large-scale
PDD	Project Design Document
PoA	Programme of Activity
PoA-DD	Programme of Activity Design Document
SSC	Small-scale
SWH	Solar water heating
UNEP	United Nations Environment Programme
UNFCCC	United Nations Framework Convention on Climate Change
WBT	Water boiling test

1 Introduction

Climate Change and CDM

The need for efficient, measurable and sustainable solutions for mitigating greenhouse gas emissions has created manifold of mitigation instruments and strategies of which one very famous is the Clean Development Mechanism (CDM). The CDM is one of the flexible instruments under the Kyoto Protocol for mitigating GHG emissions. "The CDM allows emission-reduction projects in developing countries to earn certified emission reduction (CER) credits, each equivalent to one tonne of CO_2. These CERs can be traded and sold, and used by industrialized countries to meet a part of their emission reduction targets under the Kyoto Protocol. The mechanism stimulates sustainable development and emission reductions, while giving industrialized countries some flexibility in how they meet their emission reduction limitation targets."[1]

CDM project activities

More than 4,000 CDM project activities have been registered to date expecting more than 2,000 million Certified Emission Reductions (CERs) until the end of 2012[2]. Various technologies and project types are eligible under the CDM. So far mainly single projects have been developed. However, the concept of Programme of Activities (PoAs) allows combining an unlimited number of single projects under one framework[3]. Single project activities and PoAs applying for CDM approval at UNFCCC need to be additional, meaning that such an activity would not have occurred in the absence of the CDM. This usually needs to be proven based on strict and pre-defined procedures. In order to maintain environmental integrity of the CDM and for being most conservative regarding the generation of emission reductions, a rather comprehensive system of rules and regulations has emerged over time for monitoring and verification of emission reductions achieved under the CDM scheme.

Sampling as part of the determination of emission reductions

One important requirement is the generic requirement for each CDM project activity to calculate and document achieved emission reductions in a conservative and transparent manner (usually referred to as monitoring), allowing verification of emission reductions by an independent third party. For many projects involving a high number of appliances that maybe even spread over a certain geographical area (e.g. cook stoves, solar lamps, biogas-digester, solar water heater), the conduction of the monitoring in a practical and feasible way can only be done based on a sampling approach. Sampling is a very important element for measuring emission reductions of CDM projects and PoAs. A CDM project may consist of several thousand single appliances (e.g. LED lamps) in various consumer groups (e.g. households, commercial buildings) and could be spread over a whole country. For those project types, solid sampling is a decisive instrument to keep transaction costs for monitoring at a feasible level and simultaneously to not question the accuracy of results.

Focus of this manual

The matter of sampling for a PoA, with its high number of projects and appliances included, becomes even more relevant than for single CDM project activities. The key objective of this Sampling Manual is to support CDM project developers (participants) and stakeholders responsible for the sampling for monitoring in providing guidance on how to choose the

---

[1] UNFCCC (2012): About CDM: Available at: http://cdm.unfccc.int/about/index.html

[2] Status as of April 2012: UNFCCC (2012): CDM in Numbers: Available at: http://cdm.unfccc.int/Statistics/index.html

[3] KfW Bankengruppe (2010): PoA Blueprint Book: Guidebook for PoA coordinators under CDM/JI, 2nd Edition

right sampling approach. Sampling is not automatically required for every CDM project activity, though. It depends mainly on the applied CDM methodology and additional CDM specific requirements if sampling is allowed or even required.

Typical project types requiring sampling

Sampling is typically very important if a relative high number of units or appliances are replaced, modified or installed under the project activity. The sampling approach is taken to enable such projects and make them feasible in terms of transaction costs and practically implementable (it would not be possible or would require high costs to monitor all single appliances like a million of energy efficient lamps in a single project). Typical CDM project types that require sampling approaches include

- Compact fluorescent lamps
- Improved cook stoves
- Small domestic biogas plants
- Solar water heating systems
- Building refurbishment

Furthermore, the following project types may also require sampling approaches depending on the specific project circumstances:

- Domestic or industrial boilers
- Efficient air conditioning
- Efficient refrigerators
- Methane utilisation/destruction

Sampling under CDM and PoAs

While several methodologies and some specific project types require a sampling and survey approach, specifically Programme of Activities (PoAs) can be designed in a way that makes sampling relevant. Under the umbrella of the PoA theoretical an unlimited number of CDM Programme Activities (CPAs), which are the actual underlying projects achieving the emission reductions, can be included. In case a high number of CPAs are included under the PoA, the number of single appliances further increase compared to a single CDM project and hence increase the need for proper sampling approaches within the PoA. Under a PoA the sampling should not only consider the sampling within each individual CPA (similar to one single project) but rather sampling for the whole PoA since the total number of samples can be assumed to be lower under most circumstances. A PoA that is set up for enabling the distribution of efficient cook stoves may involve hundreds of individual CPAs, each CPA consisting of several thousand cook stoves. If the underlying conditions of all CPAs are similar (homogeneous), a sample could be taken out of the total population drawn from all CPAs.

In the case of PoAs, sampling may be required on various levels and in various stages of the PoA development. Taking the example above, on the level of each CPA, sampling may be applied according to the underlying CDM methodologies. Imaging a PoA consists of 3 CPAs where each CPA covers 15,000 cook stoves. For monitoring of each CPA only a certain sample of these 15,000 stoves would need to be monitored per CPA. Another approach could be to have a cross-CPA sample (i.e. a certain sample of the total number of 45,000 stoves).

Sampling at inclusion of CPAs and at verification

The PoA concept is usually structured in a way that once the PoA is registered from the UNFCCC as CDM PoA, it is possible to develop and include further CPAs into the PoA without an individual registration of each CPA from UNFCCC. However, a third party, called Designated Operational Entity (DOE), needs to check whether the CPA to be included complies with the eligibility criteria as defined in the registered PoA. These eligibility criteria are technology and PoA specific and are specified in the registered PoA documents. The DOE needs to check whether all CPAs to be included in the PoA are in line with such criteria. This is another step where sampling may be relevant. Therefore a reasonable sampling approach would need to be defined that comply with the requirements according to the CDM regulatory framework and the underlying methodology(ies). Another relevant step for sampling is the verification of emission reductions achieved under the PoA. The Coordinating/Managing Entity (CME) may use sampling on the level of the PoA. In addition the DOE may conduct sampling in a similar approach than for inclusion of CPAs during verification.

Sampling under the CDM is not always of the same type. Basically, two different types can be differentiated:

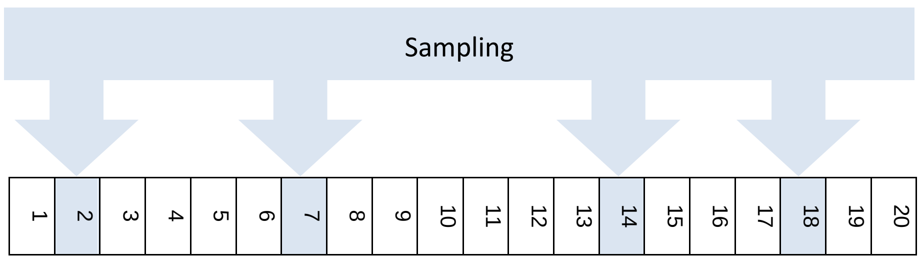
a) Sampling instead of a complete census

b) Sampling as periodical measurements

Sampling replacing a census

Firstly, sampling can be conducted instead of a complete census as shown in Figure 1. Assuming you need to check a specific parameter for a huge number of elements (e.g. the operation of improved cook stoves in all households of a specific region or CPA) you could apply a sampling procedure and based on that only visit a certain number of households (e.g. 100) instead of visiting every household (e.g. 30,000). Thus it would not only reduce the related efforts and costs significantly, but in many cases it would be the only reasonable and realistic approach. Otherwise the effort and costs for monitoring and data gathering would make it totally unfeasible for project participants to realize CDM projects and PoAs.

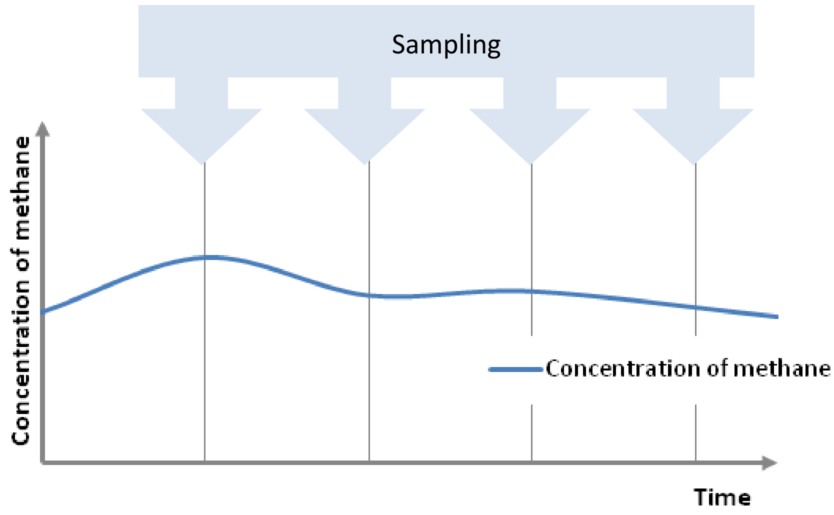
Figure 1: Sampling instead of a complete census



Sampling as periodical measurement / determination

Secondly, sampling is often applied for periodical measurements or determination of a specific parameter (e.g. instead of a continuous measurement of the methane concentration in a biogas stream you conduct one measurement per month). Figure 2 displays such a periodical sampling.

Figure 2: Periodical sampling



Although both types of sampling share several similar characteristics, preparation and execution of the sampling approach is different. The technical implementation of a periodical measurement by far is much easier than preparation of a sampling and survey procedure based on a population of thousands or even millions of elements. Handling of such huge numbers of elements alone is a challenge. Any mistake during preparation or execution of the sampling and survey procedure can easily result in tremendous additional costs due to the still relatively high number of samples. This manual shall assist project participants in setting up a robust sampling and survey plan that satisfies at the same time official requirements defined by the CDM framework and expectations of project participants to have everything under control. Only with a solid and workable sampling plan in practice, the monitoring and hence the expected emission reduction will finally materialize.

This manual shall provide required theoretical background and knowledge to enable project participants to start with setting up a solid sampling and survey procedure. Nevertheless, intensive reading of the most recent UNFCCC rulings remains essential. All practical examples in this manual are intended to provide practical support to CDM project participants. However, its application requires intensive consideration by project participants in order to adjust the approaches in the examples properly to the specific circumstances of the individual project(s). This is especially relevant since sampling approaches in general require taking project specific circumstances into consideration.

1.1 Outline of this sampling manual

Chapter 1: Introduction

The information in this manual is structured as follows. Chapter 1 provides basic information on the purpose of this manual and the concept of sampling. A brief overview on important CDM institutions relevant to sampling is presented and the typical project cycles of single CDM project activities and programmatic CDM project activities are described. In order to understand and use the manual the reader should be fairly familiar with the CDM process as described in this chapter.

Chapter 2: Background on statistics and sampling

Chapter 2 provides the statistical background that is required to set-up a well-functioning sampling procedure. If you are applying sampling instead of a complete census it has to be ensured that the results of the sampling reflect as close as possible the situation in the total population of interest. The indicators to determine accuracy of sampling are precision and level of confidence, which are explained in chapter 2.

Chapter 3: Sampling in the CDM

Chapter 3 deals with sampling in the CDM. The CDM framework with its hierarchy of documents defines the requirements for sampling in the CDM. The documents relevant to sampling are presented and explained in this chapter. A core document on sampling is the *Standard for sampling and surveys for CDM project activities and programme of activities*, which has been initially adopted by EB65 in November 2011. Nevertheless, several approved CDM methodologies define specific sampling requirements. Chapter 3 also includes a description of the commonly used sampling approaches in the CDM, which are simple random sampling, stratified random sampling, cluster sampling, systematic sampling and multi-stage sampling.

Chapter 4: Sampling in practice

Chapter 4 offers practical information on how the sampling requirements defined by UNFCCC can be implemented on the ground. Typical pitfalls are highlighted in order to facilitate smooth-running project implementation and materialisation of emission reductions in form of CER issuance. The development of a proper sampling plan is the core task to be elaborated and implemented according to CDM requirements and successful project implementation. Hence, chapter 4 provides some hints on available software that might be useful in this context. Finally, the sampling procedure has to be validated by DOEs. Therefore, a section on sampling on the DOE level is included in chapter 4.

Chapter 5: Sampling blueprints

Chapter 5 provides practical examples of projects that include a sampling procedure. The examples are presented in a way that allows potential project participants to easily adjust these examples to their specific project circumstances. The examples are selected in a way that four different project types applying most commonly used sampling approaches are shown. The sample cases taken reflect most relevant project types applied for PoAs and involving need for solid sampling procedures. The selected project types are cook stoves, efficient lighting, solar water heaters and rural biogas.

1.2 CDM projects

CDM institutions

The CDM framework is a rather demanding structure of institutions, standards, decisions and rules. A brief overview of important CDM institutions and entities is presented in Table 1.

Table 1: Important CDM institutions and entities

Institution	Abbreviation	Function
CDM Executive Board	CDM EB	Supervises the Kyoto Protocol's clean development mechanism under the authority and guidance of the CMP. The Board is the ultimate point of contact for CDM project participants for the registration of projects and the issuance of certified emission reductions.
Methodologies Panel	Meth Panel	Established to develop recommendations to the Board on guidelines for methodologies for baseline and monitoring plans and for new/revised methodologies.
Small-Scale Working Group	SSC WG	Established to prepare recommendations on submitted proposals for new baseline and monitoring methodologies for CDM small scale project activities.
Designated Operation Entities	DOEs	Private certifiers who validate projects and verify emission reductions. It has two key functions. A DOE (i) validates and subsequently requests registration of a proposed CDM project activity and (ii) verifies emission reductions of a registered CDM project activity, certifies as appropriate and requests the Board to issue certified emission reductions accordingly.

Source: adapted from UNFCCC (available at: http://cdm.unfccc.int/EB/governance.html)[4]

Scale of CDM projects

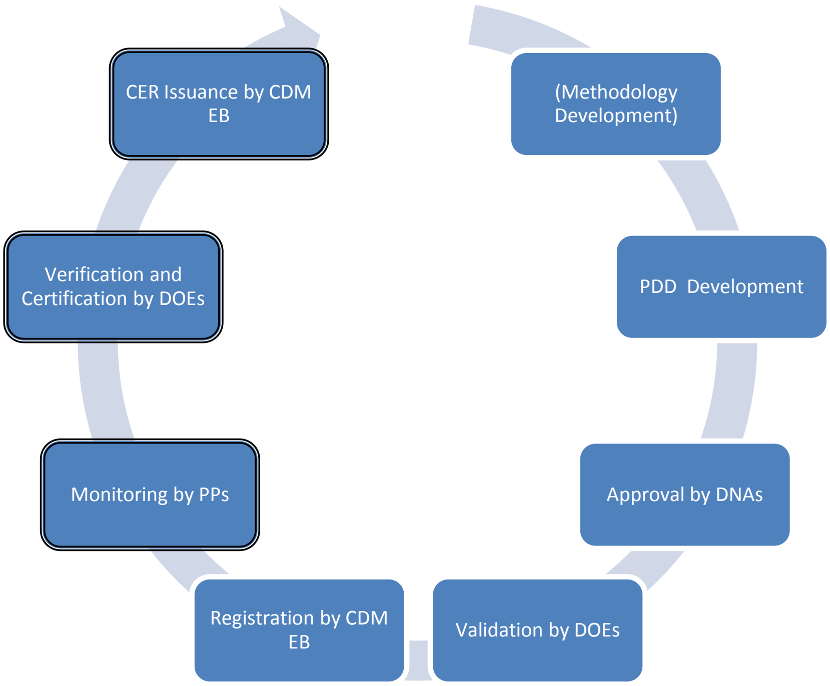
In order to apply the correct rules to the specific project activity or PoA it is important to differentiate between small-scale project activities (SSC) and large scale project activities (LSC). Small-scale project activities are limited in regard to installed capacity, achieved energy savings or achieved emission reductions depending on the project type[5].

CDM project cycle

Before a CDM project can achieve issuance of related CER volumes, several steps have to be passed. A typical sequence of a CDM project is shown in Figure 3.

[4] In order to understand and use this sampling manual you should be fairly familiar with at least the important CDM institutions and entities described in Table 1.
[5] Further information on the limits of small scale project activities can be found at: https://cdm.unfccc.int/Reference/Guidclarif/ssc/methSSC_guid06.pdf

Figure 3: The CDM project cycle



Sampling procedure already important in the PDD

Every CDM project activity[6] has to apply an approved CDM methodology[7] or a combination of approved methodologies. If an approved methodology is not available for a specific technology or measure, project participants can develop a new methodology and propose it to the regulatory body of the CDM. In order to start the validation and registration process of a CDM project, a Project Design Document (PDD) has to be developed. Based on the PDD, approval by DNAs and validation by a DOE will take place[8]. Finally, if the project fulfils all CDM requirements, registration by the CDM EB can be requested. The design of the project activity which is described in the PDD is fixed with the registration. Any (significant) change afterwards (e.g. project is not implemented as described in the PDD) would require additional effort and interaction with a DOE and in most cases a renewed approval by the CDM EB (with the remaining risk of non-acceptance). In order to prevent additional cost and time required, any sampling and survey procedure should be well elaborated and documented in the PDD and its accompanying documents that will be submitted for validation so that risk of post-registration changes are avoided.

[6] The recommendations of this manual refer in general to CDM project activities and PoAs/CPAs if not otherwise stated.

[7] A CDM methodology basically defines the framework for a CDM project activity providing relevant information on the applicability of the methodology, on the procedure to determine the baseline scenario and to demonstrate additionality, on the approach to calculate emission reductions based on project, baseline and leakage emissions and on the required monitoring.

[8] Approval by DNAs and validation by DOEs may be prepared partially in parallel.

Sampling during monitoring and verification

Monitoring, verification and CER issuance are recurring steps for each monitoring period. Normally, project participants choose a monitoring period of one year but also other frequencies may be possible, if not stated otherwise in the applied methodology. For each monitoring period the related emission reductions have to be determined based on the applied monitoring procedures as defined in the monitoring plan. Therefore, it is essential that the sampling and survey approach that is presented in the monitoring sections of the PDD is appropriate, robust and at the same time cost-effective. A too laborious approach will result in high costs while a too rudimentary approach might result in a loss of CERs as the achieved emission reduction cannot be verified as initially expected.

This manual provides practical guidance on sampling approaches for both phases, the validation phase (prior to CDM registration of the project activity) and for the monitoring and verification phase (after the CDM registration of the project activity).

Further reading:	
UNFCCC	http://cdm.unfccc.int/about/index.html
UNEP	http://www.cd4cdm.org/
IGES	CDM in Charts: http://www.iges.or.jp/en/cdm/

1.3 Programme of activities

PoAs shall promote CDM activities

While some countries like India, China or Brazil were able to realize a high number of CDM projects, many other countries were not able to develop any or only few CDM project activities. Similar observations have been done regarding applied project types. While project types like renewable energy production, methane avoidance or industrial and commercial energy efficiency are rather common under the CDM, others are rarely seen. Especially project types that have a huge sustainable benefit and have a very positive impact on local communities, but are very difficult to be implemented and monitored (e.g. cook stoves, LED lighting), have not taken up under the CDM in a similar scale, yet. The reasons are mainly CDM related. Those project types involve a high number of individual appliances and are usually dispersed over a wide geographical area. The effort and cost for distributing and especially monitoring the emission reductions is very high, while the expected emission reductions are comparably low. The dilemma is that those project types with high environmental and social co-benefits are generally well suited for the Least Developed Countries (LDCs).

In order to overcome some of the existing deficits and to enable underrepresented project types and CDM activities in underrepresented countries, the concept of Programme of Activities was introduced under the CDM. The EB worked on the operationalisation of the concept and established the basic rules, as well as the initial procedures for registration and issuance. The opportunity for project participants to utilise the new concept was finally opened up in 2007.

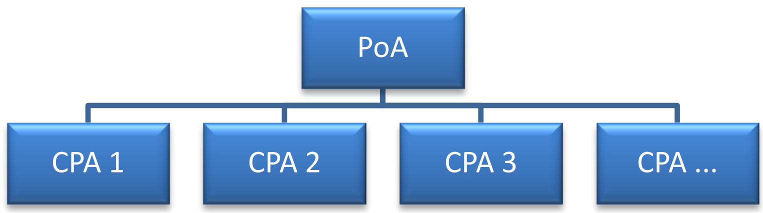
Similarity between a single CDM project activity and a CPA

The PoA concept was ought to reduce the CDM related transaction costs for implementing CDM activities and enable the scaling-up of CDM activities even across countries. As the PoA rules have been adapted from the single CDM project rules, the treatment of each CPA is still largely akin to that of a single CDM project. This is mainly due to the fact that the underlying methodologies and tools have to be applied more or less in the same way for CPAs under a PoA than for single projects. This has had implications especially for project types that have faced difficulties so far under the standard CDM and that usually have high CDM transaction costs (e.g. due to complex project design, surveys, monitoring requirements, complex sampling approaches). Taking those aspects into consideration, it becomes obvious, that a robust and well elaborated sampling approach will be a very relevant instrument for reducing the overall CDM transaction costs for those mentioned technologies under a PoA.

