

CDM-EB79-AA-A10

Road-testing and revised guidelines for quality assurance and quality control of data used in the establishment of standardized baselines

Version 01.0



United Nations
Framework Convention on
Climate Change

COVER NOTE

1. Procedural background

1. The information note on the lessons learned from road-testing of guidelines for quality assurance and quality control of data used in the establishment of standardized baselines (QA/QC guidelines) and the draft revised QA/QC guidelines are developed in accordance with the 2014 workplan of the Executive Board of the clean development mechanism (hereinafter referred to as the Board).
2. The secretariat developed the revised draft of the guidelines based on lessons learned from: (i) evaluation of bottom-up submissions (road-testing) of standardized baselines; (ii) practitioners who have been involved in the development of sector-specific standardized baselines and participated in the practitioners' workshop organized by the secretariat on 9 July 2013; and (iii) inputs received from the Methodologies Panel (Meth Panel) and Small-Scale Working Group (SSC WG).
3. The Board at its seventy-fifth meeting reviewed the first draft revision and requested further improvement, focusing on simplicity, internal consistency of the guideline and clarity in the best practice examples. The Board at its seventy-sixth meeting requested the secretariat to clarify how stakeholder consultation should be conducted and what is expected from such consultations.
4. As per the request of the Board at its seventy-eighth meeting (EB78 report, paragraph 10), this revised draft of the guidelines is submitted for the Board's consideration as a joint recommendation of the Meth Panel and the secretariat.

2. Purpose

5. The purpose of the information note is to inform the Board on:
 - (a) The lessons learned through road-testing of QA/QC guidelines up to March 2014 since their first adoption in March 2012; and
 - (b) The key issues that have been considered in the revision of the QA/QC guidelines, which is submitted for the consideration of the Board at this meeting.
6. The purpose of the revision of the QA/QC guidelines is to:
 - (a) Improve clarity through editorial changes while bringing further simplicity, consistency, objectivity and transparency to the guideline;

- (b) Improve user-friendliness by providing further guidance on practical challenges identified, by presenting more flexible best practice examples and by introducing a checklist for step-by-step quality assurance/quality control activities and modified samplers.

3. Key issues and proposed solutions

3.1. Lessons learned from road-testing

7. Twelve standardized baselines (SBs) have been submitted for consideration. Four SBs were prepared using the “Guidelines for the establishment of sector specific standardized baseline” (SB guidelines) for charcoal sector, cement sector, rice mill sector and power sector, while eight SBs were prepared for the power sector using the “Tool to calculate emission factor for an electricity system” (grid tool). At present, a total of four SBs have been approved by the Board, which includes one each in the charcoal and rice mill sectors using the SB guidelines and the remaining two in the power sector using the grid tool.
8. Six SBs in the waste sector for landfill gas recovery and methane destruction are undergoing an initial assessment while three designated national authorities (DNAs) have expressed their interest in developing the SBs in the waste sector.
9. Issues were identified while assessing these submissions and key lessons were incorporated in the revision of the QA/QC guidelines. These include the following:
 - (a) **Primary versus secondary data sources:** Secondary sources of data were predominantly used for the establishment of the submitted SBs. Only one SB for energy use in the rice mill sector of Cambodia (PSB0004) used primary data obtained from a survey. The SB for clinker production in Ethiopia (PSB0002) used a combination of primary and secondary data sources. Instead of prioritizing one data source over another, the QA/QC guidelines should provide more flexibility to DNAs with clear guidance on how to select credible data sources by taking into account the country-specific or sector-specific situations;
 - (b) **Conservativeness:** Where data were missing and not recorded, conservative approaches were applied, by using values of the top 20% best-performing facilities in the country (e.g. PSB0002), applying conservative values (e.g. Uzbekistan grid emission factor applies default efficiency of the power generation technology as per the “Tool to calculate emission factor for an electricity system”) or IPCC default values (e.g. multiple SBs on grid emission factor apply IPCC default values for net calorific values and emission factors of fossil fuels). This implies that SB developers have ensured the data quality objectives through various conservative approaches as per the examples of the current QA/QC guidelines. The revision of the QA/QC guidelines may further elaborate best practices for conservativeness in a flexible manner;
 - (c) **Data template:** The SB procedure (“Procedure for development, revision, clarification and update of standardized baselines”) requires SB developers to use a data template published by the secretariat or to submit a new template for its publication before the submission of SBs. However, eight SBs for the power sector used the data template available from the grid tool and three early SBs were submitted before the SB procedure included this requirement. There are not

enough cases to identify key issues/lessons learned related to data templates yet. However, the QA/QC guidelines should provide further guidance on how to develop a data template and highlight why the use of a data template is recommended as a best practice;