PoA structure

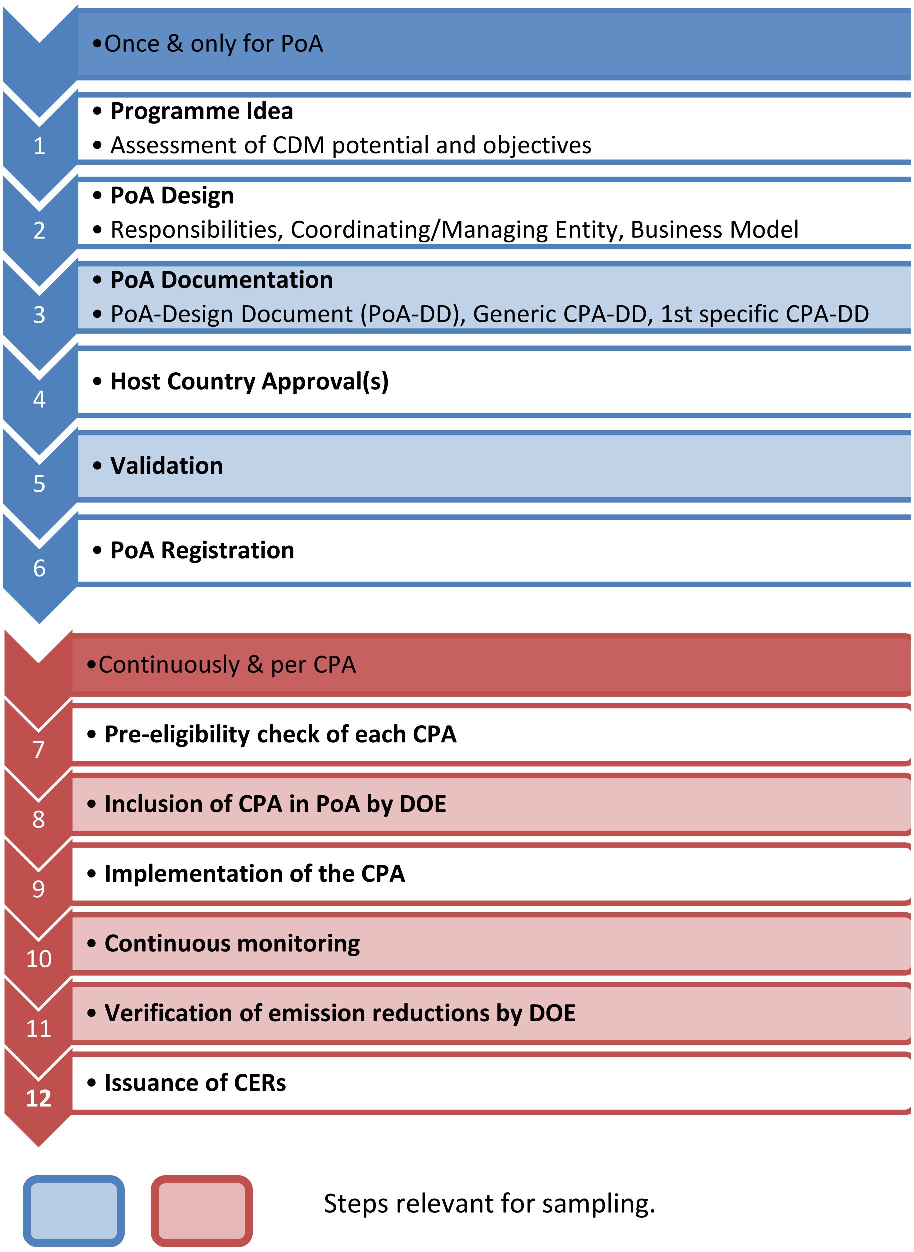
Under a PoA, a framework is developed that allows the inclusion of many individual activities (CPAs). Transaction costs are thereby shared by the number of CPAs under the PoA and partially moved to the level of the PoA. A PoA might be structured as shown in Figure 4. If there are changes of the general PoA layout over time a new version of the PoA may arise and from that point in time onwards new CPAs are included following the new version of the PoA.

Figure 4: PoA structure



Once the PoA is approved with the first CPA, further CPAs can be added to the PoA. Typical steps in order to set up a PoA and to include CPAs under the PoA are presented in Figure 5.

Figure 5: Typical work steps for setting up PoAs and inclusion of CPAs



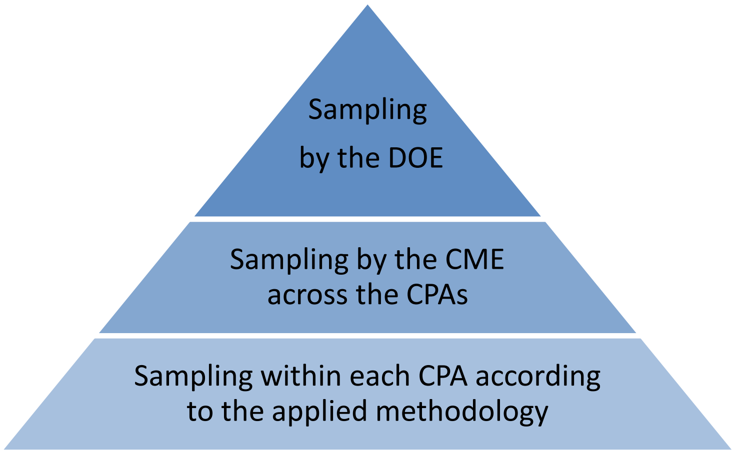
The role of the CME

The Coordinating/Managing Entity (CME) is the focal point of a PoA, mainly responsible for the interaction with UNFCCC, DNAs and DOEs. The CME has to setup the procedures for cooperation with the CPA developers and is normally responsible for negotiation of emission reduction purchase agreements with CER buyers. The CME may be at the same time the exclusive CPA developer, may develop CPAs under the PoA in parallel to other CPA developers or act solely as CME and does not develop CPAs under the PoA itself.

Sampling in PoAs

Sampling in PoAs may occur at different timing and on different levels. It is a requirement from the CDM rules and procedures that a PoA provides a sampling plan, if sampling is required. If sampling is relevant for a PoA, the PoA documentation should already provide sufficient description of the sampling approaches applied under the PoA. During validation of the PoA and during inclusion of CPAs, DOEs may apply sampling approaches for keeping efforts at an adequate level. During monitoring of emission reductions the CME and/or individual CPA developer have to apply sampling procedures if relevant for the PoA. Afterwards, the DOE may again apply a sampling approach for the verification of emission reductions. Hence, the following levels of sampling can occur within a PoA (see Figure 6).

Figure 6: Sampling levels in PoAs



The standard for sampling and surveys provides separate sections for sampling applicable to PoAs and sampling conducted by DOEs (see Chapter 3.2.2).

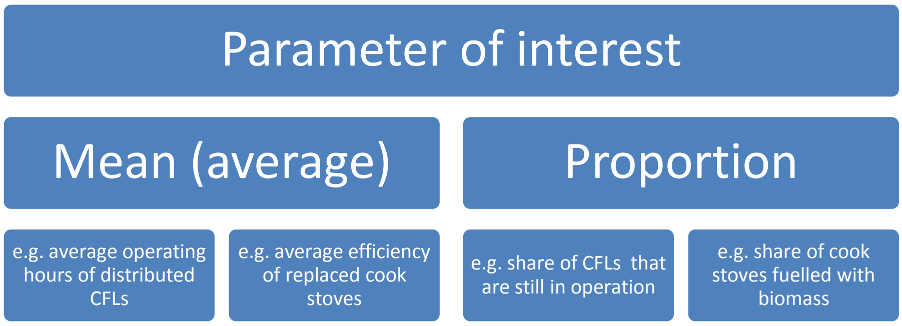
Further reading:	
UNFCCC	http://cdm.unfccc.int/ProgrammeOfActivities/index.html

2 Background on statistics and sampling

Sampling in the CDM is applied for the determination of the value of specific parameters in case the parameter has to be determined based on a huge number of elements. By applying a sampling approach, fewer elements have to be evaluated. The parameters of interest can be divided into two categories as shown in Figure 7.[9]

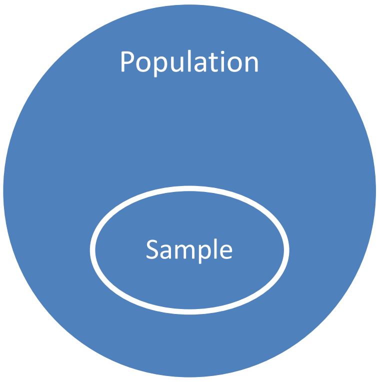
Means and proportions

Figure 7: Parameters of interest



Means are calculated based on continuous variables. Theoretical the value of such a variable or parameter X is within the range $-\infty < X < \infty$. However, also the range of $0\% < X < 100\%$ can be observed in practice (e.g. average efficiency of a technical system). Proportions are calculated based on observations where an attribute of a specific parameter is either true or false. An appliance (e.g. a cook stoves or CFL) is either properly working and still in place or it has a malfunction and/or is no longer in operation. A cook stove is either fuelled with biomass or fossil fuel. In practice also proportions with more than two options are relevant (e.g. share of lighting technology used (incandescent lamps, compact fluorescent lamps, LED, etc.).

Figure 8: Characteristics of a population and a sample



Population:

- Number of elements (N)
- Mean (μ)
- Variance (σ^2)
- Standard deviation (σ)

Sample:

- Number of elements (n)
- Mean ($\bar{x}$)
- Variance (s^2)
- Standard deviation (s)

Please note: Symbols in parenthesis are used to denote the parameters in this manual but may differ in other publications.

[9] If not otherwise stated, the information provided in this chapter is adopted from Schwarze (2001) and Papula (1997)

Mean, variance and standard deviation

Figure 8 presents the main characteristics of a total population and a sample thereof. All elements (e.g. all households participating in the CDM project) comprises the total population. Among all these elements a parameter such as the operating hours of a light appliance may vary significantly. This variation is expressed by the variance and the standard deviation respectively. If one would do a complete census (i.e. monitoring every household in the total population the average operating hours (i.e. the mean), the variance and the standard deviation could be determined exactly applying the equations below while using all values of the total population. For the sample, the following equations are used to calculate the mean (see equation (1)), the variance (see equations (2) and (3)) and the standard deviation (see equation (4)) of a continuous variable (e.g. operating hours of a light appliance). Variance and standard deviation, describing the average deviation of the individual values compared to the mean value within one population or one sample are only applicable to a parameter of the type "mean".

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{1}$$

Where:

$\bar{x}$	=	Mean or average value
x_i	=	Value of each element *i*
n	=	Number of elements

$$s^2 = \frac{\sum_{i=1}^{N} (x_i - \bar{x})}{n} \tag{2}$$

Where:

s^2	=	Variance
$\bar{x}$	=	Mean or average value
x_i	=	Value of each element *i*
n	=	Number of elements

$$s^2 = \frac{\sum_{i=1}^{N} (x_i - \bar{x})}{(n-1)} \tag{3}$$

$$s = \sqrt{s^2} \tag{4}$$

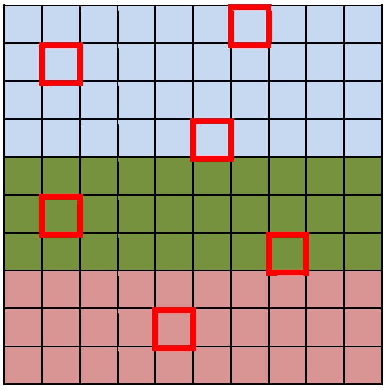
Where:

s	=	Standard deviation
s^2	=	Variance

The difference of the variance calculated by equation (2) and (3) is that equation (2) assumes a total population while equation (3) determines the variance of a sample that can be used to estimate the variance of the total population, meaning that the result of the equation represents an estimate for the variance of the total population. Hence, in the following examples equation (2) has been applied for the calculation of the variance of the total population and equation (3) has been applied to calculate the variance based on the sample as an estimate for the variance of the total population.

Estimation for the total population based on a sample

Figure 9: Mean of a population and the sample



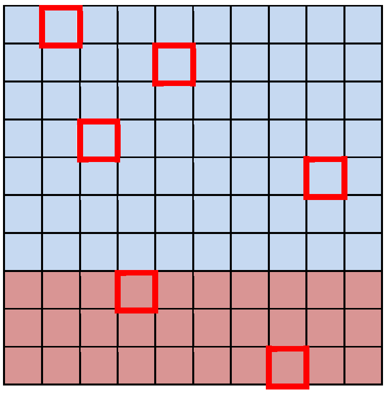
Population:
N = 100
μ = 2.95
σ^2 = 0.1725
σ = 0.4153

Sample:
n = 6
$\bar{x}$ = 2.8333
s^2 = 0.1666
s = 0.4082

Element that is operated 2.5 hours per day
Element that is operated 3 hours per day
Element that is operated 3.5 hours per day
Element that is sampled

From Figure 9 above and Figure 10 below it becomes obvious, that the characteristics of population can be estimated by the characteristics of the sample. However, the estimates from the sample are not identical to the real values of the population.

Figure 10: Proportion of a population and the sample



Population:
N = 100
$P_{(1)}$ = 70/100 = 0.7
$P_{(0)}$ = 30/100 = 0.3

Sample:
n = 6
$\hat{p}_{(1)}$ = 4/6 = 0.6666
$\hat{p}_{(0)}$ = 2/6 = 0.3333

Element that is fully functional (1)
Element that is not functional (0)
Element that is sampled

Level of confidence and precision

As the estimates derived from sampling are not exactly matching to the real parameter of the total population it is important to define indicators reflecting the accuracy of the sampling results compared to the real parameter. These indicators are the level of confidence and the precision.

Precision

In line with the S*tandard for sampling and surveys for CDM project activities and programme of activities* (in the following the *Sampling standard*) the precision in this manual is always applied as the relative precision if not otherwise stated. Table 2 provides an example of the application of precision.

Table 2: Precision according to the *Sampling standard*

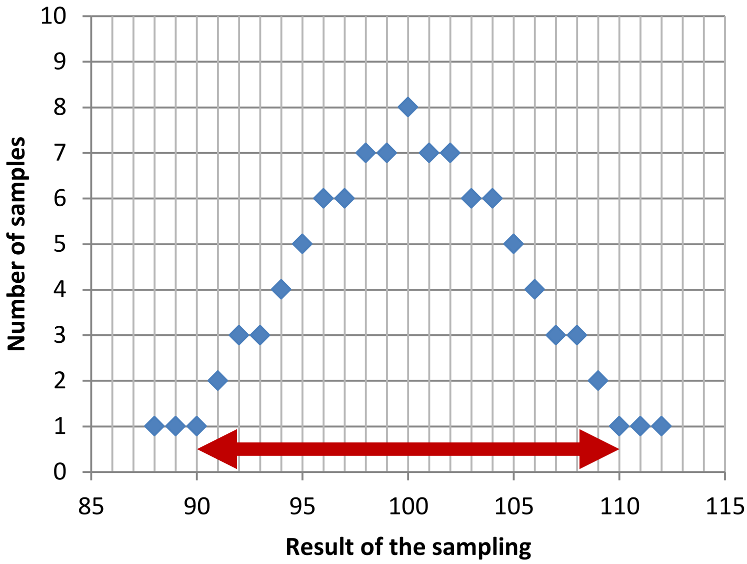
Precision	Lower bound	Parameter	Upper bound
	Determination of a mean		
10%	4.5 180	5 200	5.5 220
	Determination of a proportion		
10%	0.36 0.63	0.4 (40%) 0.7 (70%)	0.44 0.77

Level of confidence

The level of confidence can be interpreted as the frequency that a sample will derive the result within the accepted range of precision. Assuming you apply a hundred times a sampling of the same population. A level of confidence of 95% is reflected by the observation that 95 of the samples will obtain a result that is within the range of the required precision while only 5 out of 100 samples would return a value that is out of the pre-defined range. Figure 11 below visualises the concept of the level of confidence. Assuming a mean value of 100, a level of confidence of 96%[10] and a precision of 10%, the sampling repeated for one hundred times should return in 96 cases a value within the accepted range of the parameter (here: 90 to 110). Only in four cases the sampling would derive a value that is out of the accepted range.

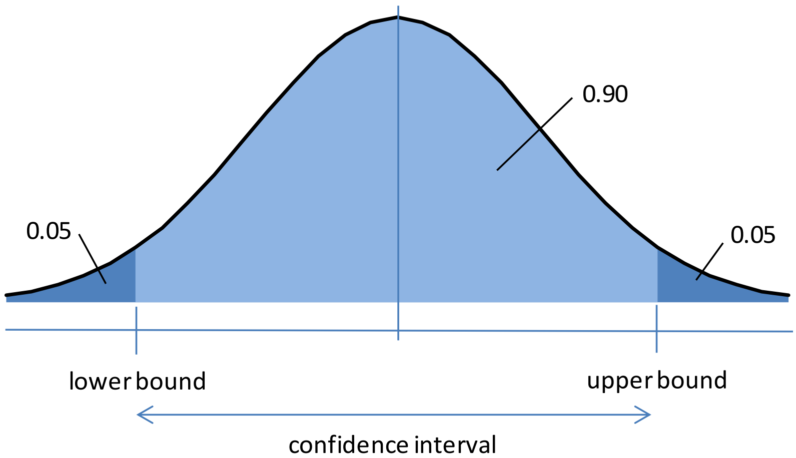
[10] A level of confidence of 96% has been selected exemplary to result in integers.

Figure 11: Sampling results based on a specific level of confidence/precision



The level of confidence or confidence interval can also be understood as area below the distribution curve as shown in Figure 12. The light blue area represents an area of 90% while the dark blue area comprises for the missing 10%.

Figure 12: Confidence interval of 90%



The lower and upper bound of the confidence interval for the determination of a mean value based on sampling is defined by equation (5).

$$\bar{x} - z\frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + z\frac{\sigma}{\sqrt{n}} \qquad (5)$$

Where:

$\bar{x}$	=	Mean or average value of the sample
z	=	Constant which relates to the level of confidence

n	=	Number of elements in the sample
σ	=	Standard deviation of the total population. If this parameter is unknown, it needs to be assumed (e.g. by the standard deviation of a sample)

The lower and upper bound of the confidence interval for the determination of a proportion based on sampling is defined by equation (6).

$$\hat{p} - \frac{z}{n}\sqrt{n\hat{p}(1-\hat{p})} \leq p \leq \hat{p} + \frac{z}{n}\sqrt{n\hat{p}(1-\hat{p})} \tag{6}$$

Where:

$\hat{p}$	=	Proportion determined by the sample
z	=	Constant which relates to the level of confidence
n	=	Number of elements in the sample
p	=	Proportion of the total population

The value of the constant *z* depends on the applied level of confidence (see Table 3 for often used levels of confidence).

Table 3: Constant that relates to the level of confidence

Level of confidence	Constant
90%	1.645
95%	1.96
~100%	~3.9

Constants related to the level of confidence

The constant is normally derived from the standard normal distribution. Common values applied under the CDM are 90% and 95% and can be found in Table 3. If required, other values would need to be derived from tables depicting the surface area below the standard normal distribution. One needs to be careful as these tables normally show the area of a one-sided interval and not as a two-sided interval as required. Equation (7) can be used to convert the two-sided interval into a one-sided interval. The value of the one-sided interval can afterwards be searched in the standard normal distribution table and the related constant can be found.

$$\text{interval}_{1s} = \frac{\text{interval}_{2s} + 1}{2} \tag{7}$$

Where:

interval_{1s}	=	One-sided interval that equals the required two-sided interval
interval_{2s}	=	Desired two-sided interval

For a two-sided confidence interval of 90% equation (7) returns a value of 0.95 as shown below:

$$\text{interval}_{1s} = \frac{\text{interval}_{2s} + 1}{2} = \frac{0.9 + 1}{2} = 0.95$$

The constant related to the interval or area of 0.95 can then be searched in the standard normal distribution table. The following values can be found:

0.9495 → 1.64

0.9505 → 1.65

Applying linear interpolation the following result can be derived:

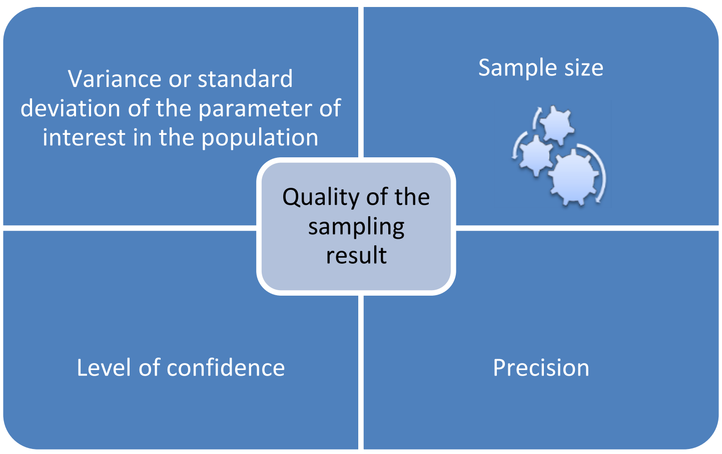
0.95 → 1.645

The value of 1.645 can then be used to calculate the required sample size as shown in the following chapters.

Determinants of the quality of a sample

The level of confidence and the precision of the sampling results are defined by CDM rulings (i.e. by methodologies and/or the *Sampling standard*). The standard deviation within the population under investigation is inherently predetermined by the characteristics of the population itself. If a proportion has to be determined by sampling, the standard deviation would be not applicable. Hence, the only adjustment that can be done by project participants is the number of elements in the sample size. This number has to be determined properly to ensure that the required quality of the sampling result is finally achieved (see Figure 13).

Figure 13: Determinants of the quality of the sampling result



Determination of the sample size

The determination of the sampling size requires basically an estimate/assumption for the expected mean and the variance (or standard deviation). In case of sampling to determine a proportion an estimate/assumption for the expected proportion is necessary. Different sampling approaches result in a different determination of the required minimal sampling size. Hence, the approach for determination of the minimal sampling size is provided in Chapter 3.3 for each of the specific sampling approaches.

3 Sampling in the CDM

Sampling requirements are defined by several UNFCCC documents

Several project types require sampling in order to calculate emission reductions. This is reflected by a considerably amount of rules developed by UNFCCC. There is not only one single comprehensive document existing that includes all the relevant rules regarding sampling. Hence, it is important to understand the different levels of ruling and the interaction between the different existing documents. This chapter provides the required information to understand and apply the existing regulatory framework for sampling in the CDM.

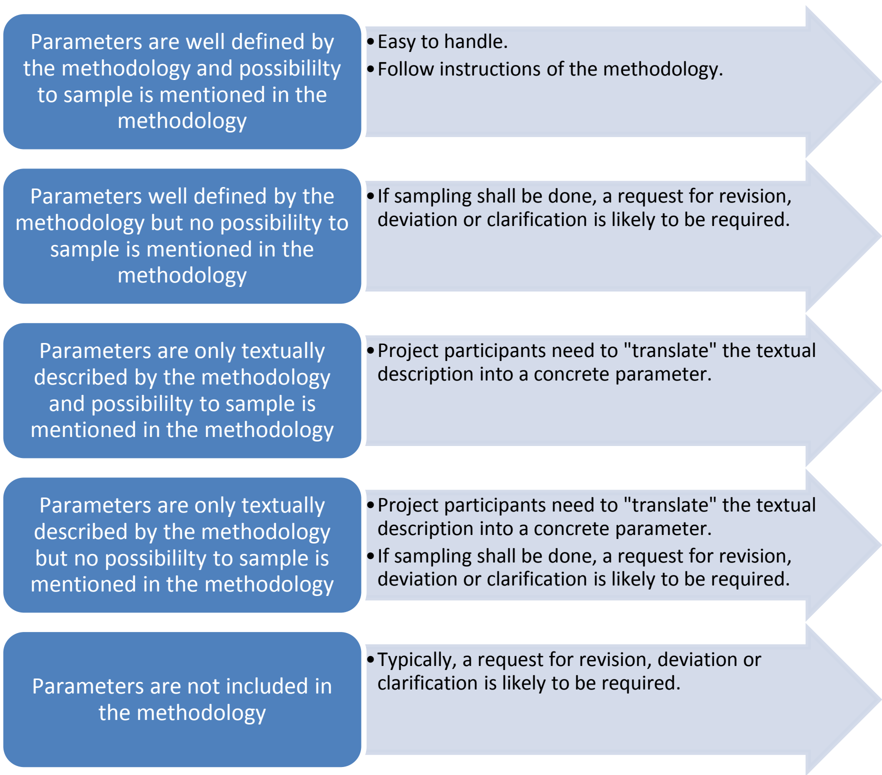
3.1 Types of parameters

Generally, two different types of parameters are applied in emission reduction calculations under the CDM:

- Data and parameters that are available at validation
- Data and parameters monitored

While the first type is determined only once at the start of the project activity (or at the start of each crediting period) the second type is monitored throughout the crediting period. The frequency of monitoring can vary typically from continuously to biennial (every two years) and is normally defined by the methodologies.

Figure 14: Determinants of the quality of the sampling result



Some of the parameters are very accurately defined by the methodology and some are not. For example, several methodologies require an annual or biennial check if the installed equipment is still in operation. However, no related parameter is defined and included in the equations provided for the emission reduction calculation. In such situations project participants need to define appropriate parameters on their own. Depending on the project situation and the applied methodologies some effort might be required by the project developer do adjust the parameters according to their specific project conditions (see Figure 14).

3.2 UNFCCC ruling

New CDM standards and procedures

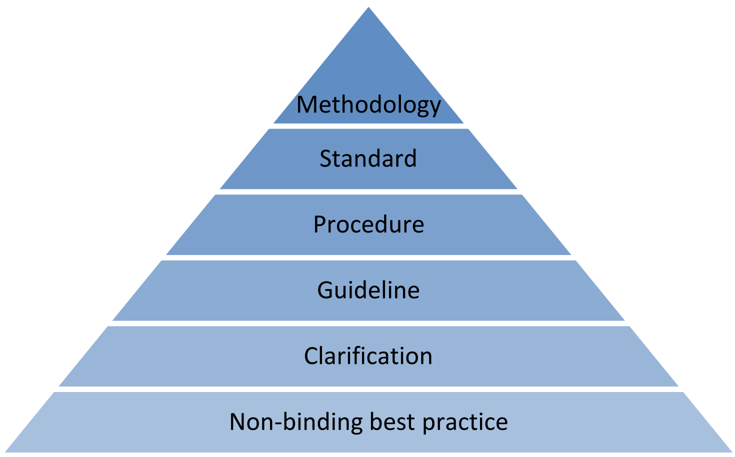
There are currently three core documents that can be seen as the general basis of the new UNFCCC regulatory framework. These documents are the

- Clean development mechanism project standard (PS);
- Clean development mechanism validation and verification standard (VVS);
- Clean development mechanism project cycle procedure (PCP).

Hierarchy of UNFCCC documents

Furthermore, there are various types of documents available from UNFCCC. It is important to comply with the hierarchy of the different documents. First of all, the applied methodology has the highest hierarchy, meaning that the provisions provided by the methodology overrule provisions of documents with lower hierarchy. If a methodology requests sampling at confidence/precision 99/10 then this has to be done accordingly even if the *Sampling standard* would require only 95/10 or 90/10. Guidelines and clarifications provide normally additional explanation on how the methodology or a standard has to be applied. Although, non-binding best practice examples are of low hierarchy, these types of documents provide very precise and usable guidance on how to apply provisions of methodologies and standards.

Figure 15: Hierarchy of documents



Source: Adopted from UNFCCC: Document hierarchy & types: available at: http://cdm.unfccc.int/Reference/index.html

3.2.1 Sampling according to methodologies

Methodologies are first source of information regarding sampling

The sampling requirements as provided by the applied methodology have to be followed. Even in a situation where these requirements differ from the sampling standard. There are several methodologies that provide specific information regarding the sampling requirements while other methodologies, although applicable to sampling approaches, do not specify any sampling provisions at all. In such cases the sampling standard provides information on the general sampling requirements.

There are about 20 methodologies that include direct or indirect elements of specific sampling requirements when applying the respective methodology. Several other methodologies may require sampling depending on the project situation although not explicitly stated within the methodology itself. In the following, only a selection of possible methodologies have been analysed regarding their sampling provisions. The selection has been made to meet the following criteria:

- Sampling is very likely to happen within a project activity applying the methodology
- The methodology is commonly used, especially for PoAs
- There is future potential for CDM projects

As a result of the criteria above, the methodologies provided in the table below are covered by this chapter.

Table 4: Methodologies that typically require sampling approaches

Meth #	Title
AMS-I.C	Thermal energy production with or without electricity
AMS-I.E.	Switch from Non-Renewable Biomass for Thermal Applications by the User
AMS-II.C.	Demand-side energy efficiency activities for specific technologies
AMS-II.G.	Energy efficiency measures in thermal applications of non-renewable biomass
AMS-II.J.	Demand-side activities for efficient lighting technologies

AMS-I.C. Thermal energy production with or without electricity[11]

AMS-I.C.

There are already 167 CDM projects and three PoAs registered applying AMS-I.C. Many more projects are currently under validation of which about 35 projects are PoAs[12]. The methodology is applicable to renewable energy technologies that supply users with thermal energy that displaces fossil fuel use including technologies such as solar thermal water heaters and dryers, solar cookers, energy derived from renewable biomass and other

[11] Information provided is valid for AMS-I.C. version 19.0

[12] As of March 2012. Source UNFCCC project databases available at: http://cdm.unfccc.int/Projects/projsearch.html

technologies that provide thermal energy that displaces fossil fuel. In the context of sampling the following technologies are most relevant:

- Solar water heaters/cookers
- Small biogas applications

According to the methodology sampling may be applied under specific conditions for the parameters presented in Table 5.

Table 5: Parameters of AMS-I.C. applicable to sampling/survey methods

Parameter	**Sampling provisions***
Continuous operation of the equipment/system	a, b
Net quantity of thermal energy supplied by the project activity	c
Quantity of electricity generated/supplied	c
Efficiency of the baseline units	b
Efficiency of the project equipment	b

**Sampling provisions according to the methodology:*
a = General guidelines for sampling and surveys for SSC project activities
b = Representative sample
c = Not explicitly specified

As the methodology does not specify confidence and precision explicitly for the parameters above, one has to assume the provisions of the *General guidelines for sampling and surveys for SSC project activities and the sampling standard* and the *Sampling standard* respectively. For small-scale CDM project activities this would mean a level of confidence/precision of 90/10.

As monitoring or metering of the produced electricity and/or thermal energy might be very cost intensive in case of a huge number of small appliances these parameters are often found to be replaced by less cost intensive parameters. For example, the operating hours might be monitored/metered. Operating hours multiplied by the capacity of the appliances then results in the generated energy.

AMS-I.E. Switch from non-renewable biomass for thermal applications by the user[13]

AMS-I.E.

There are eight CDM projects but no PoA registered applying AMS-I.E. Several more projects are currently under validation of which about 12 projects are PoAs[14]. The methodology is applicable to activities that displace the use of non-renewable biomass by

[13] Information provided is valid for AMS-I.E. version 4.0
[14] As of March 2012. Source UNFCCC project databases available at: http://cdm.unfccc.int/Projects/projsearch.html

introducing renewable energy technologies. Examples of these technologies include but are not limited to biogas stoves, solar cookers, passive solar homes, renewable energy based drinking water treatment technologies (e.g. sand filters followed by solar water disinfection; water boiling using renewable biomass). According to the methodology sampling may be applied under specific conditions for the parameters presented in Table 6.

Table 6: Parameters of AMS-I.E. applicable to sampling/survey methods

Parameter	Sampling provisions*
Continuous operation of the equipment/system	a, b
Quantity of woody biomass that is substituted or displaced	c
Efficiency of the baseline units	b
The use/diversion of non-renewable woody biomass (leakage)	c

**Sampling provisions according to the methodology:*
a = at least biennial (once every two years)
b = representative sample
c = not explicitly specified
d = confidence/precision 90/30

The methodology provides the following general information regarding sampling:

Representative sampling methods according to AMS-I.E.:

A statistically valid sample of the locations where the systems are deployed, with consideration, in the sampling design, of occupancy and demographics differences can be used to determine parameter values used to determine emission reductions, as per the relevant requirements for sampling in the *General guidelines for sampling and surveys for small-scale CDM project activities*. When biennial inspection is chosen a 95% confidence interval and a 5% margin of error requirement shall be achieved for the sampling parameter. On the other hand when the project proponent chooses to inspect annually, a 90% confidence interval and a 10% margin of error requirement shall be achieved for the sampled parameters. In cases where survey results indicate that 90/10 precision or 95/5 precision is not achieved, the lower bound of a 90% or 95% confidence interval of the parameter value may be chosen as an alternative to repeating the survey efforts to achieve the 90/10 or 95/5 precision.

Source: UNFCCC

The provisions provided by AMS-I.E. can be summarized as follows:

Table 7: Level of confidence and precision according to AMS-I.E.

Monitoring (frequency)	Sampling provisions*
Biennial (once every two years)	Confidence/precision = 95/5
Biennial not achieving required level of confidence and precision	Lower bound of Confidence/precision = 95/5
Annual	Confidence/precision = 90/10
Annual not achieving required level of confidence and precision	Lower bound of Confidence/precision = 90/10
Annual or biennial not achieving required level of confidence and precision (as an alternative)	Repetition of the survey to achieve related level of confidence/precision

Source: adopted from UNFCCC

AMS-II.C. Demand-side energy efficiency activities for specific technologies[15]

AMS-II.C.

There are ten CDM projects and 2 PoAs registered applying AMS-II.C. About 50 projects are currently under validation of which about 16 projects are PoAs[16]. The methodology is applicable to activities that encourage the adoption of energy-efficient equipment/appliance (e.g., lamps, ballasts, refrigerators, motors, fans, air conditioners, pumping systems) at many sites. These technologies may replace existing equipment or be installed at new sites. According to the methodology sampling may be applied under specific conditions for the parameters presented in Table 8.

Table 8: Parameters of AMS-II.C. applicable to sampling/survey methods

Parameter	Sampling provisions*
Number and power of replaced devices	b
Power, operating hours or energy use of installed devices	c
Continuous operation of the equipment/system	a

**Sampling provisions according to the methodology:*
a = annual
b = allowing verification by the DOE
c = not explicitly specified

[15] Information provided is valid for AMS-II.C. version 13.0
[16] As of March 2012. Source UNFCCC project databases available at: http://cdm.unfccc.int/Projects/projsearch.html

As the methodology does not specify confidence and precision explicitly for the parameters above one has to assume the provisions of the *General guidelines for sampling and surveys for SSC project activities and the sampling standard* and the *Sampling standard* respectively. For small-scale CDM project activities this would mean a level of confidence/precision of 90/10.

AMS-II.G. Energy efficiency measures in thermal applications of non-renewable biomass[17]

AMS-II.G.

There are four CDM projects and three PoAs registered applying AMS-II.G. Many more projects are currently under validation of which about 30 are PoAs[18]. The methodology is applicable to efficiency improvements in thermal applications of non-renewable biomass. Examples of these technologies and measures include the introduction of high efficiency biomass fired cook stoves or ovens or dryers and/or improvement of energy efficiency of existing biomass fired cook stoves or ovens or dryers. According to the methodology sampling may be applied under specific conditions for the parameters presented in Table 9. These are similar to the parameters according to AMS-I.E (see above).

Table 9: Parameters of AMS-II.G. applicable to sampling/survey methods

Parameter	**Sampling provisions***
Continuous operation of the equipment/system	a, b
Checking efficiency of the equipment/system	a, b
Quantity of woody biomass used in the absence of the project activity	c
Efficiency of the baseline units	b
The use/diversion of non-renewable woody biomass (leakage)	c

**Sampling provisions according to the methodology:*
a = at least biennial (once every two years)
b = representative sample
c = not explicitly specified
d = confidence/precision 90/30

The methodology provides the following general information regarding sampling:

[17] Information provided is valid for AMS-II.G. version 3.0
[18] As of March 2012. Source UNFCCC project databases available at: http://cdm.unfccc.int/Projects/projsearch.html

Representative sampling methods according to AMS-II.G.:

A statistically valid sample of the locations where the systems are deployed, with consideration, in the sampling design, of occupancy and demographics differences can be used to determine parameter values used to determine emission reductions, as per the relevant requirements for sampling in the *General guidelines for sampling and surveys for small-scale CDM project activities*. When biennial inspection is chosen a 95% confidence interval and a 5% margin of error requirement shall be achieved for the sampling parameter. On the other hand when the project proponent chooses to inspect annually, a 90% confidence interval and a 10% margin of error requirement shall be achieved for the sampled parameters. In cases where survey results indicate that 90/10 precision or 95/5 precision is not achieved, the lower bound of a 90% or 95% confidence interval of the parameter value may be chosen as an alternative to repeating the survey efforts to achieve the 90/10 or 95/5 precision.

Source: UNFCCC

The provisions provided by AMS-II.G. can be summarized as follows:

Table 10: Level of confidence and precision according to AMS-II.G.

Monitoring (frequency)	Sampling provisions*
Biennial (once every two years)	Confidence/precision = 95/5
Biennial not achieving required level of confidence and precision	Lower bound of Confidence/precision = 95/5
Annual	Confidence/precision = 90/10
Annual not achieving required level of confidence and precision	Lower bound of Confidence/precision = 90/10
Annual or biennial not achieving required level of confidence and precision (as an alternative)	Repetition of the survey to achieve related level of confidence/precision

AMS-II.J. Demand-side activities for efficient lighting technologies

AMS-II.J.

There are already 16 CDM projects and two PoAs registered applying AMS-II.J. About 60 more projects are currently under validation of which about 12 are PoAs[19]. The methodology is applicable to activities that lead to efficient use of electricity through the adoption of self-ballasted compact fluorescent lamps (CFLs) to replace incandescent lamps (ICLs) in residential applications. Eligible self-ballasted CFLs have integrated ballasts as a non-removable part. The CFLs adopted to replace existing equipment must be new equipment and not transferred from another activity. According to the methodology sampling may be applied under specific conditions for the parameters presented in Table 11.

[19] As of March 2012. Source UNFCCC project databases available at: http://cdm.unfccc.int/Projects/projsearch.html

Table 11: Parameters of AMS-II.J. applicable to sampling/survey methods

Parameter	Sampling provisions*
Operating hours	a
Lamp failure rate	a

**Sampling provisions according to the methodology:*
a = Please consult the methodology

The methodology provides the following general information regarding sampling:

Representative sampling methods according to AMS-II.G.:

The following survey principles shall be followed for activities related to determining number of CFLs placed in service and operating under the project activity and, if required, determining the number of operating hours of baseline and project lamps:

- The sampling size is determined by minimum 90% confidence interval and the 10% maximum error margin; the size of the sample shall be no less than 100;
- Sampling must be statistically robust and relevant ,i.e. the survey has a random distribution and is representative of target population (size, location);
- The method to select respondents for interviews is random;
- The survey is conducted by site visits;
- Only persons over age 12 are interviewed;
- The project document must contain the design details of the survey.

Source: UNFCCC

3.2.2 Sampling according to the standard for sampling and surveys

The *Sampling standard*

The *Standard for sampling and surveys for CDM project activities and programme of activities* (in the following the *Sampling standard*) can be seen as the most important source of ruling referring to sampling and surveys in the context of CDM. The version 02.0 of the *sampling standard* has been adopted by EB65 in November 2011. It is valid for CDM project activities and programme of activities and defines at the same time rules for the validation and verification of CDM projects and PoAs by DOEs. For projects applying small scale methodologies the sampling requirements are set to be 90/10 (confidence/precision) and for large scale methodologies 95/10. Nevertheless, sampling requirements defined by a specific methodology are overruling and can be seen as the minimum requirement that has to be met. The *sampling standard* consists of the following main sections besides a description of the background, scope and applicability (Section I. and II.):

- Section III: Sampling Requirements
- Section IV: Sampling Requirements for PoAs
- Section V: Validation and Verification of Sampling Plans
- Appendix 1: Essential Sampling Terminology
- Appendix 2: Common Types of Sampling Approaches
- Appendix 3: Recommended outline of a Sampling Plan
- Appendix 4: Recommended Practices for Unbiased Estimates
- Appendix 5: Recommended Evaluation Criteria for Validation

Section III: Sampling Requirements

Sampling requirements

The *Sampling standard* defines the following requirements in case sampling is used to obtain a (mean) value of a parameter:

- CDM-PDDs and CDM-PoA/CPA-DDs shall include a sampling plan;
- The value of the parameter obtained by sampling shall be an unbiased and reliable estimate;
- Sampling requirements by a specific methodology applied for the project activity have precedence over the Sampling standard;
- In the absence of specific guidance by the applied methodology the level of confidence and precision as shown in Table 12 below shall be applied;
- Precision shall be interpreted as relative precision (see Table 2 in Chapter 2);
- The sample size shall be calculated to meet the minimum requirements (i.e. level of confidence and precision);
- If sampling is used to determine a parameter that shall, according to the methodology, not be obtained from sampling, project participants are requested to submit a request for revision, deviation or clarification;
- If not otherwise stated by the applied methodology, the sample mean or proportion shall be used for the emission reduction calculations and not the lower or upper bound of the confidence interval.

Table 12: Level of confidence and precision according to the *Sampling standard.*

Type of project	Level of confidence and precision
Small scale CDM/PoA	90/10
Large-scale CDM/PoA	95/10
Small-scale PoA if a group of CPAs is covered by one single sampling plan	95/10
Large-scale PoA if a group of CPAs is covered by one single sampling plan	Not allowed

Source: based on the Sampling standard

Section IV: Sampling Requirements for PoAs

Sampling requirements for PoAs

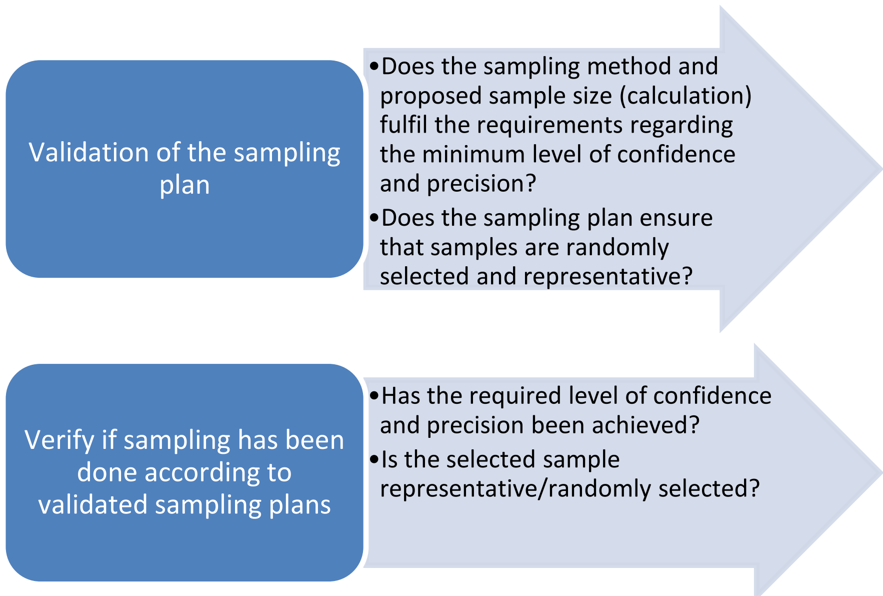
This section provides information regarding sampling options within a PoA. Sampling can be done either by sampling of a parameter separately for each CPA or, in case of applying small-scale methodologies, by sampling a parameter within a group of CPAs. Sampling based on a group of CPAs is referred to as cross-CPA-sampling within this manual. Table 12 above presents the related required level of confidence and precision. Cross-CPA-sampling shall be done only if homogeneity related to a specific parameter across different CPAs can be demonstrated. Otherwise, the differences need to be taken into account in the calculation of the sample size.

Section V: Validation and Verification of Sampling Plans

Validation and verification of sampling plans

The requirements for DOEs while checking the sampling plan are outlined in this section of the *Sampling standard*. DOEs are requested to validate/verify the sampling plan as shown in Figure 16.

Figure 16: Validation and verification of sampling plans



Furthermore, the Sampling standard allows DOEs applying sampling approaches during validation/verification if the project developer has not applied sampling. This might be relevant for multi-site CDM project activities or PoAs. Please consult the Sampling standard for more details on this approach.

Appendix 1: Essential Sampling Terminology

Sampling terminology

Appendix 1 provides definitions and explanation for the following terms:

- Sample / population;
- Parameter;

- Representative;
- Types of data (mean vs. proportion);
- Continuous variable
- Attributes (yes or no data or binary data);
- Unbiased and reliable estimates.

Please consult the *Sampling standard* for further details on the terminology.

Appendix 2: Common Types of Sampling Approaches

Sampling approaches

Appendix 2 describes the different sampling approaches such as simple random sampling, stratified sampling, systematic sampling, cluster sampling or multi-stage sampling. Please refer to Chapter 3.3 of this manual for further details on the different approaches.