- (d) **Public consultation:** The majority of the submissions did not conduct a public consultation process but in many cases publicly available sources which present aggregated data were used. The QA/QC guidelines should provide further guidance on the public consultation process;
- (e) **Further simplification:** SBs for landfill gas recovery and destruction in the waste sector do not have a data-intensive process to be followed and most of the SB submissions in this sector did not need to provide all documents as per the documentation provisions of the approved QA/QC guidelines. The QA/QC guidelines should be further simplified for such situations.

3.2. Draft guidelines for quality assurance and quality control of data used in the establishment of standardized baselines

10. The development of standardized baselines requires a data-intensive process where DNAs face a variety of challenges, particularly in collecting data and ensuring data quality. Through lessons learned from road-testing, feedback from stakeholders and other practitioners, and inputs from the Meth Panel and SSC WG, the following key issues are identified and solutions are proposed:

- (a) **Limited data availability and accessibility:** The most common and difficult challenge is limited data availability and accessibility. The draft revision includes some proposals for DNAs to address this issue, such as possible institutional arrangements and partnerships with sectors;
- (b) **Preference of primary data:** As learned from road-testing, in certain conditions where primary data are not available or difficult to obtain, secondary data could be used. However, primary data should be preferred whenever feasible. The draft revision removes the order of priority of data sources, provides further guidance on the selection of credible data sources, highlights the benefits of primary data sources and addresses some issues (e.g. inconsistency or updating of data) raised when secondary data are used;
- (c) **Best-practice examples and checklist:** To facilitate the implementation of the guidelines, the draft includes best-practice examples for five key areas and a checklist for each stage of the QA/QC process (planning, pre-submission QC, post-submission QC, compiling/outcomes and QC report) for DNAs by taking into account practicability. The secretariat plans to upload best practices on data management on the CDM webpage at regular intervals based on lessons learned from real cases;
- (d) **Uncertainty:** It is difficult to estimate the level of uncertainty and to address uncertainty in an efficient manner, which requires further technical and innovative approaches. The proposed draft highlights the main sources of uncertainty and some options to reduce uncertainty. However, the risk of human error and uncertainty in data collection/processing/management cannot be totally avoided or corrected, so the proposed draft highlights the importance of transparency

(e.g. public consultation). Further work may be taken up in future to develop some robust approaches to address issues related to uncertainty;

- (e) **Further flexibility:** In order to promote simplicity and flexibility, the general provisions have been removed and more practical guidance is provided. Another reason to delete the section on general provisions is that most of the guidance provided in this section is covered under the recently approved standard for determining coverage of data and validity of standardized baselines. To avoid confusion between best practices/examples and requirements, the clarity is improved;
- (f) **Further clarity and consistency:** More objective and concise guidance and examples are included and consistency is improved.

4. Impacts

- 11. The revision will contribute to the better understanding and easier implementation of data collection/management for the establishment of standardized baselines through improved clarity and user-friendliness.
- 12. The revision of general provisions and documentation provisions will require the revision of the “Procedure for development, revision, clarification and update of standardized baselines”, in particular the revision of its appendix 1.

5. Subsequent work and timelines

- 13. The secretariat will revise the relevant procedure once the Board adopts this document in 2014. The assistance provided by the secretariat (including through the Regional Collaboration Centres) to DNAs in their bottom-up submissions of SBs also helps to promote the effective implementation of QA/QC guidelines and develop their capacity. This work will continue as a routine operation.
- 14. The secretariat, together with the Meth Panel/SSC WG, will continue to capture lessons based on road-testing which can eventually be used for further improvements.

6. Recommendations to the Board

- 15. The secretariat recommends that the Board take note of the information presented in the note.
- 16. The secretariat recommends that the Board adopt the revised QA/QC guidelines, which will come into effect immediately upon adoption.