Appendix 3: Recommended outline of a Sampling Plan

Outline of the sampling plan

This appendix of the *Sampling standard* provides the recommended outline of a sampling plan. A summary of the outline of a sampling plan and what would be required to be included in a sampling plan outline can be found below:

Recommended outline for a Sampling Plan according to the *Sampling standard.*

- Sampling Design:
  - Objectives and Reliability Requirements
    - Objective of the sampling effort
    - Timeframe
    - Estimated parameter value(s)
    - Confidence/precision
  - Target Population
  - Sampling Method (e.g. simple random or stratified sampling)
  - Sample Size
  - Sampling Frame
- Data:
  - Field Measurements
    - Identify all variables
    - Timing and frequency of measurements
    - Conservative and necessary corrections
  - Quality Assurance/Quality Control
    - Training of staff
    - Non-response
    - Handling of outliers
  - Analysis
    - How will the data be used
- Implementation:
  - Implementation plan
    - Define a schedule for implementing the sampling effort

Source: adopted from UNFCCC

Appendix 4: Recommended Practices for Unbiased Estimates

Unbiased estimates

In order to achieve unbiased estimates Appendix 4 provides recommended practices for the following topics:

- Defining precisely the sampling objectives and target population and the measurements to be taken and/or data collected;
- Deciding on the sampling design and the size of the sample;
- Developing the sampling frame;
- Randomizing cases and drawing sample;
- Selecting the most effective information-gathering method;
- Conducting surveys/measurements;
- Minimizing non-response and adjusting for its effects;

Further explanation and recommendations in order to achieve unbiased estimates can be found in Chapter 1 of this manual.

Appendix 5: Recommended Evaluation Criteria for Validation

Evaluation criteria for validation

Appendix 5 of the *Sampling standard* provides questions that need to be answered by the DOE during validation. A summary of the questions can be found below:

Recommended evaluation criteria for DOE validation according to the *Sampling standard*.

- Does the sampling plan present a reasonable approach for obtaining unbiased, reliable estimates of the variables?
- Is the population clearly defined, and how well does the proposed approach to developing the sampling frame represent that population?
- Is the proposed sampling approach clear?
- Is the proposed sample size adequate to achieve the minimum confidence/precision requirements? Is the ex-ante estimate of the population variance needed for the calculation of the sample size adequately justified?
- Is the sample representative?
- Is the data collection/measurement method likely to provide reliable data given the nature of the parameters of interest and project, or is it subject to measurement errors?
- Are the procedures for the data measurements well defined and do they adequately provide for minimizing non-sampling errors?
- Does the frame contain the information necessary to implement the sampling approach?

Source: UNFCCC

3.2.3 Sampling according to guidelines and best practice examples

Besides the sampling requirements defined by specific methodologies and by the *Sampling standard*, there are some few more guidelines that deal with the topic of sampling. These are briefly summarised in the following section.

Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation[20]

Practical application of the *Sampling standard*

While the *Sampling standard* defines some general requirements regarding sampling this document guides project developers on how to fulfil and practically implement the requirements. The document focuses on examples in order to illustrate the calculation of the sampling size. An overview of the examples of the document can be found in Table 13.

Table 13: Examples of sample size calculations according to *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation*

No.	Scale	Parameter	Technology	Sampling approach
1-4	SSC	Continued use of appliances (Proportion)	Cook stoves	1: Simple random sampling 2: Stratified random sampling 3: Cluster sampling 4: Multi-stage sampling
5-8	SSC	Operating hours (Mean)	CFLs	5: Simple random sampling 6: Stratified random sampling 7: Cluster sampling 8: Multi-stage sampling
9	SSC	Strength of bricks (Mean)	Brick production	9: Systematic sampling
10	SSC	COD content of wastewater (Mean)	Waste water/ biogas	10: Systematic sampling
-	SSC	-	-	Other calculated sample sizes
11-14	LSC	Proportion of passengers (Proportion)	Public transport	11: Simple random sampling 12: Stratified random sampling 13: Cluster sampling 14: Multi-stage sampling
15-18	LSC	Journey length (Mean)	Public transport	15: Simple random sampling 16: Stratified random sampling 17: Cluster sampling 18: Multi-stage sampling
19	LSC	Journey time (Mean)	Public transport	19: Systematic sampling

Source: Adopted from UNFCCC

In addition, the *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation* provides a chapter on the determination of the sampling size in case of the following conditions:

- The total number of elements in the population (N) is low;
- The proportion is very low or very high.

[20] As of March 2012 this document is only a draft and a call for public input has been launched on this document. Only after approval by the EB, the information provided by the document can be seen as coming into force.

In such situations project participants are requested to check if a special approach to calculate the sample size has to be applied. Such a special approach is required if one of the following equations result in a value lower than 10.

$$N * p > 10 \quad (8)$$

$$N * (1 - p) > 10 \quad (9)$$

Where:

N	=	Total number of elements in the population
p	=	Proportion

If one of the equations above results in a value lower than 10, a different approach shall be applied for the determination of the sample size. Please consult the related document for further guidance.

Reliability of the sampling results

Furthermore, a chapter on reliability calculations is included in the *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation.* The sample size is calculated based on an estimation/expectation of the proportion or the mean value and the standard deviation of a specific parameter. These assumptions of expected proportions, means or standard deviations might be wrong. As a result, the sampling size would be over or underestimated and the value of a parameter derived by sampling would be over or underestimated, too. The chapter on reliability calculations provides some approaches to check the reliability of the sampling results. Please consult the related document for further guidance.

General guidelines for sampling and surveys for small-scale CDM project activities

The *Sampling standard* supersedes the *General guidelines for sampling and surveys for small-scale CDM project activities*. Hence, it is important to apply the provisions of the *sampling standard* for new project activities especially in regard to the sampling plan requirements in order to comply with the currently approved rulings. Nevertheless, the *general guidelines* provide some useful information (e.g. description of sampling approaches, list of references (literature) and some brief information on sampling software). Furthermore, the guidelines remain important for projects that have been registered referring to this document.

Draft non-binding best practices examples to illustrate the application of sampling guidelines

This document provides an evaluation of the different sampling approaches (i.e. advantages and disadvantages) and formulas for the calculation of the sampling size. As the document is only a draft (SSC WG 30) the document can only be used as assistance or additional source of information. It is unlikely that the document will become approved as a separate document as most of the content is also included in the *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation* as described above.

Documents under development

Further documents are currently under development by UNFCCC including:

- Methods, if any, to deal with missed reliability targets without compromising conservative estimates for emission reduction;
- Best practice examples for DOE validation/verification for sampling and surveys.

These documents can be expected to provide very important input on sampling in practice.

3.3 Sampling approaches applied in the CDM

Applied sampling approaches in the CDM

Sampling approaches in the CDM follow approaches that originate mainly from population census, scientific experiments (e.g. effectiveness of medicine) and quality assurance (e.g. control of produced goods to fulfil a specific requirement). The *Sampling standard* includes specifically the following sampling approaches which are described in more detail in the following section:

- Simple random sampling
- Stratified random sampling
- Systematic sampling
- Cluster sampling
- Multi-stage sampling

There are several more approaches available that might gain importance in the CDM in the future. However, these are not subject of this manual. Among the above mentioned approaches, simple random sample is most often used.

3.3.1 Simple random sampling

Simple random sampling

Simple random sampling can be seen as the most commonly used and most simple sampling approach. In applying the approach, project participants do not need to stratify or cluster the total population. The sample is directly drawn from the sampling frame which is in most cases identical to the total population. If there is no homogeneity within the total population related to a specific parameter (i.e. high variance / standard deviation) the approach might result in a higher sample size as compared to other approaches. However, a correctly calculated sampling size can be expected to achieve the required results. The description of the sampling approach according to the *Sampling standard* can be found below.

Simple random sample according to the *Sampling standard*

A simple random sample (SRS) is a subset of a population (e.g. villages, individuals, buildings, pieces of equipment) chosen randomly, such that each element (or unit) of the population has the same probability of being selected. The sample-based estimate (mean or proportion) is an unbiased estimate of the population parameter. SRS is conceptually straightforward and easy to implement - provided that a sampling frame of all elements of the population exists. Its simplicity makes it relatively easy to analyze the collected data. It is also appropriate when only minimum information of the population is known in advance of the data collection. SRS is suited to populations that are homogeneous. In many instances a large population size and dispersed nature of population may cause a lack of homogeneity, while in some cases those factors may have relatively low impact on homogeneity (e.g. a large number of biogas digesters located in varying altitudes and temperature zones may be less conducive for SRS to determine the average amount of biogas production per digester, while the usage hours of light bulbs across wide geographic areas and among large populations with similar socioeconomic circumstances connected to a single or similar grid/s may be sufficiently homogeneous for SRS). The costs of data collection under SRS could be higher than other sampling approaches when the population is large and geographically dispersed.

Source: adopted from UNFCCC

Determination of the sample size for simple random sampling

In order to calculate the sample size (n) required for simple random sampling project participants need an estimate or assumption for the following parameters:

- Total number of elements in the population / the sampling framework (N);
- Expected proportion (p) or mean;
- Expected standard deviation (SD) if the parameter of interest is a mean.

Determination of the sample size for simple random sampling

According to the *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation*, the following equations shall be used to determine the required sample size and in case the parameter of interest is a proportion (e.g. proportion of stoves still in operation).

$$n \geq \frac{z^2 \times N \times V}{(N-1) \times precision^2 + z^2 \times V} \tag{10}$$

Where:

$$\mathrm{V} = \frac{p \times (1-p)}{p^2} \tag{11}$$

n	=	Number of elements to be sampled
N	=	Total number of elements in the population
p	=	Proportion
z	=	Constant referring to the level of confidence (e.g. 1.645 for 90 % confidence and 1.96 for 95 % confidence.
precision	=	Required precision (e.g. 10% = 0.1)

If the parameter of interest is a mean value (e.g. average number of operating hours of CFLs) the following equations have to be applied:

$$n \geq \frac{z^2 \times N \times V}{(N-1) \times precision^2 + z^2 \times V} \quad (12)$$

Where:

$$V = \left(\frac{SD}{mean}\right)^2 \quad (13)$$

n	=	Number of elements to be sampled
N	=	Total number of elements in the population
mean	=	Average value of the parameter that is expected in the total population
SD	=	Standard deviation of the parameter that is expected in the total population
z	=	Constant referring to the level of confidence (e.g. 1.645 for 90 % confidence and 1.96 for 95 % confidence.
precision	=	Required precision (e.g. 10% = 0.1)

The formulas from above result in the minimum required sample size (i.e. number of elements to be sampled). Project participants are recommended to increase the number of elements to be sampled in order to compensate for non-response effects if applicable.

Approximate equation to calculate the sample size

The *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation* provides also an approximate equation were the actual size of the total population (N) can be ignored under specific circumstances. However, as the total number of elements should be well known in almost all cases this simplification is normally not required.

Examples for simple random sampling

Examples 1, 5, 11 and 15 of the *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation* demonstrate the calculation of the sampling size for a simple random sampling approach.

3.3.2 Stratified random sampling

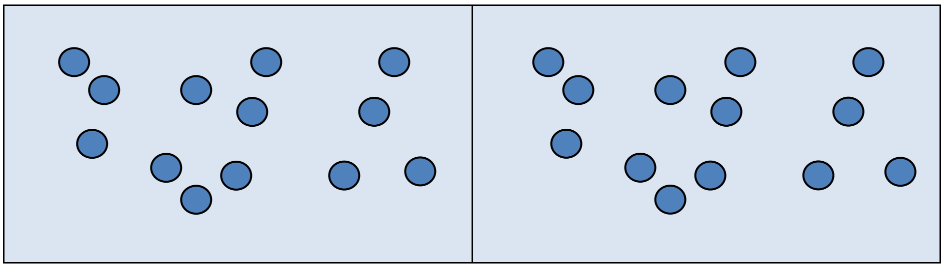
Stratified random sampling

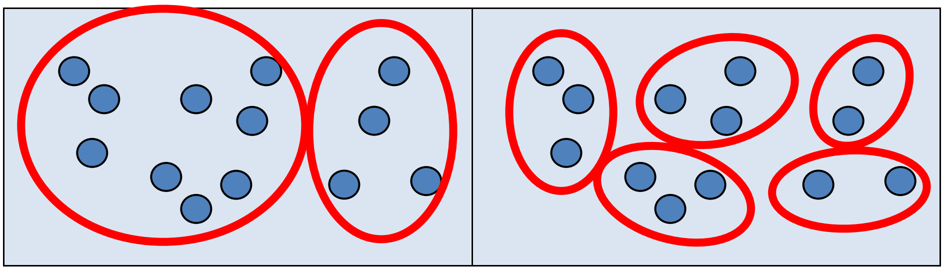
A stratified random sampling approach requires some kind of partitioning or categorization. It must be possible to allocate all elements in the total population or sampling frame uniquely to this stratification (i.e. each element belongs to one and only one stratum). One example for such stratification could be the allocation of each household in the total population or sampling frame to only one district or town. The different districts or towns are then the strata for the sampling approach. The approach of stratification, grouping or clustering is similar for the stratified random and the cluster sampling. Nevertheless, there are very significant differences as shown in Figure 17 below.

Comparison of stratified and cluster sampling

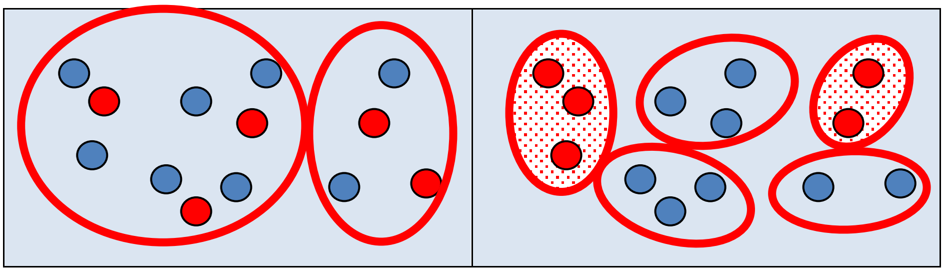
Figure 17: Comparison of stratified and cluster sampling

Stratified sampling	**Cluster sampling**

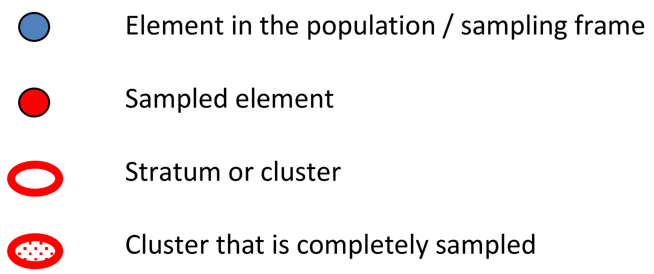




The total population is grouped into stratas (in this example only 2 stratas are depicted).	The total population is grouped into clusters (in this example only 5 clusters are depicted).



The required number of samples is drawn from each strata. The number of sampled elements per strata is proportional to the number of elements in each strata.	Instead of drawing the sample from the elements, the sample as drawn from the clusters. Within a sampled cluster, all elements have to be sampled.



As a result of the differences as described above, the number of strata or clusters and the number of elements in each stratum/cluster are typically behaving contrarily (see Table 14). Cluster sampling with very few clusters (e.g. 5 clusters) and a huge number of elements in each cluster (e.g. several hundred elements per cluster) would not work properly, as drawing a sample out of only 5 clusters would be problematic and the sampling of hundred elements of each selected clusters would not be beneficial. A stratified sampling approach with a huge number of strata and a small number of elements would not be helpful as well.

Table 14: Number of strata/clusters and number of elements

Sampling approach	Number of strata/clusters	Number of elements within each strata/cluster
Stratified random sampling	low	high
Cluster sampling	high	low

The description of the stratified random sampling approach according to the *Sampling standard* can be found below.

Stratified random sample according to the *Sampling standard*

When the population under study is not homogeneous but instead consists of several subpopulations which are known (or thought) to vary, then it is better to take a simple random sample from each of these sub-populations separately. This is called stratified random sampling. The subpopulations are called the strata. When considering stratified random sampling it is important to note that when identifying the strata no population element can be excluded and every element must be assigned to only one stratum. For example, the population of participants in a commercial lighting programme might be grouped according to building type (e.g. restaurants, food stores, and offices). Stratified random sampling is most applicable to situations where there are obvious groupings of population elements whose characteristics are more similar within groups than across groups (e.g. restaurants are likely to be more similar to one another in terms of lighting use than they are to offices or food stores). It requires that the grouping variable be known for all elements in the sampling frame. For example, the sampling frame would require information on the building type for each case in the population to allow stratification by that characteristic. Stratification helps to ensure that estimates of a population characteristic are accurate, especially if there are differences amongst the strata. For example, if lighting use within office buildings tends to be lower (on average) than in food stores then this can be taken into account when estimating the overall average number of hours of operation. Equally, if the cases within each stratum are more homogeneous than across strata, then the estimated number of hours of operation will be more precise than if a simple random sample of the same size had been taken.

Source: UNFCCC

Determination of the sample size for stratified random sampling

Determination of the sample size for stratified random sampling

For the calculation of the required sample size for a stratified random sampling approach the following steps shall be conducted:

1. Stratification of the total population;
2. Calculation of the overall proportion and standard deviation or the overall mean and standard deviation;
3. Calculation of the overall required sample size;
4. Determination of the sample size within each stratum.

Step 1: Stratification of the total population:

A stratified sampling approach is only meaningful, if the value of a parameter is different for different strata (e.g. the proportion of cook stoves that is found still in operation is different in the different strata). The following tasks shall be done for the stratification:

- Define strata (e.g. geographical regions such as districts or towns but also different types of elements, e.g. different types of buildings, vehicles etc.);
- Allocate all elements in the total population to the strata and check if every element is included in one (and only one) stratum;
- Estimate the characteristics for each stratum:
    - For a proportion parameter: the expected proportion for each stratum;
    - For a mean parameter: the expected mean value and standard deviation for each stratum.

Step 2: Calculation of the overall proportion and standard deviation or the overall mean and standard deviation:

Depending on the type of parameter (proportion or mean), the following values have to be calculated based on the estimates of step 1:

- Proportion: Overall proportion and overall standard deviation;
- Mean: Overall mean and overall standard deviation.

The calculation takes into account the number of elements of each stratum. The formulas that need to be applied can be found in the *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation* (equations as shown in Table 15).

Table 15: Equations used for stratified random sampling

Type of parameter	Parameter	Equation numbers
Proportion	Overall proportion	7 (SSC) and 46 (LSC)
	Overall standard deviation	6 (SSC) and 45 (LSC)
Mean	Overall mean	24 (SSC) and 61 (LSC)
	Overall standard deviation	23 (SSC) and 63 (LSC)

Equation numbers refer to the Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation

Step 3: Calculation of the overall required sample size:

The overall sample size is then calculated similar to the random sampling approach but taking the parameters of the overall proportion or mean and the overall standard deviation into account.

$$n \geq \frac{z^2 \times N \times V}{(N-1) \times precision^2 + z^2 \times V} \qquad (14)$$

Where:

$$\mathrm{V} = \left(\frac{SD}{mean}\right)^2 \quad \text{or} \quad \mathrm{V} = \left(\frac{SD}{\bar{p}}\right)^2 \qquad (15)$$

n	=	Number of elements to be sampled
N	=	Total number of elements in the population
mean	=	Overall mean
$\bar{p}$	=	Overall proportion
SD	=	Overall standard deviation
z	=	Constant referring to the level of confidence (e.g. 1.645 for 90 % confidence and 1.96 for 95 % confidence.
precision	=	Required precision (e.g. 10% = 0.1)

Step 4: Determination of the sample size within each stratum:

Based on the number of elements in each stratum, the total number of required samples is dived to the different strata. The resulting sample size for each stratum shall be rounded up.

$$n_i = \frac{g_i}{N} \times n \qquad (16)$$

Where:

n_i	=	Number of elements to be sampled in the stratum i
g_i	=	Number of elements in the stratum i
n	=	Total number of elements to be sampled
N	=	Total number of elements in the population

Examples for stratified random sampling

Examples 2, 6, 12 and 16 of the *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation* demonstrate the calculation of the sampling size for a stratified random sampling approach.

3.3.3 Systematic sampling

Systematic sampling

This sampling approach can be typically found as quality assurance approach for checking of production outputs. Under the CDM, the approach could be used as follows:

- As alternative to a simple random sampling approach: Instead of using a random number generator if it is ensured that the list of elements is randomly composed;
- To determine a parameter of a production output such as quality of produced bricks, efficiency of manufactured cooking stoves, etc.;
- To determine a parameter based on periodical measurements/monitoring (e.g. concentration of a specific substance in a continuous flow such as wastewater or landfill gas).

The description of the systematic sampling approach according to the *Sampling standard* can be found below.

Systematic sampling according to the *sampling standard*

Systematic sampling is a statistical method involving the selection of elements from an ordered sampling frame. The most common form of systematic sampling is an equal-probability method, in which every k^{th} element in the frame is selected, where k, the sampling interval (sometimes known as the "skip"), is calculated as:

k = population size (N) / sample size (n)

Using this procedure, each element in the population has a known and equal probability of selection. The project participant shall ensure that the chosen sampling interval does not hide a pattern. Any pattern would threaten randomness. A random starting point must also be selected. Systematic sampling is to be applied only if the given population is logically homogeneous, because systematic sample units are uniformly distributed over the population. Systematic sampling is applicable in a number of situations. If there is a natural ordering or flow of subjects in the population, such as output of bricks in a manufacturing process, then it is typically easier to sample every k^{th} unit to test for quality as they are produced. In all cases, it is important that the list of subjects or the process is naturally random, in the sense that there is no pattern to its order.

Source: UNFCCC

Determination of the sample size for systematic sampling

Determination of the sample size for systematic sampling

The sample is derived by applying the following two steps:

1. Calculation of the required sample size;
2. Calculation of the "skip".

Typically, systematic sampling is applied for determination of mean values of parameters as it is normally applied to monitor characteristics of a production output (e.g. produced bricks). As a continuous production output would normally result in a huge number of elements (e.g. produced bricks per year) the approximate formula according to the *Draft*

Best Practices Examples: Focusing on Sample Size and Reliability Calculation can be applied (Assumption N >5,000). For step 1 the following equation can be applied:

$$n \geq \frac{z^2 \times V}{precision^2} \tag{17}$$

Where:

$$V = \left(\frac{SD}{mean} \right)^2 \tag{18}$$

n	=	Number of elements to be sampled
mean	=	Average value of the parameter that is expected in the total population
SD	=	Standard deviation of the parameter that is expected in the total population
z	=	Constant referring to the level of confidence (e.g. 1.645 for 90 % confidence and 1.96 for 95 % confidence.
precision	=	Required precision (e.g. 10% = 0.1)

The determination of the “skip” is done according to equation (19).

$$skip \leq \frac{N}{n} \tag{19}$$

Where:

skip	=	Determiner that defines frequency of sampling (i.e. every k^{th} element will be sampled)
n	=	Number of elements to be sampled
N	=	Total number of elements (e.g. number of produced bricks,

In case of sampling a continuous flow (e.g. waste water or landfill gas) a different approach has to be applied in order to calculate the sampling frequency. Example 10 of the *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation* provides the related approach.

Examples for systematic sampling

Examples 9, 10 and 19 of the *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation* demonstrate the calculation of the sampling size for a systematic sampling approach.

3.3.4 Cluster sampling

Cluster sampling

Similar to stratified random sampling, cluster sampling requires subdividing the total population into subgroups (strata in case of stratified sampling, clusters in case of cluster sampling)[21]. Typically, all clusters should be of approximate the same size (i.e. number of elements in each cluster). After this step of disaggregation, the clusters can be seen as new elements from which the sample is drawn. The description of the cluster sampling approach according to the *Sampling standard* can be found below.

Cluster sampling according to the *sampling standard*

Clustered sampling refers to a technique where the population is divided into sub-groups (clusters), and the sub-groups are randomly selected (sampled), rather than the individual elements which are to be studied. The data are then collected on all the individual elements in the selected sub-groups. Cluster sampling is used when "hierarchical" groupings are evident in a population, such as villages and households within villages, or buildings and appliances within buildings. For example, suppose a project installs high-efficiency motors in new apartment buildings, with several motors typically in each building. In order to estimate the operating hours of the motors, one might take a sample of the buildings instead of the motors, and then meter all of the motors in the selected buildings. In contrast to stratified sampling, where the equipment of interest is grouped into a relatively small number of homogeneous segments, there are many clusters of motors (i.e. apartment buildings), and there is no expectation that the motors in each building are more homogeneous than the overall population of efficient motors. Cluster sampling is useful when there is no sampling frame at the lowest level of the hierarchy but there is one at the cluster level, as in the case above where a ready list of all motors would not be available, but a list of all new apartment buildings would be. In many applications to monitor efficient equipment, the units occur naturally in clusters, with a different number of elements per cluster. For example, a building or plant location might constitute a natural cluster, with varying numbers of pieces of equipment per location. A cluster sampling approach can offer cost advantages. For instance, if a significant component of the cost of data collection is travel time between buildings, but there is minimal cost to collect data on units within a building, then it is more cost-effective to collect data on all units within a sample of buildings than to take a simple random sample across all units in the study. It will, however, usually be necessary to meter more pieces of equipment (sample more clusters) to achieve the same level of precision as the simple random sampling, but the reduction in cost and other benefits may more than offset this apparent increase in effort.

Source: UNFCCC

[21] Differences between stratified and cluster sampling are explained in Figure 17 and Table 14.

Determination of the sample size for cluster sampling

Determination of the sample size for cluster sampling

For the calculation of the required sample size for a cluster sampling approach, the following steps shall be conducted:

1. Clustering of the total population;
2. Calculation of the average proportion and standard deviation or the cluster mean and standard deviation;
3. Calculation of the required clusters to be sampled.

Step 1: Clustering of the total population:

The following tasks shall be conducted for step 1:

- Define clusters (e.g. geographical regions such as villages but also different types of elements, e.g. different types of buildings, etc.);
- Allocate all elements in the total population to the clusters and check if every element is included in one (and only one) cluster;
- Estimate the characteristics for each cluster:
  - For a proportion parameter: the expected proportion for each cluster;
  - For a mean parameter: the expected mean value of the cluster rather than the mean value of the individual elements.

Step 2: Calculation of the average proportion and standard deviation or the average mean and standard deviation:

Depending on the type of parameter (proportion or mean), the following values have to be calculated based on the estimates of step 1:

- Proportion: Average proportion and standard deviation that results from the different values of the clusters;
- Mean: Average mean and standard deviation that results from the different values of the clusters.

The formulas that need to be applied can be found in the *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation* (equations as shown in Table 16).

Table 16: Equations used for cluster sampling

Type of parameter	Parameter	Equation numbers
Proportion	Average proportion	13 (SSC) and 53 (LSC)
	Standard deviation of clusters	13 (SSC) and 53 (LSC)
Mean	Average mean	31 (SSC) and 70 (LSC)
	Standard deviation of clusters	31 (SSC) and 71 (LSC)

Equation numbers refer to the Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation

Step 3: Calculation of the overall required sample size:

The sample size is then calculated similar to the random sampling approach but taking the parameters of the average proportion or mean and the standard deviation of the clusters into account.

$$m \geq \frac{z^2 \times M \times V}{(M-1) \times precision^2 + z^2 \times V} \tag{20}$$

Where:

$$V = \left(\frac{SD}{mean}\right)^2 \quad \text{or} \quad V = \left(\frac{SD}{\bar{p}}\right)^2 \tag{21}$$

m	=	Number of clusters to be sampled[22].
M	=	Total number of clusters
mean	=	Cluster mean
$\bar{p}$	=	Average proportion
SD	=	Standard deviation of the clusters
z	=	Constant referring to the level of confidence (e.g. 1.645 for 90 % confidence and 1.96 for 95 % confidence.
precision	=	Required precision (e.g. 10% = 0.1)

As a result of the equation above, the number of clusters that needs to be sampled is determined. In case of determination of a mean value, the sampling result (i.e. the mean value determined based on the monitored sampled) refers to the clusters and not to the individual elements within each cluster. Hence, this has to be taken into account while using the sampled parameter for emission reduction calculations.

Example for cluster sampling

Examples 3, 7, 13 and 17 of the *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation* demonstrate the calculation of the sampling size for a cluster sampling approach.

3.3.5 Multi-stage sampling

Multi-stage sampling

Multi-stage sampling consists basically of at least a combination of two of the approaches mentioned above, conducted in a sequence. Typically, cluster sampling and simple random sampling are combined. The approach is described as follows in the *Sampling standard*:

[22]The *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation* use “c” instead of “m”.

Multi-stage sampling according to the *sampling standard*

Multistage sampling is a more complex form of cluster sampling. Measuring all the elements in the selected clusters may be prohibitively expensive, or not even necessary. In multistage sampling, the cluster units are often referred to as primary sampling units and the elements within the clusters secondary sampling units. In contrast to cluster sampling where all of the secondary units are measured, in multi-stage sampling data are collected for only a sample of the secondary units. For example, in a study of efficient lighting, if the operation hours of motors within any one building are thought likely to be similar across all motors then - especially if the cost of measuring them is relatively high - there is not much to be gained by metering all of them. It might be better to draw a sample of buildings, and then only measure a sample of motors from within each selected building. On the other hand, if the measurements are inexpensive once a technician is on-site, then it may make sense to monitor all of the fixtures. Multi-stage sampling can be extended further to three or more stages. For example, one might group the population into building complexes, then buildings, and finally fixtures. Multi-stage sampling can be extended further to three or more stages. For example, one might group the population into building complexes, then buildings, and finally fixtures.

Source: UNFCCC

Determination of the sample size for multi-stage sampling

Please consult the examples provided by the *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation* in order to calculate the required sample size.

Example for multi-stage sampling

Examples 4, 8, 14 and 18 of the *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation* demonstrate the calculation of the sampling size for a multi-stage sampling approach.

4 Sampling in practice

This chapter provides useful information to develop an appropriate sampling procedure. It will be shown that it is not sufficient to comply with the rules defined by the methodologies and the available sampling standards. Many different factors have impact on a successful implementation of a sampling approach. At the end, the applied sampling approach should be analysed also in regard to costs and benefits, meaning that the project participants overall objective is the issuance of the maximum of the achieved emission reductions by the project.

4.1 Development of an appropriate sampling procedure

In order to set up an appropriate sampling procedure, it is very useful to check what have been the problems in the past and what could be a problem applying the new rules. Hence, in the following typical pitfalls which occur when designing sampling plans and conducting sampling are presented.

The Sub-Chapter "How to develop an appropriate sampling approach" deals with the question on how to decide on the applied sampling approach (e.g. simple random sampling versus cluster sampling).

4.1.1 Typical pitfalls in sampling

This chapter provides an overview of typical pitfalls that have been occurred in the past within real CDM projects and that may remain or arise in future projects. This manual reviews the following pitfalls as shown in Table 17.

Table 17: Typical pitfalls

#	Description
Pitfall 1	No specific or insufficient description of sampling in the project design documents
Pitfall 2	Disregarding provisions of methodologies and other rules
Pitfall 3	Wrong determination of the sampling size
Pitfall 4	Lack of sufficient background information of the characteristics of the population
Pitfall 5	No evaluation of the selected sampling approach
Pitfall 6	Improper monitoring frequency
Pitfall 7	Single sampling approach for various parameters
Pitfall 8	Insufficient implementation of the sampling approach including insufficient training of staff
Pitfall 9	Insufficient software applications
Pitfall 10	Insufficient check of the sampling results

Pitfall 1: No specific description of sampling in the project design documents

Pitfall 1

There are several CDM projects and PoAs that provide only very limited information regarding sampling as follows:

“Sampling will be done in compliance with the UNFCCC requirements (e.g. the *sampling standard*)”

Although some of these projects achieved registration, one can expect significant problems during the verification and issuance process. As the *Sampling standard* explicitly asks for a sampling plan, this pitfall will lose importance. Nevertheless, there is a general tendency among project participants to postpone fixing of the sampling procedures for different reasons. First of all, simply time constraints may be the reason for such behaviour. Achieving registration of the project as soon as possible is deemed to be more important than finalizing the detailed sampling procedure. But often the fixing of the sampling approach is delayed as the development of the sampling procedure is a very difficult task or because the project participants try to maintain flexibility of the sampling procedure as long as possible and as long as the final project design is not fully finalized (which is often the case prior to registration). This behaviour then may results in significant problems at a later point in time. Typically, during verification the difficulties arise and project participants become aware that some of the problems cannot be solved ex post. For example, data that would be required to be collected during implementation of the project has not been collected as the project developer was not aware that such data would be required. Hence, the risk of being too constricted by a detailed sampling plan in the design documents can be seen as less problematic than missing required decisions in regard to the sampling plan. If it happens to turn out that the specification of the sampling approach cannot not be implemented as foreseen in the design documents, required adjustments should be possible applying a request for deviation or revision of the monitoring plan.

Project participants are recommended to provide a complete sampling plan as requested by the *Sampling standard*.

Pitfall 2: Disregarding provisions of methodologies and other rules

Pitfall 2

As described in Chapter 3.2 the UNFCCC rules are quite complex, following a strict hierarchy. If there are several documents dealing with the same topic one needs to analyse which document supersedes the other documents. Such a situation can occur if the methodology specifies provisions regarding sampling while at the same time the *Sampling standard* provides other rules. As a result project participants need to check which rules needs to be finally followed for the specific project activity. As a general rule, the provisions of the applied methodologies overrule the provisions of the *Sampling standard*. But, if the methodology does not provide information regarding sampling requirements, the rules of the *Sampling standard* apply. In a situation, where several documents provide rules for sampling, project participants are not allowed to select the rules depending on their own preferences but they need to analyse, which document overrules the other documents. To be even more confusing, in a situation of transition of rules and methodologies, it may become even more complex to detect the applicable ruling.

Project participants shall follow all relevant requirements taking into account the hierarchy of the different types of documents providing rulings.

Pitfall 3: Wrong determination of the sampling size

Pitfall 3

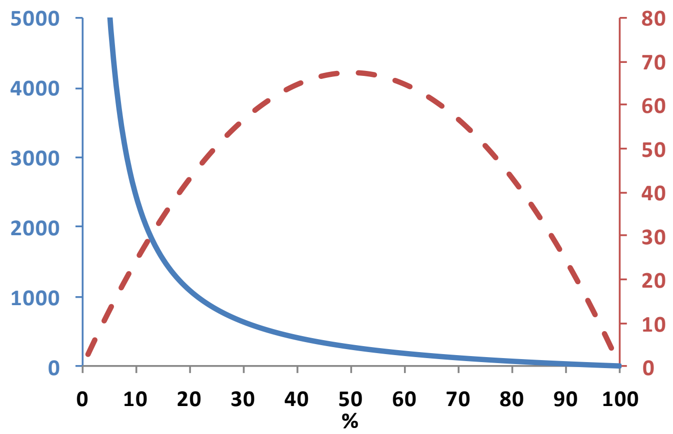
The determination of the sample size is essential in order to comply with the sampling requirements as defined by the *Sampling standard*. The guidance provided by the *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation* shall be applied in order to calculate the correct sample size. While applying the formulas of this document, it is important to use and follow the related provisions. A very prominent mistake is related to the application of the precision requirements. Under UNFCCC rules, the precision is applied as relative error margin. If this is disregarded in the calculation of the sampling size, e.g. by using sampling tools applying an absolute error margin, the sample size might be wrongly determined (see Figure 18). Based on the red dashed curve project participants might think that assuming a 50% proportion would be conservative as it results in the highest possible sample size. However, this is only true assuming an absolute error margin (e.g. 10 percentage points). In fact, UNFCCC rules would result in the blue curve.

Another typical mistake is the assumption that the calculated sample size is automatically sufficient to comply with the requirements. However, the calculated sample size is the minimum requirement. As an example, the following effects are not a priori reflected:

- Non-response of households;
- Failure of metering equipment;
- Failure of data collection (e.g. interviewer does not properly note down).

In order to compensate for such effects, the calculated sample size should be increased accordingly.

Figure 18: Sample size for a proportion depending on the expected value



Blue curve: Sample size calculated based on a relative error margin of 10%.
Red dashed curve: Sample size calculated based on an absolute error margin of 10 percentage points.

Finally, the calculation of the sample size needs to be done with the correct equations. These need to be selected according to the type of parameter (e.g. proportion versus mean) and according to the applied sampling approach.

Project participants shall also check if the requirements of the calculation approaches (i.e. sufficient size of the total population and the sample) are fulfilled. A minimum sample size of at least 30 is recommended in order to achieve general applicability of the approaches presented in this manual.[23]

Pitfall 4: Lack of sufficient background information of the characteristics of the population

Pitfall 4

The calculation of the sample size requires knowledge of the characteristics of the sample size. Prior to the determination of the proportion or a mean value through sampling, one needs solid assumptions or estimates for the expected proportion or the expected mean. In addition, in most situations the standard deviation among the different elements needs to be known. In the absence of robust assumptions or estimates, the calculation of the appropriate sampling size is not possible project participants should carefully evaluate their knowledge in advance (e.g. by conducting a pre-assessment and/or a survey). All available internal and public available sources of information should be taken into consideration. In case of insufficient knowledge even a pre-sampling study might be helpful as especially at the start of a project no data from previous monitoring periods is available. Such a pre-sampling study is often less expensive than repeating the complete sampling for a specific monitoring period.

[23] See Schwarze (2001) and Papula (1997)

Pitfall 5: No evaluation of the sampling approach

Pitfall 5

Although it might be sufficient to simply select a sampling approach (i.e. simple random sampling or one of the other approaches) for example for one of the following reasons:

- A registered project has already applied the same approach;
- Simple random sampling would not require additional information.

Project participants should evaluate all of the possible sampling approaches by conducting a cost benefit analysis. To decide simply on the resulting number of samples is usually not sufficient because other factors, such as travel times (especially in remote areas) or the overall costs and efforts also affect the evaluation of a specific sampling approach.

Pitfall 6: Improper monitoring frequency

Pitfall 6

Formulas and procedures as described in methodologies, the *Sampling standard* or other official UNFCCC documents ensure realistic and conservative estimations of emission reductions. Nevertheless, they do not guarantee materialisation of the maximum possible emission reduction achieved under a project. Hence, it might be worth to monitor a parameter that needs to be monitored only once a year according to the methodology more frequently within the project. For example, the parameter "appliances still in operation" (e.g. share of CFLs that are still in operation) has typically a great impact on the emission reduction calculation. Project participants should simulate the specific situation of their projects (e.g. set-up fictitious data for a small amount of appliances (e.g. 100 cook stoves) and check the impact of the monitoring frequency of specific parameters applied for monitoring and emission reduction calculation).

Especially in the following situations more frequent monitoring might be advantageous:

- At the start of the first crediting period because accumulated knowledge and understanding of the project is low;
- In all situations where a parameter has a significant impact on the emission reduction calculation and knowledge of the development of the parameter over time is unknown;
- In all situations where a parameter has a significant impact on the emission reduction calculation and the parameter is fluctuating or changing significantly in short time periods.

Pitfall 7: Single sampling approach for various parameters

Pitfall 7

Typically, several parameters need to be sampled. In such a situation it is often not sufficient to develop only one sampling approach. For each parameter the most appropriate sampling approach should be defined. Afterwards, one should check, if combining the approach for several parameters is possible and useful. Therefore, not only the calculated minimum sample size has to be taken into consideration, but also the requirements to increase the sample size due to expected non-response or other effects.

Furthermore the monitoring frequency is very important to check if a combined approach is possible.

Pitfall 8: Insufficient implementation of the sampling approach including insufficient training of staff

Pitfall 8

The efforts to sample are not finalised achieving a validated sampling plan. The implementation of the sampling plan has to be done accordingly. This may include the following activities:

- Development of specific meters;
- Recruitment and training of staff;
- Data collection during distribution of appliances (e.g. cook stoves or CFLs);
- Development of appropriate software applications.

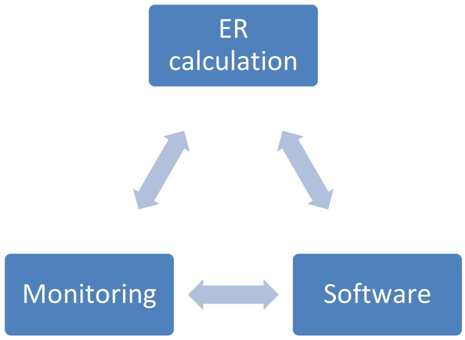
Training of staff involved in the sampling and in the distribution is of highest importance. If wrong data is collected during the distribution of a huge number of elements, any required corrections would result in tremendous additional effort. Staff, not trained on the applied software applications may also cause significant extra work to correct wrongly entered or processed data. Quality assurance procedures should be developed and put in operation to minimize any problems resulting from wrong data handling of staff.

Pitfall 9: Lack of appropriate software

Pitfall 9

Project participants' main focus before validation and registration of a CDM project or PoA is a successful and hopefully smooth validation and registration process under the CDM. Hence, a selection of appropriate software applications for sampling and early set-up of the overall software portfolio is typically postponed and not seen to be of high priority prior to registration. Maybe, the project implementation has even started before the project developer starts thinking about an appropriate software solution. This may be very risky, since it might be extremely laborious to adjust a software application afterwards to the specified formulas and procedures as described in the design documents (PDD, PoA-DD, CPA-DD). Hence, matching of the elements as shown in Figure 19 should be in the focus of project participants in an early stage of the process, but at latest once the design of the project implementation is clear. While DOEs and UNFCCC will check only CDM requirements, it is in the interest of the project developer, to ensure that adequate software is available right from the start of the project.

Figure 19: Software as part of ER calculation and monitoring



CDM projects and PoA may run for more than 20 years. In such a time span a software solution is likely to experience an update, new versions are developed while today existing versions may be discontinued. The applied software applications should be checked in regard to their continued availability. If no information can be found on the guaranteed and continuous support for a specific software application, risk of disappearance is high. Especially internet based software applications may be removed without any notification (e.g. a specific random number generator).

If the development of software is not taken into consideration from the start of the project and carefully regarded throughout the whole crediting period, projects may be exposed to significant risks as follows:

- Significant delay during verification / issuance;
- Additional and very high effort in order to transfer the data set to the new software versions;
- Partial or total loss of data resulting in partial or total loss of verifiable emission reduction.

See also Annex I or further information.

Pitfall 10: Insufficient check of the sampling results

Pitfall 10

After achieving the sampling results, project participants are recommended to check the appropriateness of the sampling results in advance of the verification process. Even a perfectly developed sampling approach, applying the formulas correctly to calculate the sampling size may fail as the estimates to determine the sample size might be wrong. A first check could be a comparison of the estimates with the sampled results. In case of a significant difference (e.g. the sampled mean of a parameter is outside of the estimated mean and the accepted error margin or estimated standard deviation) project participants should check, whether corrections would be required. In addition, the reliability calculations as provided by the *Best Practices Examples: Focusing on Sample Size and Reliability Calculation* should be applied. In turn, the sampling plan should already provide procedures to correct the sampling results accordingly. Such procedures could include:

- Repetition of the sampling;

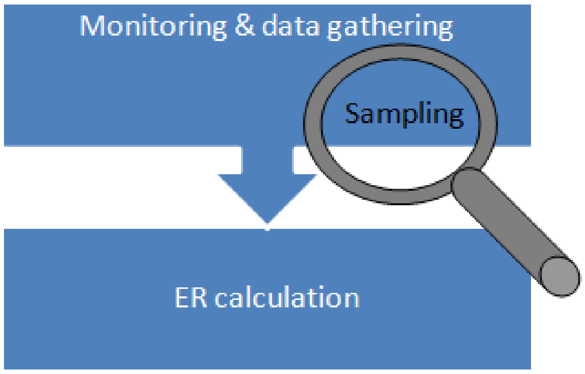
- Collection of additional data from a buffer group;
- Increasing the sample size by drawing additional elements to be sampled.

4.1.2 How to develop an appropriate sampling approach

Sampling is (only) part of the overall monitoring procedures

Sampling is carried out to achieve reliable data in case monitoring/metering of all elements would result in unreasonable effort. First of all, the development of the sampling approach has to fulfil the requirements set by UNFCCC. However, the selection of the sampling approach should be seen in the overall context of the project registration, the implementation and the cost/benefits of conducting the sampling procedures in practice. Sampling is only one part of the monitoring and of the data gathering procedures. Furthermore, the overall aim of the project activity is to achieve issuance of the achieved emission reductions (see Figure 20). In order to take this into consideration, the development of an appropriate sampling approach should be carried out in a stepwise procedure, not limited to whether UNFCCC requirements are fulfilled.