7. References

- 17. The guidelines are related to, and hence should be implemented in conjunction with the latest version of:
 - (a) The “Guidelines for establishment of sector-specific standardized baselines”;
 - (b) The “Procedure for development, revision, clarification and update of standardized baselines”;

- (c) The “Standard for determining coverage of data and validity of standardized baselines”;
- (d) The “Guideline for the establishment of standardized baselines for afforestation and reforestation project activities under the CDM”.

TABLE OF CONTENTS	Page
APPENDIX 1. INFORMATION NOTE ON ROAD TESTING	8
APPENDIX 2. REVISED GUIDELINES FOR QUALITY ASSURANCE AND QUALITY CONTROL OF DATA USED IN THE ESTABLISHMENT OF STANDARDIZED BASELINES.....	13
ATTACHMENT 1. DATA DELIVERY PROTOCOL (SAMPLER)	28
ATTACHMENT 2. DATA TEMPLATE (SAMPLER).....	30
ATTACHMENT 3. QUALITY CONTROL (QC) REPORT (SAMPLER).....	31
ATTACHMENT 4. ASSESSMENT REPORT (SAMPLER) CHECKLISTS FOR DNAS	32

DRAFT

Appendix 1. Information note on road testing

1. Introduction

1. The approved management plan (MAP) for 2014–2015 and the 2014 workplan of the Executive Board of the clean development mechanism (hereinafter referred to as the Board) requires that when revising the guidelines for quality assurance and quality control of data used in the establishment of standardized baselines (QA/QC guidelines), lessons learned from the use of these guidelines should be implemented in the revision of the QA/QC guidelines and a report on road-testing of these guidelines should be submitted for the Board's consideration.

2. Lessons learned from road-testing and corresponding revision of the QA/QC guideline

2.1. Lessons learned through evaluation and assessment of standardized baseline proposals

2. Twelve standardized baselines (SBs) have been submitted for consideration. Four SBs were prepared using the "Guidelines for the establishment of sector specific standardized baseline" (SB guidelines) for the charcoal sector, cement sector, rice mill sector and power sector, while eight SBs were prepared for the power sector using the "Tool to calculate emission factor for an electricity system" (grid tool).
3. At present, a total of four SBs have been approved by the Board, which includes one each in the charcoal and rice mill sectors using the SB guidelines and the remaining two in the power sector using the grid tool.
4. While assessing these submissions and communicating with stakeholders and other practitioners, several lessons were learned and some key issues were identified by the secretariat, the Methodologies Panel (Meth Panel) and the Small-Scale Working Group (SSC WG).
5. The following is an analysis of lessons learned while assessing the SB submissions.

2.1.1. Charcoal sector SB submission (PSB0001)

6. This SB was developed using the SB guidelines and was the first submission evaluated by the secretariat and the SSC WG. The lessons learned were as follows.
7. The key parameters collected were the charcoal yield and methane generation for a specific set of charcoal-making technologies based on publicly available secondary data sources. Although the data in the sources were collected for different countries, they were proven to be relevant to Uganda because these countries have the same level socio-economic conditions with Uganda and the activity data collected were only for the traditional unimproved charcoal-making technologies which were the same as those producing over 90% of the charcoal used by households and small/medium enterprises (SMEs) in Uganda.

8. This leads to the conclusion that under certain conditions secondary data sources for some technologies may be representative for several countries as long as relevance is proven and justified.
9. The data sources used have been developed by various international institutions and agencies such as international development institutions, universities in developed or developing countries, multilateral agencies, especially United Nations agencies, research institutes, and ministries or agencies of national governments. Based on the quality of data and the criteria, only credible data sources were used for establishing the SB with justification.

2.1.2. Cement clinker sector SB submission (PSB0002)

10. This proposed standardized baseline (PSB), developed based on the SB guidelines, is still under evaluation. The experience of this SB led to several lessons learned as follows.
11. The most important parameters (such as total clinker production; total fuel consumption as per type of fuel; kiln technology used in the host country and its investment cost; installed capacity of the cement plants in the host country; date of commencement of the cement plants in the host country) were collected from the primary data sources from the respective cement manufacturing plants in Ethiopia through the Ministry of Industry and sector institutions, i.e. the Climate Resilient Green Economy (CRGE) unit. Some other key parameters such as emission factor of each type of fuel and net calorific value of each type of fuel were collected based on a secondary data source, i.e. the IPCC 2006 report.
12. This indicates that primary and secondary data sources can be used in a combined manner as long as the quality of data sources can be justified and a combination of the data sources does not lead to biases.
13. Where data for a specific parameter related to clinker production and fuel consumption were not provided from individual plants, the values were determined by a conservative approach – average value of the top 20% best cement plants in the region.
14. It is to be noted that the approach for handling missing data by applying conservative values was widely used and considered appropriate.