Figure 20: Sampling as part of monitoring and data gathering



One specific sampling approach, following exactly the UNFCCC requirements, might nevertheless be not advantageous compared to an alternative approach which is also in line with the UNFCCC requirements, but also taking into consideration aspects that enable a proper application in practice.

In order to develop an appropriate sampling approach in line with the sampling requirements defined by UNFCCC documents and as well in line with the objective to implement a robust but efficient and less cost intensive, the step-wise procedure as presented in Table 18 can be applied.

Table 18: Procedure to develop an appropriate sampling approach

Step	Task
1	Check/decide on existing regulatory framework
2	Identify parameters that need to be sampled
3	Identify sampling requirements for the identified parameters according to the applied methodology and other UNFCCC rules (e.g. Sampling standard)
4*	Identify which of the possible sampling approaches would be applicable for sampling of the parameters
5*	Collect all available data and information in order to derive good estimates which are required to calculate the sample size
6*	Conduct a sampling size determination for the parameters for each sampling approach that have been deemed to be applicable
7	Check the impact of the sampled parameters in regard to the emission reduction calculation.
8*	Check software requirements required for the different sampling approaches
9*	Evaluate the different sampling approaches and select the most advantageous sampling approach (e.g. cost-benefit-analysis)
10*	Conduct a sampling cross-check
11*	Develop the sampling plan
12	Start preparation of the implementation of the sampling plan

** Further explanation of this step is provided in the following.*

In the following further guidance is provided for selected steps that require further explanation according to Table 18.

Step 4: Identification of applicable sampling approaches for parameters

Step 4

For each parameter that is going to be determined by sampling, project participants should check which sampling approaches would be suitable. It is very difficult to provide general rules that guide project participants to the best suitable sampling approach. The specific situation of the project has a huge impact on the possible sampling approaches. Please consult the Appendix of the *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation* for further information.

Example: Different sampling approaches for a cook stove project

Assuming an efficient cook stove project that will distribute wood stoves and charcoal stoves various option of sampling are possible. For example, the determination of the efficiency of the new stoves may be determined via sampling as follows:

Option 1: Simple random sampling applied for the total population
Option 2: Stratified random sampling for the total population
Option 3: Simple random sampling applied separately for the group of wood stoves and for the group of charcoal stoves

Cross-CPA sampling

In case of PoAs, another question needs to be answered by the project participants. The question, if sampling shall be conducted separately for each CPA or rather across CPAs. Cross-CPA sampling may combine a certain number of CPAs or even all CPAs of the PoA. Generally, the Sampling standard allows cross-CPA sampling for PoAs applying small-scale methodologies. The different options regarding sampling of PoAs are shown in Figure 21.

Cross-CPA sampling according to the *Sampling standard*

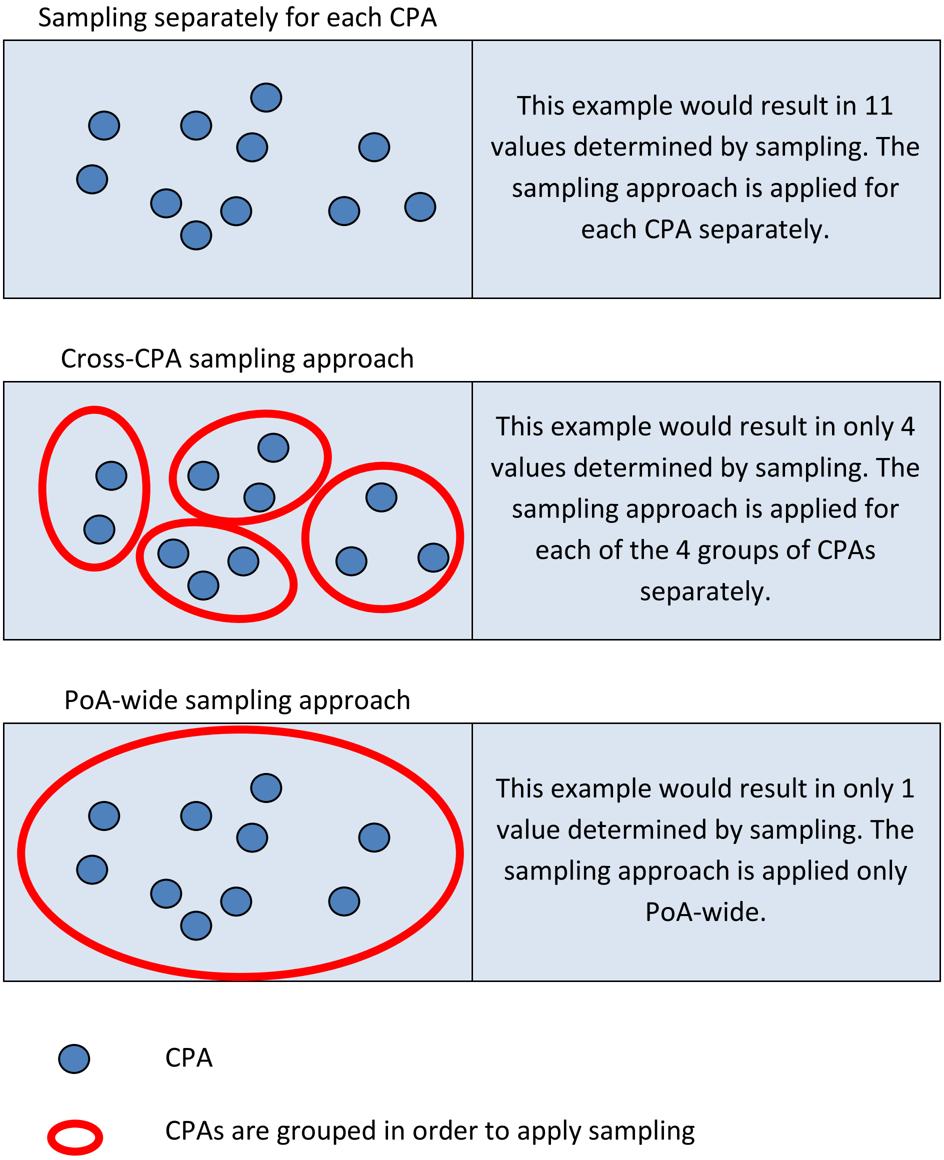
Parameter values shall be estimated by sampling in accordance with the requirements in the applied methodology separately and independently for each of the CPAs included in a PoA except when a single sampling plan covering a group of CPAs* is undertaken applying 95/10 confidence/precision** for the sample size calculation.

* That is, the populations of all CPAs are combined together, the sample size is determined and a single survey is undertaken to collect data e.g. if the parameter of interest is the daily usage hours of light bulbs, it may be feasible to undertake a single sampling and survey effort spread across geographic regions of several CPAs when either homogeneity of included CPAs relative to the light usage hours can be demonstrated or the differences among the included CPAs is taken into account in the sample size calculation. Currently PoAs applying large scale CDM methodologies are not included for applying single sampling plan covering a group of CPAs pending further analysis.

** This is consistent with the approach in many approved methodologies to aim at higher confidence/precision when the sampling/survey effort is undertaken less frequently e.g. AMS-I.E, AMS-II.G or AMS-I.J.

Source: UNFCCC

Figure 21: Cross-CPA sampling



Step 5: Collect all available data and information in order to derive good estimates which are required to calculate the sample size

Step 5

The calculation of the sampling size requires estimates for the parameter of interest (i.e. proportion, mean and standard deviation). Project participants are recommended to put some effort into this step as wrong estimates may result in an inadequate sampling size. Especially the determination of the sampling size for the first monitoring period requires special attention as reliable data based on previous monitoring periods is not available. Although additional data collection (e.g. a pre-study) would result in extra costs and effort, it is a good investment if no other data sources are available. The *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation* suggest the following options for data collection:

Estimates of the parameter of interest according to the *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation*

Estimates of the parameter of interest (proportion, mean and standard deviation) are required for sample size calculations. There are different ways to obtain these:

a) We may refer to the result of previous studies and use these results;
b) In a situation where we do not have any information from previous studies, we could take a preliminary sample as a pilot and use that sample to provide our estimates;
c) We could use “best guesses” based on the researcher’s own experiences.

Note that if the standard deviation is unknown but the range (maximum – minimum) is known then a rough “rule of thumb” is that the standard deviation can be estimated as the range divided by 4.

Source: UNFCCC

Step 6: Conduct a sampling size determination for the parameters for each sampling approach that have been deemed to be applicable

Step 6

The formulas for the calculation of the sample size are provided in Chapter 3.3. It should be noted that the calculated minimum sample size should be further assessed in order to treat non-response effects, outliers and other factors resulting in loss of feedback of sampled elements. There are several options to handle such conditions. Partially, these are defined in relevant UNFCCC documents such as methodologies, standards or guidelines. Hence, project participants need to check if there are already provisions in the applied UNFCCC documents. In order to achieve the minimum required sample size, the calculated minimum sample size should be increased to compensate for non-response and other effects. This could be done as follows:

Option 1: Increase the sample size and apply directly the increased number of samples

Option 2: Create a buffer group that will be included if the required sample size is not achieved

Option 3: Draw an additional sample

Option 4: Use the lower/upper bound of the confidence interval in case the required sample size could not be achieved

Illustrative example of increasing the sample size

The calculated required sample size shall be 100. The project developer expects an effect of non-response for 10 to 20 elements.

For Option 1 the project developer would draw a sample of 120 elements and would directly sample all 120 elements.

For Option 2 the project developer would draw a sample of 120 elements and would than sample initially only 110 elements. If this would not result into the required sample size the additional 10 elements would be included into the sample.

For Option 3 the project developer would draw a sample of 110 elements and would sample these. If the required sample size is not achieved, an additional sample of 10 elements will be drawn and included in the sample.

Option 4 would be possible if the options above would not result in a sufficient sampling size. Project participants should check the applied methodologies regarding provisions of this option.

Please note: It is recommended to sample at least the required number of samples increased by the minimum expected non-response rate.

Step 8: Check software requirements required for the different sampling approaches

Step 8

Sampling of several thousands of elements requires application of appropriate software. Project participants should dedicate sufficient effort to the selection of software applications. It is very unlikely that project participants will find a “ready-to-use” and “all-in-one” software solution for the purpose of sampling, conducting the related surveys, analysing the results and calculation of the resulting emission reductions. Hence, in many cases, project participants need to compose different software applications or have to develop a tailor-made individual software package which is often very costly and time consuming. Selection of improper software and resulting lack of required data may cause significant loss of verified emission reductions. Typically, the following types of software applications might be used during sampling:

- Type of software application;
- Spreadsheet programs;
- Web-based number generators;
- Web-based sample size calculators;
- Database programs;
- Web-based databases;
- Statistical programs;
- Supporting programs.

More information on the different software applications can be found in Annex I.

The selection of software applications used in the context of the sampling procedure should be undertaken carefully taking into account at least the following considerations:

- Availability of experienced staff: During development/set-up of the software application/portfolio experienced staff is required. Later on, staff is required for data input (e.g. staff that visits households as part of the monitoring procedures), data processing and analysis (e.g. calculation of the sampling results and the emission reduction calculation).
- Functionality of the software applications/portfolio: Does the software application/portfolio offer all required functions (e.g. database of households, sampling procedures, calculation of emission reductions)?
- Linkage within the software applications/portfolio: Normally data has to be exchanged between the different applied software applications. Hence, it is required to check if data import and export is compatible.
- Version development: A CDM project or PoA may run for more than 20 years. During such a time period software is typically enhanced. New versions may be developed while existing versions are discontinued. This problem does not only exist for the specific software in use, but also for computer operating systems or utility programs (e.g. data input via the internet is affected by browser developments).
- Multi-user functionality: Multi-user functionality requires in most situations tailor-made software solutions. Especially in case of CDM projects or PoAs with thousands of elements such functionality is essential especially in cases where the data has to be surveyed for the first time (i.e. no existing electronic data set that can be used).
- Costs: This includes costs for the development/set-up of the software application/portfolio, licences, maintenance, training of staff, etc. Furthermore, also related hardware requirements should be considered (e.g. web based solutions would require an appropriate internet service provider).

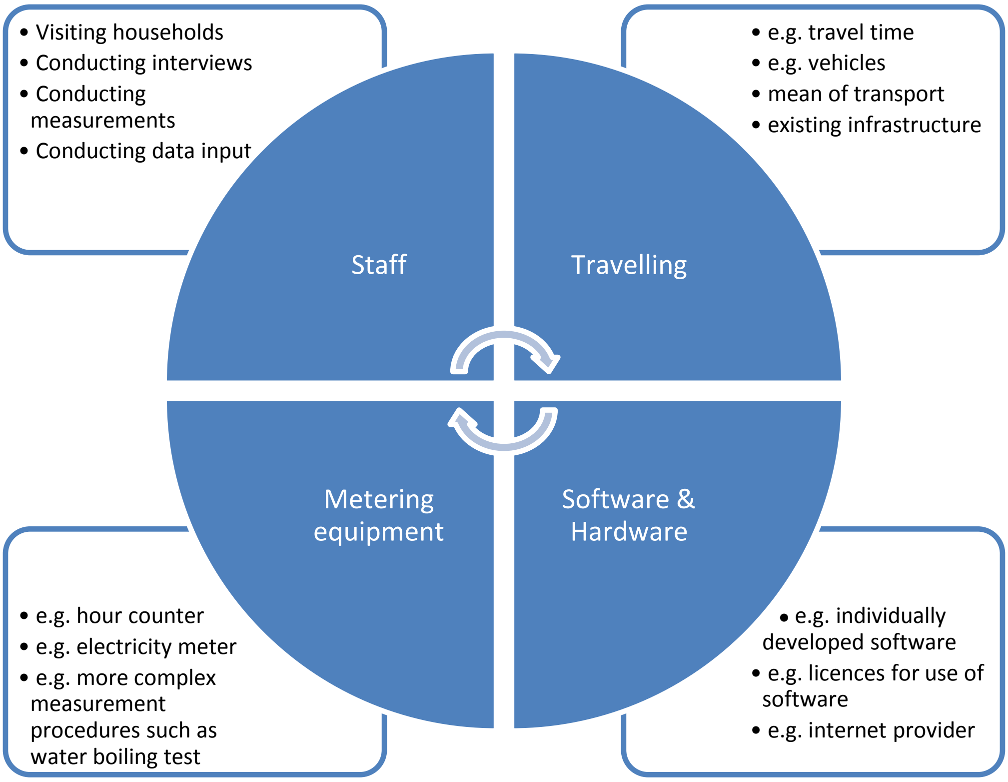
Step 9: Evaluate the different sampling approaches (e.g. cost-benefit-analysis)

Step 9

The evaluation of the different sampling approaches shown above, simply related to the resulting sample size would not be sufficient. The overall effort of the sampling approach can usually be determined as costs. However, these costs are impacted by various factors as shown in Figure 22.

Depending on the specific project conditions and the specific parameters of interest, the outcome of an evaluation of the different sampling approaches will vary. A general rule is therefore not available. On the other hand, the parameters that are sampled will be used to calculate the achieved emission reductions. Hence, it needs to be checked, if some additional effort for sampling may increase the resulting emission reductions that will be issued afterwards and hence may be worth the effort.

Figure 22: Main cost categories related to sampling



Step 10: Conduct a sampling cross-check

Step 10

Before elaborating the sampling approach further, project participants should double-check if the selected approach is suitable. The following checklist may be applied (see Table 19).

Table 19: The sampling cross-check

No.	Topic	☑
1	Provisions of the applied methodology are correctly followed	☐
2	Sampling requirements according to the Sampling standard and other UNFCCC documents are correctly followed	☐
3	The sampling approach has been developed based on an overall evaluation of the different possible options	☐
4	Available information is sufficient to develop the required estimates for the sampled parameters	☐
5	If No. 4 is not satisfied, an additional data collection approach (e.g. pre-sampling study) is envisaged	☐
6	The sampling plan has been developed according to the requirements of the *Sampling standard*	☐
7	The sampling approach/plan is robust in order to be smoothly validated and to be prepared for effects during verification (e.g. handling of non-response)	☐
8	Implementation of the sampling approach is envisaged (e.g. training of staff, development of software applications, etc.)	☐

Step 11: Develop the sampling plan

Step 11

The requirements of the sampling plan are defined by the sampling standard. It should be noted that the sampling plan should not only be seen as relevant for the validation and registration of the project activity. DOEs will check the sampling plan initially during validation for compliance with UNFCCC rulings. But, later on, the DOE will also check, if the sampling plan has been correctly implemented and executed as described in the design documents. This aspect should be taken into consideration when developing the sampling plan.

4.2 The role of the DOEs

Validation and verification by DOEs

According to the *Modalities and Procedures for a Clean Development Mechanism* as defined in Article 12 of the *Kyoto Protocol* (CDM M&P) and the *Clean Development Mechanism Validation and Verification Manual* (CDM VVM[24]), a DOE shall, inter alia, validate proposed CDM project activities and verify/certify reductions in anthropogenic

[24] The VVM will be replaced by the Clean development mechanism validation and verification standard (VVS).

emissions by sources of greenhouse gases. Hence, the sampling approach as defined by the sampling plan will be subject to evaluation by the DOE at two different stages of the CDM cycle. First of all, general compliance with UNFCCC ruling is checked during the validation phase and secondly, the correct implementation and execution of the sampling plan within each verification will be assessed. Please note that changing of rules needs to be taken into account. In addition, further documents such as the following

- Methods, if any, to deal with missed reliability targets without compromising conservative estimates for emission reduction
- Best practice examples for DOE validation/verification for sampling and surveys

are currently under development according to the *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation.* The content of these documents may be included in a following edition of this manual.

5 Sampling blueprints

In the following, three project examples are provided that demonstrate the development and documentation of an appropriate sampling approach in practice. The project types for the examples have been taken with prior considering of the importance of these project types for PoAs, the need for sampling and for enabling a variety of different sampling procedures. These finally selected examples are

a) A cook stoves PoA following AMS-II.G. and applying a simple random sampling approach;
b) A solar water heater PoA following AMS-I.J and applying a multi-stage sampling approach;
c) A CFL SSC CDM project activity following AMS-II.J and applying a stratified random sampling approach.

While the first example is presented with comprehensive documentation of the sampling approach, the other examples only focus on the differences due to the different methodologies and/or different sampling approaches.

The examples apply the *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation*. As this document is still a draft, project participants are requested to consult the most recent and approved versions of documents available at the UNFCCC website. As content and status of approved methodologies and/or UNFCCC rules evolve, these examples shall only be used as indication and project participants need to develop a suitable and robust sampling approach based on the specific circumstances of their projects following the rules that are approved and active at this point in time.

The following structure is used to show the sampling blueprints:

Part of the project design documents that need to be prepared for the validation/registration.

[Text provided in square brackets indicates specific text elements that need to be developed by project.]

> Further guidance and explanation that is only meant for the reader of this manual.

5.1 Sampling for a cook stoves project

Project description[25]

Project description:	
Type:	PoA
Project type:	Cook stoves
Methodology:	AMS-II.G version 3.0
Country:	Multi-country
Sampling approach:	Simple random sampling
Cross CPA-sampling:	Yes (partially)

The purpose of the PoA is the dissemination of improved cooking stoves (ICS) in a number of different countries. The Programme will promote stoves fuelled with wood or charcoal that replace existing less efficient cooking stoves using woody-biomass.

The improved cooking stoves (ICS) to be distributed are more efficient in transferring heat from the fuel to the pot when compared to the stoves typically being used in the host countries. By replacing inefficient stoves, the PoA will save on consumption of woody biomass (either wood or charcoal made of wood) which is the dominant fuel used for cooking in the selected countries. The ICSs applied in the PoA have been designed to match the traditional utensils and cooking habits of the people in countries.

It is assumed that in the absence of the project activity, the baseline scenario would be the use of fossil fuels for meeting similar thermal energy needs. Therefore, by reducing the total amount of fuel required for cooking, the replacement of less efficient stoves with more efficient ICS reduces the amount of Green House Gases (GHG) emitted into the atmosphere. Certified Emission Reductions (CERs) are calculated following version 3.0 of methodology AMS.II.G on the basis of the mass of non-renewable woody biomass saved by the ICSs.

Sampling plan

Due to the large number of ICS envisaged to be distributed as part of the CPAs to be included in the PoA, it is not economically feasible to monitor each individual ICS unit distributed. Therefore, representative sampling will be undertaken as part of a PoA-wide Sampling Plan that is designed in line with the requirements of AMS II.G v3 and the "Standard for sampling and surveys for CDM project activities and programme of activities" (the *Sampling standard*). The Sampling Standard (paragraph 19, footnote 13) allows for sampling across a group of CPAs, provided the homogeneity of population can be demonstrated, or differences are taken into account in the sample size calculation and 95/10 confidence/precision is applied. The methodology requires 90/10 confidence/precision if annual sampling is applied, or 95/5 confidence/prevision if biennial (every two years) sampling is applied.

Flexibility to apply cross-CPA sampling is likely to be critical for the feasibility of the proposed PoA due to the large number of CPAs envisaged. In particular, this is the case

[25] Project description partially based on a cook stove PoA that is currently in validation.

for the parameter η_{new} which involves carrying out water boiling tests in the field. For this parameter, there is likely to be a very high level of homogeneity since the ICS to be distributed have been designed to meet stringent efficiency specifications and are manufactured in modern factories. There is no reason to think the actual efficiency of similar ICS models will vary significantly from CPA to CPA or even country to country.

> Cross-CPA sampling is typically deemed to be advantageous in regard to the overall number of elements that need to be sampled as long as sufficient homogeneity can be assumed.

Sampling design: i. Objectives and reliability requirements

The objective is to obtain a reliable estimate of the following key variables over the course of the crediting period and meeting the indicated confidence/precision levels. All of the parameters presented below will be determined separately for the wood cook stoves and the charcoal stoves.

The most stringent confidence/precision levels will be applied as required by the circumstances (and indicated below):

Parameter	**Description of parameter**	**Confidence/precision level (frequency of sampling)**[b]	**CPA grouping**
η_{new}	The thermal efficiency of the ICS distributed (%)	95/10 will be applied as cross-CPA sampling and annual sampling will be applied.	Sampling will be applied across all CPAs up to the PoA level.
SOF	The Stove Operating Fraction, i.e. the fraction (up to 1.0) of users using the ICS	90/10 will be applied as sampling will be applied annual for each CPA.	Sampling for each CPA
u_{old}	The amount of fuel that continues to be used in the replaced stoves (kg)	90/10 will be applied as sampling will be applied annual for each CPA.	Sampling for each CPA
f_{old}	The fraction of stove users still using baseline (replaced) stoves (up to 1.0)	90/10 will be applied as sampling will be applied annual for each CPA.	Sampling for each CPA

> It is important to identify correctly the relevant sampling requirements (i.e. level of confidence and precision) in case of different options available according to UNFCCC documents.

Sampling design: ii. Target population

The overall target population are the ICS distributed as a result of the CPAs implemented under the PoA. The ICS to be sampled will be drawn from the list of individual ICS serial ID numbers contained in the PoA Distribution and Monitoring Database, which is maintained by the CME. Each ICS is assigned to a CPA in the PoA Distribution and Monitoring Database and linked to an end user whose premises will be visited during monitoring. See Section below on Sampling Frame for a detailed discussion of the approach for differentiating the target population during sampling, so as to ensure homogeneity.

> Typically, under the CDM the target population and the sampling frame are identical (i.e. every element in the total population is also part of the sampling frame and can hence be selected during the sampling approach).

Sampling design: iii. Sampling method

The CME will draw a single sample for each defined sampling frame. It is likely that the required sample size of the parameters will be different, since the variance in values can be expected to differ (see below for discussion of sample size). The required number of ICS to be selected for sampling of each parameter will be determined by the CME according to the level of reliability required for that parameter.

Simple random sampling will be used whenever the homogeneity of the population within the sampling frame can be expected to be sufficiently high. The definition of the sampling frame as described below can help ensure an appropriate level of homogeneity. It is also possible that cross-country (PoA-wide) sampling may be considered by the CME in the case of monitoring η_{new} (for estimating the efficiency of the same stove models), since the variance in values obtained during sampling is not likely to differ between countries. If necessary, or deemed to be appropriate by the CME, other sampling methods could also be applied.

To ensure a random selection, random number generators shall be applied. Each ICS in the target population is uniquely identifiable by its Serial ID number. Each ICS can thus be allocated a Sample Selection Number in each monitoring period, starting at 1 and increasing up to the total number of ICS in the Database for that pre-defined sampling frame (see below – this could be defined according to the country-level DO-level , or user-group etc). Applying the random number generators, the ICS can then be randomly chosen from the defined population up to the required sample size as calculated by the CME. The CME will also account for the differences in ICS vintages, by ensuring a representative share of stove vintages in the sample selected. It is not envisaged that geographical or demographic representativeness is required, given the highly homogenous characteristics of the end users within each defined sampling frame.

To determine the parameters, sampling will involve the following approaches (outcome in brackets):

- η_{new}: ICS will be tested using WBTs. (ICS efficiency)
- SOF: visual inspection of the premises to see if ICS is operational and in use. Interview with end user if required to verify that ICS is still in use. (Yes/No)
- f_{old}: visual inspection of the premises to see if baseline (replaced) stove continues to be used. Interview with end user if required to verify that baseline (replaced) stove is still in use. (Yes/No)
  The parameter will be estimated by measuring the fraction of end users not using baseline (replaced) stoves, such that:
  $f_{old} = 1 - f_{nonold}$
- u_{old}: interview with end user to establish the share of cooking that is done using baseline stove compared with the scenario prior to receipt of the ICS, multiplied by total annual fuel consumption (kg/year). This will be done by using a simple estimation technique such as:

$$u_{old} = \frac{MPM_{after\ ICS}}{MPM_{before\ ICS}} \cdot \text{Total annual fuel consumption (kg)}$$

Where:

$MPM_{after\ ICS}$ meals per month cooked using the baseline (replaced) stove after the receipt of the ICS

$MPM_{before\ ICS}$ meals per month cooked using the baseline (replaced) stove before the receipt of the ICS

Note: in the case of estimating u_{old} for charcoal stoves, the above value would then need to be converted into an amount of woody biomass using the IPCC conversion factor of 6.

More ICS will be selected for sampling than is required by the sample size, to ensure that if there are any ICS that are unable to be reached the required accuracy is still achieved. The size of the buffer will be driven by the required sample size – if the sample size is a relatively small number (e.g. 30), then a relatively large buffer may be necessary (say, 20-30%); if the sample size is a relatively large number (e.g. 100 or more), then a smaller buffer may be sufficient (e.g. 10%). The CME may choose to stop monitoring a particular parameter once the required level of confidence/precision has been reached, as long as the calculated minimum number of samples has been achieved. The following steps could logically be followed for the case of applying a 30% buffer:

1. Visit first 10% of premises required for the 30% buffer. If the number of responses is sufficient to achieve the required reliability level, then stop sampling.
2. If step 1 is not sufficient to achieve the required reliability level, then visit the next 10% of premises (increases the additional sampling to 20% of the 30% buffer). If this additional sampling is sufficient, then stop sampling.
3. If step 2 is not sufficient to achieve the required reliability level, then complete the final 10% of the additional sampling buffer (bringing the total to 30%).

An additional round of sampling may be completed if necessary in order to achieve the required confidence/precision level (as is discussed below under Quality Assurance/Quality Control.

> Simple random sampling can be assumed to be applicable in most cases. However, the approach does not automatically result in the lowest possible sample size and sampling effort.

Sampling design: iv. Sample size

The size of the sample for each sampling frame is determined by the requirement to achieve 90/10 or the 95/10 confidence/precision (as appropriate) for the estimation of the proportion or mean value of the parameter investigated. Whenever the CME applies cross-CPA sampling, it will select a sample size sufficient to achieve 95/10 confidence/precision.

An overview of the estimated sample sizes for a hypothetical population of 100,000 ICS units applying a level of 90/10 and 90/10 is provided below. Note: of the four parameters to be monitored, two are proportions/percentages (SOF and f_{old}) and two are mean values (η_{new} and u_{old}).

In order to calculate the required sample size estimates for the proportions and the mean values are required. Furthermore, the standard deviation needs to be assumed in case of sampling for a mean value. For the first monitoring period, the values as described below are applied. For the following monitoring periods, the estimates shall be adjusted taken the results of the previous monitoring period(s) into account.

For the parameters SOF and f_{old} the following equation[26] is applied:

$$n \geq \frac{z^2 \times N \times V}{(N-1) \times precision^2 + z^2 \times V} \qquad (22)$$

Where:

$$\mathrm{V} = \frac{p \times (1-p)}{p^2} \qquad (23)$$

n	=	Number of elements to be sampled
N	=	Total number of elements in the population
p	=	Proportion
z	=	Constant referring to the level of confidence (e.g. 1.645 for 90 % confidence and 1.96 for 95 % confidence.
precision	=	Required precision (e.g. 10% = 0.1)

For the parameters η_{new} and u_{old} the following equation[27] is applied:

---

[26] Equation according to the *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation (Agenda of EB 66)*

[27] Equation according to the *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation (Agenda of EB 66)*

$$n \geq \frac{z^2 \times N \times V}{(N-1) \times precision^2 + z^2 \times V} \tag{24}$$

Where:

$$V = \left(\frac{SD}{mean}\right)^2 \tag{25}$$

n	=	Number of elements to be sampled
N	=	Total number of elements in the population
mean	=	Average value of the parameter that is expected in the total population
SD	=	Standard deviation of the parameter that is expected in the total population
z	=	Constant referring to the level of confidence (e.g. 1.645 for 90 % confidence and 1.96 for 95 % confidence.
precision	=	Required precision (e.g. 10% = 0.1)

The calculation of the required sample size is illustrated below for a 90/10 and 95/10 level of confidence and precision.

η_{new}:

$$n \geq \frac{1.96^2 \times 100{,}000 \times V}{(100{,}000-1) \times 0.1^2 + 1.96^2 \times V} \geq 14 \tag{26}$$

Where:

$$\mathrm{V} = \left(\frac{0.02}{0.21}\right)^2 \tag{27}$$

A sample size of 30 would be sufficient to achieve the required confidence/precision for η_{new} values ranging from 0.2 to 0.326. The anticipated mean of η_{new} for ex-ante emissions reduction purposes is 0.263 and the standard deviation is 0.0315.

SOF:

$$n \geq \frac{1.96^2 \times 100{,}000 \times V}{(100{,}000-1) \times 0.1^2 + 1.96^2 \times V} \geq 21 \tag{28}$$

Where:

$$V = \frac{0.95 \times (1 - 0.95)}{0.95^2} \quad (29)$$

A sample size of 100 would be sufficient to achieve the required confidence/precision for SOF values ranging from 1.0 to 0.8. (The anticipated value of SOF for ex-ante emissions reduction purposes is in the order of 0.95; to be applied in the emissions reduction calculations at the CPA level).

u_{old}:

$$n \geq \frac{1.96^2 \times 100{,}000 \times V}{(100{,}000 - 1) \times 0.1^2 + 1.96^2 \times V} \geq 20 \quad (30)$$

Where:

$$V = \left(\frac{28}{124}\right)^2 \quad (31)$$

In the case of CPAs involving residential charcoal users for example, a sample size of 50 would be sufficient to achieve the required confidence/precision for u_{old} values ranging from 8kg to 240kg per year, which is equivalent to a range of 1-33% of the average annual charcoal consumption of households of 726kg per household. In this example, the anticipated mean of u_{old} for ex-ante emissions reduction purposes is 124 kg and the standard deviation is 28 kg. (An assumed value will be applied to each CPA for ex-ante emissions reduction calculation purposes, however, the actual values estimated during monitoring will help determine the values to be assumed for future CPAs).

f_{old}:

It is expected that logically the majority of end users will not use the old stoves after they have received the new and more efficient stoves (in order to make the decision to purchase the new stove, the end user has perceived an opportunity to reduce fuel costs by making investing that will only pay off if they stop cooking with their inefficient stove). Hence, the fraction of end users continuing to use baseline (replaced) stoves (f_{old}) is related to the fraction of end users *not* continuing to use baseline (replaced) stoves ($f_{non,old}$):

$$f_{old} = 1 - f_{nonold}$$

The parameter will be determined by sampling as follows.

$$n \geq \frac{1.96^2 \times 100{,}000 \times V}{(100{,}000 - 1) \times 0.1^2 + 1.96^2 \times V} \geq 43 \quad (14)$$

Where:

$$V = \frac{0.9 \times (1-0.9)}{0.9^2} \tag{15}$$

A sample size of 100 would be sufficient to achieve the required confidence/precision assuming a range of the proportion of end users not using the old stoves ($f_{non,old}$) from 1.0 to 0.8. (The anticipated value of the proportion of end users not using the old stoves for ex-ante emissions reduction purposes will be applied at the CPA level, and is expected to be in the order of 0.9. Experience gained during the actual monitoring of this parameter will help inform the value assumed for future CPAs).

> The calculation of the sample size follows the *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation*.

Sampling design: v. Sampling frame

Generally, the above mentioned parameters are sampled separately for the group of wood stoves and charcoal stoves.

Parameter	Sampling frames
η_{new}	There are only two sampling frames for which sampling is applied separately: • All wood stoves distributed under the PoA; • All charcoal stoves distributed under the PoA.
SOF	There are only two sampling frames for each CPA: • All wood stoves distributed under the specific CPA; • All charcoal stoves distributed under the specific CPA.
u_{old}	There are only two sampling frames for each CPA: • All wood stoves distributed under the specific CPA; • All charcoal stoves distributed under the specific CPA.
f_{old}	There are only two sampling frames for each CPA: • All wood stoves distributed under the specific CPA; • All charcoal stoves distributed under the specific CPA.

> The specific sampling frames result from the different types of stoves distributed and the different CPAs included in the PoA.

Data: i. Field measurements

The following parameters will be measured as indicated below:

Parameter	Timing (indicative)	Frequency (required by AMS II G)	Methods to be applied	Comments on seasonal fluctuation
η_{new}	First monitoring will occur	Annually	Water boiling test (WBT)	Not due to any seasonal fluctuation

	within 12 months of ICS distribution at the latest, but will include ICS of different vintages			
SOF	First monitoring is likely to occur within 12 months of ICS distribution, but will include ICS of different vintages	Annually	Visit to premises, visual inspection and interview with owner of ICS if required	Unlikely to be due to seasonal fluctuation
u_{old}	First monitoring is likely to occur within 12 months of ICS distribution, but will include ICS of different vintages	Not specified in methodology – likely to be monitored once at the start of crediting period to establish average value of consumption using baseline stoves for the target population and then again later in the crediting period to update/confirm this value.	Visit to premises and interview with owner of ICS to estimate share of annual consumption accounted for by baseline stoves (approach described under Sampling Method above)	Unlikely to be due to seasonal fluctuation if interview questions are explained properly
f_{old}		Not specified in methodology – likely to be monitored simultaneously with SOF (at least every year, but may be more frequently)	Visit to premises, visual inspection and interview with owner of ICS if required (if the end user is clearly not using a baseline stove	Unlikely to be due to seasonal fluctuation if interview questions are explained properly

			then no interview is required)	

This information has been developed the provisions as specified by the applied methodology.

Data: ii. Quality assurance / Quality control

The potential for non-responses, refusals and related issues will be considered by the CME during sample selection. If the sampling results are insufficient to achieve the target reliability levels, the CME has a number of options to address this (see below). By selecting a larger than necessary sample size before commencing monitoring, the CME can help ensure that an adequate number of responses are obtained during sampling.

If it is necessary to engage third parties for carrying out field measurements, the CME will ensure that any such third parties are credible, experienced adequately trained for the tasks they are contracted for (e.g. carrying out of WBTs in line with a methodology supported by an appropriate international body such as PCIA). Training will also be provided to the parties carrying out the actual field measurements (Monitoring Agents) on how to deal with non-responses etc if necessary.

The calculation of the sample size will be carried out using estimates for proportions, mean of values and standard deviations as the actual characteristics of the population/sampling frame are unknown. In order to ensure the quality of the sampling results, the CME can draw on the provisions for reliability calculations as provided by the *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation (Agenda of EB 66)*. In the event that the sampling results do not fulfil the required level of confidence and precision, the CME can undertake additional samples. If the reliability is still not sufficient after additional samples, the sampling may be repeated with an increased sample size. Alternatively, the CME may choose to apply the lower bound of the sampling results as is allowed for by the methodology (AMS II G v3, paragraph 22).

The data contained in each individual CPA Monitoring Record and collected during field measurements will transferred to the CME by the Monitoring Agents. Either the originals of the CPA Monitoring Records or scanned copies of each Record will also be provided to the CME to enable cross-checking.

The CME will be responsible for maintaining a secure PoA Distribution and Monitoring Database, which includes all the data relating to the CPAs within the PoA. The Database will be located on the CME's secure server. The system automatically backs up on regular basis any files that have been modified. The files are backed up onto separate hard drives that are regularly swapped to ensure there is always one drive located securely offsite. The CME may improve this system over time with new technology.

Project participants need to develop this section depending on the specific circumstances of the individual projects.

Data: iii. Analysis

The data obtained from sampling of each group of CPAs will be used to estimate values for the parameters described above. The values will then be factored into the emissions reduction calculations and result in the request for issuance of CERs for that group of CPAs. The parameters are applied for emission reduction calculations as follows:

$$B_{y,savings} = B_{old} * \left(1 - \frac{\eta_{old}}{\eta_{new}}\right) \tag{14}$$

Where:

B_{old}	=	Quantity of biomass used in the absence of the project activity in tonnes/ year
η_{old}	=	Efficiency of the system being replaced. According to the methodology, a default value of 0.10 can be used if the replaced system is a three stone fire, or a conventional system with no improved combustion air supply or flue gas ventilation system, i.e. without a grate or a chimney; for other types of systems a default value of 0.2 can be used. Weighted average values will be used if more than one type of system is being replaced.
η_{new}	=	Efficiency of the system being deployed as part of the project activity (fraction) as determined by using Water Boiling Test (WBT) protocol. Weighted average values will be used if more than one type of system is being introduced by the project activity.

$$B_{old} \geq LAF * N_{all} * SOF * \left(Q_{biomass} - \left(\frac{u_{old}}{1000}\right) - f_{old}\right) * Stove_{year} \tag{14}$$

Where:

LAF	=	Net to gross Adjustment factor (0.95) applied in accordance with paragraph 13 and 23 of AMS-II. G version 03
N_{all}		Total number of stoves distributed.
u_{old}	=	Amount of woody biomass consumption that is consumed through the continued use of old stoves (kg/year) to be established through sampling.
f_{old}	=	Fraction of end users that are still using their replaced stoves during the monitoring period (established through sampling). During monitoring, it is the fraction of end users not using baseline (replaced) stoves that will be estimated using sampling. This parameter is referred to as $f_{non,old}$ and the approach is explained in the PoA Sampling Plan in Section E.7.2 of the PoA DD. Thus, $f_{old} = (1- f_{non,old})$.

$Stove_{year}$	Calculated average stove operation years in the monitoring period (years). The average use of 365 days per year would result in a value of 1 year. The average use of 274 days per year would result in a value of 0.5 years.

If more than one sample is taken during a monitoring period, the approach will be to take the values obtained during the first sampling phase and the values obtained during the second sampling phase and calculate the average of these values. This average value will then be applied for the purposes of the emissions reduction calculations.

> This section provides the reference to the emission reduction calculation according to the applied methodology.

Implementation: i. Implementation plan

It is envisaged that the CME will implement the Sampling Plan over the course of the first 12 months of the PoA, including contracting all necessary third parties who would be responsible for actual field measurements. The actual timing will depend on the speed of CPA inclusion and ICS distribution, as well as the decisions made by the CME to either hire and train direct staff to conduct field measurements or to sub-contract these responsibilities. The CME will train any such third parties to ensure that field measurements are undertaken in line with the standards required of the Sampling Plan (e.g. WBTs will follow a procedure that meets an internationally-recognized standard).

The skills and experience required for the data collection activities under the Sampling Plan include:

- Experience conducting WBTs;
- Experience conducting door-to-door surveys of biomass consumption;
- Local language skills and English language skills;
- Cultural awareness;
- Numerical proficiency;
- Data entry skills.

> Project participants need to develop this section depending on the specific circumstances of the individual projects.

5.2 Sampling for a solar water heaters project

Project description

Project description:	
Type:	PoA
Project type:	Solar water heaters (residential use)
Methodology:	AMS-I.J version 1.0
Country:	Single-country
Sampling approach:	Systematic sampling and multi-stage sampling
Cross CPA-sampling:	No

The purpose of the PoA is the distribution of solar water heating (SWH) systems in a single country. The SWH systems will replace existing electricity or fossil fuel based water heating systems in residential applications.

Sampling plan

Due to the large number of SWH systems envisaged to be distributed as part of the CPAs to be included in the PoA, it is not economically feasible to monitor each individual SWH system distributed. A representative sampling will be undertaken based on the sampling requirements as defined by AMS I.J v1 and the "Standard for sampling and surveys for CDM project activities and programme of activities" (the *Sampling standard*).

> Further information may be provided according to the specific project circumstances. Please see example Cook stoves in Chapter 5.1 for further recommendations.

Sampling design: i. Objectives and reliability requirements

The objective is to obtain a reliable estimate of the following key variables over the course of the crediting period and meeting the indicated confidence/precision levels. The methodology does not specify directly equations and parameters. Hence, based on the "system metering method" the parameters identified below will be monitored. Furthermore, the methodology requires: "For residential SWH systems, in any given year, emission reductions can only be claimed for systems that are demonstrated to be operational and in compliance with manufacturer-required maintenance procedures, on an annual or biennial (every other year) basis during the crediting period." The most stringent confidence/precision levels will be applied as required by the circumstances (and indicated below).

Parameter	**Description of parameter**	**Confidence/precision level (frequency of sampling)**	**Sampling approach**
$Q_{SWH,y}$	Energy content of hot water supplied by the SWH systems to consumers (kWh)	90/10 will be applied as data collection will be done annual for each CPA.	Systematic sampling
$FC_{SWH,y}$	Fossil fuel consumption of the SWH systems (litres or kg)	90/10 will be applied as data collection will be done annual for each CPA.	Systematic sampling
$EC_{SWH,y}$	Electricity consumption of SWH systems	90/10 will be applied as data collection will be done annual for each CPA.	Systematic sampling

	(kWh)		
$f_{compliance,y}$	Fraction of SWH systems find to be operational and in compliance with manufacturer-required maintenance procedures (up to 1.0)	90/10 will be applied as annual sampling is conducted per CPA.	Multi-stage sampling

> Further information may be provided according to the specific project circumstances. Please see example Cook stoves in Chapter 5.1 for further recommendations.

Sampling design: ii. Target population

The overall target population consists of the SWH systems installed as a result of the CPAs implemented under the PoA. Each SWH system will be identified by a unique serial number for each CPA. Based on the serial numbers, the systematic sampling and the multi-stage sampling will be applied.

> Further information may be provided according to the specific project circumstances. Please see example Cook stoves in Chapter 5.1 for further recommendations.

Sampling design: iii. Sampling method

The parameters $Q_{SWH,y}$, $FC_{SWH,y}$ and $EC_{SWH,y}$ are determined by meters installed within the SWH system. As it would be economically not feasible to equip each SWH system with a complete set of meters, only the SWH systems identified to be sampled will be equipped accordingly. A systematic sampling approach based on the serial ID numbers of the SWH systems will be applied for each CPA in order to allow installation of meters during the installation process of the SWH systems.

The parameter $f_{compliance,y}$ will be sampled annually based on a multi-stage sampling approach which consist of a combined cluster and random sampling approach. In a first step a set of clusters (which represent villages where the SWH systems are installed) will be identified to be sampled. From each selected cluster, only a specific number of SWH systems will be identified to be investigated by applying random sampling.

> Further information may be provided according to the specific project circumstances. Please see example Cook stoves in Chapter 5.1 for further recommendations.

Sampling design: iv. Sample size

The size of the sample for each sampling frame is determined by the requirement to achieve 90/10 confidence/precision for the estimation of the proportion or mean value of the parameter investigated.

An overview of the estimated sample sizes for a hypothetical population of 75,000 SWH systems applying a level of 90/10 is provided below. Note: of the four parameters to be monitored, three are mean values ($Q_{SWH,y}$, $FC_{SWH,y}$ and $EC_{SWH,y}$) and one is a proportions/percentages ($f_{compliance,y}$).