2.1.3. Rice mill sector SB submission (PSB0004)

15. This SB was developed using the SB guidelines and approved by the Board as ASB0004. The lessons learned were as follows.
16. This SB used a primary data source through a survey, based on the sampling standard (“Sampling and surveys for CDM project activities and programme of activities”) referred to the QA/QC guidelines. Although this SB was submitted before the QA/QC guidelines were approved, the SB developer as well as the DNA provided further data/information in accordance with the relevant standard.
17. It is to be noted that sampling approaches are successfully applied for small-scale scattered emitters provided that the relevant provisions of the sampling standard are followed.

2.1.4. Power sector standardized baselines (PSB0003 Southern African Power Pool, PSB0005 Uzbekistan, PSB0006 Belize, PSB0007 Cape Verde, PCB008 Mauritius, PSB0009 Sao Tome and Principe, PSB0010 Dominican republic, PSB0011 Armenia, PSB0012 Uganda)

18. One of the SBs for the power sector in Cape Verde (officially the Republic of Cabo Verde) was developed using the SB guidelines and the other eight SBs were developed using the “Tool to calculate emission factor for an electricity system”. The lessons learned from these SBs were as follows.
19. Various secondary data sources such as officially published research papers, governmental data and statistics, and data from various national agencies such as energy agencies and statistics institutes were used for deriving the key parameters.
20. Where data were missing and not recorded, conservative approaches were applied such as using the default efficiency of the power generation technology as per the “Tool to calculate emission factor for an electricity system” (e.g. Uzbekistan) and IPCC default values (e.g. many of the grid emission factor SBs apply IPCC default values for net calorific values and emission factors of fossil fuels).
21. For data obtained from secondary sources, the existing QA/QC systems of the institutions collecting the data were described and justifications for the data quality were provided based on existing data management practices (e.g. SBs for grid emission factors).

2.1.5. Waste sector standardized baselines

22. There were six submissions received for landfill gas recovery and methane destruction in the waste sector and another three DNAs expressed their interest in developing SBs in this sector. These PSBs were developed using the SB guidelines. The submissions are undergoing the first evaluation process. The following are the lessons learned so far.
23. The majority of the submissions stipulate that there are no regulations in the countries that mandate the destruction of methane from landfills quoting national institutions and sources.
24. Taking into consideration that the only information that is provided in the submission is on the existence or absence of national regulations for the capture and destruction of methane, it may be concluded that not all of the QA/QC guidelines are relevant or applicable.

2.1.6. Common lessons learned from all the SB submissions

25. None of the 12 submissions has applied a sector-specific data template for collecting, compiling and/or calculating the baselines parameters. However, eight SBs for the power sector used the data template available from the grid tool and three early SBs were submitted before the SB procedure included this requirement.
26. Most of the submissions did not conduct a public consultation process, but in many cases publicly available sources which present aggregated data were used.

27. In most of the cases the SB submissions did not develop a new QA/QC system for data collection, compilation and calculations but elaborated the design and the elements of existing QA/QC systems of the relevant authorities that have provided the data.
28. Secondary sources of data were predominantly used for establishment of the submitted SBs, such as officially published research papers, governmental data and statistics, and data from various national agencies such as energy agencies and statistics institutes. Only one SB for energy use in the rice mill sector of Cambodia (PSB0004) used primary data obtained from a survey and the SB for clinker production in Ethiopia (PSB0002) used a combination of primary and secondary data sources.
29. Most SB submissions did not need to provide all documents as per the documentation provisions of the approved QA/QC guidelines and the documentation provisions were not consistent with the SB procedure.