In order to calculate the required sample size estimates for the mean values and the proportions are required. Furthermore, the standard deviation needs to be assumed in case of sampling for a mean value. For the first monitoring period, the values as described below are applied. For the following monitoring periods, the estimates shall be adjusted taken the results of the previous monitoring period(s) into account.

For the parameters $Q_{SWH,y}$, $FC_{SWH,y}$ and $EC_{SWH,y}$ the following equations[28] are applied:

$$n \geq \frac{z^2 \times N \times V}{(N-1) \times precision^2 + z^2 \times V} \qquad (32)$$

Where:

$$V = \left(\frac{SD}{mean}\right)^2 \qquad (33)$$

n	=	Number of elements to be sampled
N	=	Total number of elements in the population
mean	=	Average value of the parameter that is expected in the total population
SD	=	Standard deviation of the parameter that is expected in the total population
z	=	Constant referring to the level of confidence (e.g. 1.645 for 90 % confidence and 1.96 for 95 % confidence.
precision	=	Required precision (10% = 0.1)

The determination of the “skip” is done according to equation (34) based on the result of equation (33). The required sample size “n” is calculated for each parameter. The highest value for “n” is then used to calculated the overall required skip as follows:

$$skip \leq \frac{N}{n} \qquad (34)$$

Where:

skip	=	Determiner that defines frequency of sampling (i.e. every k^{th} element will be sampled)
n	=	Number of elements to be sampled

[28] Equation according to the *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation (Agenda of EB 66)*

N	=	Total number of elements

For the parameter $f_{compliance}$ the following equations are applied:

$$c \geq \frac{\frac{SD_B^2}{\bar{p}^2} \times \frac{M}{M-1} + \frac{1}{\bar{u}} \times \frac{SD_W^2}{\bar{p}^2} \times \frac{(\bar{N} - \bar{u})}{(\bar{N} - 1)}}{\frac{precision^2}{z^2} + \frac{1}{M-1} \times \frac{SD_B^2}{\bar{p}^2}} \quad (35)$$

Where:

c	=	Number of clusters (i.e. villages) that should be sampled
M	=	Total number of clusters (i.e. villages) in the population
$\bar{u}$	=	Number of units to be sampled within each cluster (pre-specified by project participants)
$\bar{N}$	=	Average units per cluster
$SD_B{}^2$	=	Unit variance (variance between villages)
$SD_W{}^2$	=	Average of the group variances (average within village variation)
$\bar{p}$	=	Average proportion
z	=	Constant referring to the level of confidence (e.g. 1.645 for 90 % confidence and 1.96 for 95 % confidence.
precision	=	Required precision (e.g. 10% = 0.1)

The parameter $\bar{p}$ is calculated as follows:

$$\bar{p} = \frac{1}{M} \sum_{i=1}^{M} p_i \quad (36)$$

Where:

M	=	Total number of clusters (i.e. villages) in the population
$\bar{p}$	=	Average proportion
p_i	=	Proportion within each cluster (i.e. village)

The parameter $SD_B{}^2$ is calculated as follows:

$$SD_B^2 = \frac{1}{M-1} \sum_{i=1}^{M} (p_i - \bar{p})^2 \quad (37)$$

Where:

$SD_B{}^2$	=	Unit variance (variance between villages)
M	=	Total number of clusters (i.e. villages) in the population

$\bar{p}$	=	Average proportion
p_i	=	Proportion within each cluster (i.e. village)

The parameter $SD_W{}^2$ is calculated as follows:

$$SD_W{}^2 = \frac{1}{M}\sum_{i=1}^{M}\left(p_i\left(1-p_i\right)\right) \tag{38}$$

Where:

$SD_W{}^2$	=	Average of the group variances (average within village variation)
M	=	Total number of clusters (i.e. villages) in the population
p_i	=	Proportion within each cluster (i.e. village)

The calculation of the required sample size is illustrated below for a 90/10 level of confidence and precision.

$Q_{SWH,y}$:

$$n \geq \frac{1.645^2 \times 75{,}000 \times V}{(75{,}000-1)\times 0.1^2 + 1.645^2 \times V} \geq 8 \tag{39}$$

Where:

$$\mathrm{V} = \left(\frac{100}{600}\right)^2 \tag{40}$$

A sample size of 30 would be sufficient to achieve the required confidence/precision for $Q_{SWH,y}$. The anticipated mean of $Q_{SWH,y}$ for ex-ante emissions reduction purposes is 600 kWh and the standard deviation is 100 kWh.

$FC_{SWH,y}$:

$$n \geq \frac{1.645^2 \times 75{,}000 \times V}{(75{,}000-1)\times 0.1^2 + 1.645^2 \times V} \geq 68 \tag{41}$$

Where:

$$\mathrm{V} = \left(\frac{50}{100}\right)^2 \tag{42}$$

A sample size of 68 would be sufficient to achieve the required confidence/precision for $FC_{SWH,y}$. The anticipated mean of $FC_{SWH,y}$ for ex-ante emissions reduction purposes is 100 litres and the standard deviation is 50 litres.

$EC_{SWH,y}$:

$$n \geq \frac{1.645^2 \times 75{,}000 \times V}{(75{,}000 - 1) \times 0.1^2 + 1.645^2 \times V} \geq 11 \qquad (43)$$

Where:

$$\mathrm{V} = \left(\frac{10}{50}\right)^2 \qquad (44)$$

A sample size of 11 would be sufficient to achieve the required confidence/precision for $EC_{SWH,y}$. The anticipated mean of $EC_{SWH,y}$ for ex-ante emissions reduction purposes is 50 kWh and the standard deviation is 10 kWh.

The highest required sample size has been determined to be 68. In order to compensate for meter failures this number is increased by 10% resulting in a sample size of 75.

The skip is then calculated to be as follows:

$$skip \leq \frac{75{,}000}{75} = 1000 \qquad (45)$$

As a result the SHW systems with the serial ID numbers 1, 1001, 2001, 3001 … 74001 will be equipped with meters.

$\mathbf{f_{compliance,y}}$:

The following values are applied/result from the calculation of the sample size.

c	=	Number of clusters (i.e. villages) that should be sampled	44
M	=	Total number of clusters (i.e. villages) in the population	120
$\bar{u}$	=	Number of units to be sampled within each cluster (pre-specified by project participants)	10
$\bar{N}$	=	Average units per cluster	50
${SD_B}^2$	=	Unit variance (variance between villages)	0.0237
${SD_W}^2$	=	Average of the group variances (average within village variation)	0.2295
$\bar{p}$	=	Average proportion	0.46
z	=	Constant referring to the level of confidence (e.g. 1.645 for 90 % confidence and 1.96 for 95 % confidence.	1.645
precision	=	Required precision (e.g. 10% = 0.1)	0.1

[The numbers presented here are taken from Example 4 of the *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation (Agenda of EB 66)*. Project participants are recommended to use a spreadsheet program to calculate the related numbers and provide reverence to such a document in the sampling plan.]

[The procedure for increasing the sample size in order to compensate for non response might be taken from the example of cook stoves in Chapter 5.1.]

> Please note: due to the number of clusters (here villages) and the related required calculations in order to determine the sampling size, project participants should develop a spreadsheet to conduct and document the calculations. Especially, the number of sampled elements within each cluster will have a significant impact on the overall required number of sampled elements. Hence, the decision on how many elements per cluster will be sampled should be selected by changing the pre-defined number and comparison of the different results. See also "Example 4 Multi-stage sampling" of the *Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation (Agenda of EB 66).* Further information may be provided according to the specific project circumstances. Please see the example of cook stoves in Chapter 0 for further recommendations.

Sampling design: v. Sampling frame

Generally, the above mentioned parameters are sampled among all distributed SWH systems separately for each CPA. Hence, the overall sampling frame consists of all installed SWH systems per CPA.

> Further information may be provided according to the specific project circumstances. Please see example Cook stoves in Chapter 5.1 for further recommendations.

Data:

i. Field measurements

[needs to be developed according to the specific project circumstances]

> Please see example cook stoves in Chapter 5.1 for further recommendations.

Data: ii. Quality assurance / Quality control

[needs to be developed according to the specific project circumstances]

> Please see example cook stoves in Chapter 5.1 for further recommendations.

Data: iii. Analysis

[needs to be developed according to the specific project circumstances]

> Please see example cook stoves in Chapter 5.1 for further recommendations.

Implementation :i. Implementation plan

[needs to be developed according to the specific project circumstances]

> Please see example cook stoves in Chapter 5.1 for further recommendations.

5.3 Sampling for a CFL project

Project description

Project description:	
Type:	SSC CDM project activity
Project type:	CFLs
Methodology:	AMS-II.J version 4.0
Country:	Single-country
Sampling approach:	Stratified random sampling
Cross CPA-sampling:	Not applicable

The purpose of the project activity is the distribution of compact fluorescent lamps (CFLs) in a single country. The CFLs will replace existing incandescent lamps (ICLs) in residential applications.

Sampling plan

The methodology requires ex post monitoring surveys in order to determine a Lamp Failure Rate (LFR). Therefore, a representative sampling will be undertaken based on the sampling requirements as defined by AMS II.J v1 and the "Standard for sampling and surveys for CDM project activities and programme of activities" (the *Sampling standard*).

> Further information may be provided according to the specific project circumstances. Please see the example of cook stoves in Chapter 5.1 for further recommendations.

Sampling design: i. Objectives and reliability requirements

The objective is to obtain a reliable estimate of the LFR. According to the methodology, this will be carried out in the first year and afterwards once every 3 years.

The following requirements will be followed according to the methodology:

- The sampling size is determined by minimum 90% confidence interval and the 10% maximum error margin; the size of the sample shall be no less than 100;
- Sampling must be statistically robust and relevant i.e., the survey has a random distribution and is representative of target population (size, location);
- The method to select respondents for interviews is random;
- The survey is conducted by site visits;
- Only persons over age 12 are interviewed;
- The project document must contain the design details of the survey.

Hence, the following parameters are subject to sampling:

Parameter	Description of parameter	Confidence/precision level (frequency of sampling)[b]	Sampling approach
LFR_1	LFR in the first year.	90/10 will be applied according to the methodology.	Stratified sampling
LFR_4	LFR in year 4.	90/10 will be applied according to the methodology.	Stratified sampling

LFR_7	LFR in year 7.	90/10 will be applied according to the methodology.	Stratified sampling
LFR_{10}	LFR in year 10.	90/10 will be applied according to the methodology.	Stratified sampling

> Further information may be provided according to the specific project circumstances. Please see example Cook stoves in Chapter 5.1 for further recommendations.

<u>Sampling design: ii. Target population</u>

The overall target population consist of the CFLs distributed as part of the project activity.

> Further information may be provided according to the specific project circumstances. Please see example Cook stoves in Chapter 5.1 for further recommendations.

<u>Sampling design: iii. Sampling method</u>

The parameter LFR is sampled based on a stratified sampling approach. The strata represent the cities in which the CFLs will be distributed.

> Further information may be provided according to the specific project circumstances. Please see example Cook stoves in Chapter 5.1 for further recommendations.

<u>Sampling design: iv. Sample size</u>

The size of the sample is determined by the requirement to achieve 90/10 confidence/precision for the estimation of the proportion of the parameter investigated.

An overview of the estimated sample sizes for a population of 150,000 CFLs applying a level of 90/10 is provided below.

For the first survey, the values as described below are applied. For the following surveys the presented estimates shall be adjusted taken the results of the previous surveys into account.

Town	**Number of CFLs distributed in each town**	***LFR_1***	***LFR_4***	***LFR_7***	***LFR_{10}***
A	100,000	5%	25%	45%	85%
B	75,000	6%	26%	46%	86%
C	50,000	8%	28%	48%	88%
D	25,000	5%	25%	45%	85%

The sample size is calculated as follows:

$$n \geq \frac{z^2 \times N \times V}{(N-1) \times precision^2 + z^2 \times V} \tag{46}$$

Where:

$$\mathrm{V} = \frac{SD^2}{\bar{p}^2} \tag{47}$$

n	=	Number of elements to be sampled
N	=	Total number of elements in the population
$\bar{p}$	=	Overall proportion
SD^2	=	Overall variance
z	=	Constant referring to the level of confidence (e.g. 1.645 for 90 % confidence and 1.96 for 95 % confidence.
precision	=	Required precision (e.g. 10% = 0.1)

For the parameter SD^2 the following equation is applied:

$$SD^2 = \frac{\sum_i (g_i \times p_i (1 - p_i))}{N} \tag{48}$$

Where:

SD^2	=	Overall variance
N	=	Total number of CFLs in the population
g_i	=	Number of CFLs distributed in town *i*
p_i	=	Average proportion of CFLs found not in operation in town *I* (LFR)

For the parameter $\bar{p}$ the following equation is applied:

$$\bar{p} = \frac{\sum_i (g_i \times p_i)}{N} \tag{49}$$

Where:

$\bar{p}$	=	Overall proportion
N	=	Total number of CFLs in the population
g_i	=	Number of CFLs distributed in town *i*
p_i	=	Average proportion of CFLs found not in operation in town *I* (LFR)

In order to determine the sample size for each stratum (i.e. town) the following equation is applied:

$$n_i = \frac{g_i \times n}{N} \qquad (50)$$

Where:

n_i	=	Number of sampled CFLs within town *i*
g_i	=	Number of CFLs distributed in town *i*
n	=	Total number of CFLs that needs to be sampled
N	=	Total number of CFLs in the population

The following values are applied/result from the calculation of the sample size.

Town	**Number of CFLs distributed in each town**	***Sample size for LFR_1***	***Sample size for LFR_4***	***Sample size for LFR_7***	***Sample size for LFR_{10}***
A	100,000	1,711	310	128	18
B	75,000	1,284	232	96	14
C	50,000	856	155	64	9
D	25,000	428	78	32	5
Total	250,000	4,279	775	320	46

[Project participants are recommended to use a spreadsheet program to calculate the related numbers and provide reverence to such a document in the sampling plan.]

[The procedure for increasing the sample size in order to compensate for non response might be taken from the example cook stoves in Chapter 5.1.]

> Project participants should develop a spreadsheet to conduct and document the calculations. Further information may be provided according to the specific project circumstances. Please see the example of cook stoves in Chapter 5.1 for further recommendations.
>
> The available stratified information is used to apply a stratified sampling approach which could be more accurate than a simple random approach.

Sampling design: v. Sampling frame

The sampling frame consists of all distributed CFLs as part of the project activity.

> Further information may be provided according to the specific project circumstances. Please see the example of cook stoves in Chapter 5.1 for further recommendations.

Data: i. Field measurements

[needs to be developed according to the specific project circumstances]

> Please see the example of cook stoves in Chapter 5.1 for further recommendations.

Data: ii. Quality assurance / Quality control

[needs to be developed according to the specific project circumstances]

> Please see the example of cook stoves in Chapter 5.1 for further recommendations.

Data: iii. Analysis

[needs to be developed according to the specific project circumstances]

> Please see the example of cook stoves in Chapter 5.1 for further recommendations.

Implementation: i. Implementation plan

[needs to be developed according to the specific project circumstances]

> Please see the example of cook stoves in Chapter 5.1 for further recommendations.

Annex I

This annex provides an overview of possibly useful software applications that can be used for sampling and surveys[29].

Type of software	Description, Pros & Cons, Examples, Web sources
Spreadsheet programs	**Description:** Spreadsheet programs are widely used for various purposes. In the context of CDM projects or PoAs, typically emission reduction calculations and financial analysis (e.g. internal rate of return) are presented in spreadsheet formats. **Pros & Cons:** + Software in many situations already available and in use. + Experienced staff in many situations already available. + Linkage to emission reduction calculation relatively easy. − Limitation of spreadsheets regarding the total number of elements (e.g. the maximum number of rows per sheet in Microsoft Excel 2007 is 65,536. If a project consists of more elements, use becomes very inconvenient (e.g. wrapping of rows, use of several spreadsheets to overcome the limitation)). − Use of database and sampling functionality often requires more experienced staff (e.g. programming of Macro/Visual Basic). − Data input through a number of users (at the same time) is (normally) not impossible. **Examples:** E.g. Microsoft Excel, Open Office Calc, etc. **Web sources:** -

[29] Naming of specific software applications is only exemplary and does not imply any recommendation to use or not to use a specific product.

Type of software	Description, Pros & Cons, Examples, Web sources
Web-based number generators	**Description:** Applications that can be found on the internet. After entering information regarding the total number of elements and the required number of samples a list of randomly selected numbers is returned. **Pros & Cons:** + Simple to handle − Detailed calculation procedure often unknown. − Provides only a very small part of the overall required software. − Web source might not be available for the total required period (i.e. the total crediting period). **Examples:** See web sources below. **Web sources:** http://www.random.org/ http://www.psychicscience.org/random.aspx

Type of software	Description, Pros & Cons, Examples, Web sources
Web-based sample size calculators	**Description:** Applications that can be found on the internet. After entering information regarding the sampling requirements the related sample size is returned. **Pros & Cons:** + Simple to handle − Detailed calculation procedure often unknown. Hence, compliance with the related CDM requirements is difficult to judge. − Provides only a very small part of the overall required software. − Web source might not be available for the total required period (i.e. the total crediting period). **Examples:** See web sources below. **Web sources:** http://www.raosoft.com/samplesize.html

Type of software	Description, Pros & Cons, Examples, Web sources
Database programs	**Description:** Database programs are used for various purposes (e.g. customer database, enterprise resource planning, etc.). **Pros & Cons:** + Software in some situations already available and in use. + Experienced staff in some situations already available. + Database programs provide the required functionality regarding handling of a (huge) set of data such as a list of households with e.g. ID, address, etc. − Experienced staff required to set-up the general structure of the database in advance of the use of the database. − Use of sampling or emission reduction calculation functionality within the database program often requires more experienced staff (e.g. programming of Macro/Visual Basic/SQL). − Data input through a number of users (at the same time) is often limited. **Examples:** E.g. Microsoft Office Access, Open Office Base, Oracle Database, etc. **Web sources:** -

Type of software	Description, Pros & Cons, Examples, Web sources
Web-based databases	**Description:** **Pros & Cons:** + Allows data input from various users at the same time. + Accessibility via the internet. + Possibility to include other interfaces such as hand-held, SMS-Services etc. − Requires substantial development (programming) effort in advance. − Requires at least temporarily internet connection for staff/ users of the database. **Examples:** - **Web sources:** -

Type of software	Description, Pros & Cons, Examples, Web sources
Statistical programs	**Description:** **Pros & Cons:** + Offer (almost) all of required and useful statistical calculations. − Requires experienced staff. − Provides only a small part of the overall required software. **Examples:** E.g. SAS, Stata, SPSS, R, etc. **Web sources:** http://www.hcp.med.harvard.edu/statistics/survey-soft/#Online http://www.freestatistics.info/stat.php

Type of software	Description, Pros & Cons, Examples, Web sources
Supporting programs	**Description:** Various types of software. **Pros & Cons:** + Various different software applications are available. − Requires often experienced staff. − Provides often only a small part of the overall required software. **Examples:** E.g. Census and Survey Processing System (CSPro) **Web sources:** http://www.census.gov/population/international/software/cspro/

Annex II

A simple random number generator using Microsoft Excel/Visual Basic[30]

If the distributed elements have a consecutive ID number the following stepwise approach allows generation of list of randomly selected elements. The program code removes already duplicates. When using web based random number generators project participants should check if such function is available.

Step 1: Start Microsoft Excel and open a blank/empty workbook

Step 2: Open Visual Basic (e.g. press Alt + F11)

Step 3: Double-click on *Sheet1* in the visual basic window (Microsoft Excel Objects) and copy the following program code to the code window.

Example of a simple random number generator

```
Sub random_number_generator()

population = InputBox("Enter total number of elements ", "Total number ")
sample = InputBox("Enter sample size", "Sample size")
Randomize

For n = 1 To sample
Number = Int((population * Rnd) + 1)

For k = 1 To sample
check = 0

If Number = Worksheets(1).Cells(k, 1).Value Then
k = sample
check = 1
n = n - 1
Else: check = 0
End If
Next k

If check = 0 Then
Worksheets(1).Cells(n, 1) = Number
End If

Next n

End Sub
```

Step 4: Run the code (e.g. press F5)

[30] The program code has been developed using Microsoft Office Excel 2007. No representation or warranty is offered as to the accuracy of the assumptions or methodologies used, and no liability in respect thereof shall be taken.

Step 5: Enter the total number of elements in the population/sample frame (e.g. number of all distributed cook stoves) and press *OK.*

Step 4: Enter the required sample size (e.g. number of cook stoves that are going to be monitored) and press *OK.*

Step 5: Column *A* of *Sheet1* provides a list of a random sample (e.g. ID numbers of the distributed cook stoves)

Attention:

Before running the program code again, make sure the sample list is empty (i.e. column *A* of *Sheet1* is empty).

In order to save the document you might need to adjust your macro security settings temporarily.

References

Hayashi, Daisuke; Michaelowa, Axel; Dransfeld, Björn; Niemann, Mareike; Marr, Marc André; Müller, Nicolas; Wehner, Stefan; Krey, Matthias; Oppermann, Klaus; Neufeld, Carolyn S. (2010): PoA BLUEPRINT BOOK: Guidebook for PoA coordinators under CDM/JI; 2. Edition; KfW Bankengruppe, Frankfurt

Papula, Lothar (1997): Mathematik für Ingenieure und Naturwissenschaflter Band 3; Vieweg; Braunschweig, Wiesbaden

Schwarze, Jochen (2001): Grundlagen der Statistik II; Verl. Neue Wirtschafts-Briefe; Herne, Berlin

United Nations Framework Convention on Climate Change (2012): CDM; available at: http://cdm.unfccc.int/

Main UNFCCC documents referred to in this manual:

AMS-I.E.: Switch from Non-Renewable Biomass for Thermal Applications by the User: Version 4.0

AMS-I.J.: Solar water heating systems (SWH): Version 1.0

AMS-II.C.: Demand-side energy efficiency activities for specific technologies: Version 13.0

AMS-II.G.: Energy efficiency measures in thermal applications of non-renewable biomass: Version 3.0

AMS-II.J.: Demand-side activities for efficient lighting technologies: Version 4.0

Draft Best Practices Examples: Focusing on Sample Size and Reliability Calculation (Agenda EB66)

Standard for sampling and surveys for CDM project activities and programme of activities: Version 2.0 (EB 65)