3. Key conclusions

30. Various secondary data sources can be used for deriving baseline parameters. Since it is not always feasible to recommend the use of either primary or secondary data sources for the development of SBs, in the revised version of the QA/QC guidelines further guidance on how to select credible data sources should be provided.
31. Where data were missing and not recorded conservative approaches were applied as per the examples of the current QA/QC guidelines. The revision of the QA/QC guidelines may further elaborate best practices for conservativeness in a flexible manner.
32. There are not enough cases to identify key issues/lessons learned related to data templates yet. However, the QA/QC guidelines should provide further guidance on how to develop a data template and highlight why the use of a data template is recommended as a best practice.
33. Most of the submissions did not conduct a public consultation process, but in many cases publicly available sources which present aggregated data were used. In the draft revised QA/QC guidelines more elaboration is provided on the public consultation process.
34. Further simplification is needed, in particular focusing on the documentation provisions and a certain case where no intensive data collection is required.

4. Recommendations to the Board

35. The Board may wish to take note of the information presented in this document.

- - - - -

Document information

<i>Version</i>	<i>Date</i>	<i>Description</i>
01.0	14 May 2014	Initial publication as an annex to the annotated agenda of EB 79.
Decision Class: Operational Document Type: Information note Business Function: Governance Keywords: data collection and analysis, quality management, standardized baselines		

DRAFT

Appendix 2. Revised guidelines for quality assurance and quality control of data used in the establishment of standardized baselines

1. Introduction

1.1. Background

1. The Executive Board of the clean development mechanism (CDM) (hereinafter referred to as the Board), at its sixty-second meeting, approved the “Guidelines for the Establishment of Sector Specific Standardized Baselines” (hereinafter referred to as the SB Guidelines) to meet the request from the sixth session of the Conference of the Parties serving as the Meeting of the Parties to the Kyoto Protocol (CMP).
2. The Board at its sixty-third meeting adopted the “Procedure for submission and consideration of standardized baselines”. ~~The Board requested the secretariat to prepare an implementation plan for the procedure for consideration by the Board at its sixty-fifth meeting at the latest. As the use of standardized baselines is voluntary, the Board stressed that the level of stringency of the standardized baselines should be sufficient to ensure a high level of environmental integrity. The Board also requested the secretariat to take into account this principle while developing the work programme to implement the SB Guidelines.~~ The Board at its sixty-fifth meeting agreed to the work programme on the standardized baselines. For the effective implementation of the SB Guidelines, the Board at its sixty-sixth meeting adopted the “Guidelines for quality assurance and quality control of data used in the establishment of standardized baselines”~~se guidelines were developed to ensure the quality of data required for the establishment of the standardized baselines.~~ At its seventy meeting, the Board adopted the “Guideline for establishment of standardized baselines for afforestation and reforestation project activities under the CDM”.
3. The Board at its seventy-first meeting adopted the “CDM two-year business plan and management plan 2013–2014”, which included further revision of the regulatory framework based on the lessons learned from road-testing and based on inputs from stakeholders and relevant research. At its seventy-fifth and seventy-sixth meetings, the Board requested the secretariat to revise the “Guidelines for quality assurance and quality control of data used in the establishment of standardized baselines”, taking into account the inputs provided by the Board.
4. At its seventy-seventh meeting, the Board adopted the “Standard for determining coverage of data and validity of standardized baselines”, and agreed to the work programme on the further development and implementation of standardized baselines.

1.2. Objectives

5. The development of standardized baselines is a data-intensive process. The data required for quantifying baseline greenhouse gas (GHG) emissions in a particular sector include data on activities in a country or a group of countries such as energy usage, industrial production statistics, production technologies, demographic data, process-

related characteristics and mitigation-related practices. Some data might already be available to Designated National Authorities (DNAs) but in many cases it could be challenging for DNAs to entail the process of collect data collection from data providers and ensure the data quality. The collection of data from multiple sources can lead to inconsistencies in levels of details, data formats and data quality. It also poses difficulties in achieving data integrity, data validity and completeness. The purpose of this document is to specify provisions and processes for ensuring data quality objectives and to provide guidance and best practices/examples on practical aspects of data collection, processing, compilation and reporting of data including the use of sector-specific data templates.

6. This document aims to define best practices to meet the data quality objectives given the current state of scientific knowledge and data availability, which could help DNAs improve their institutional capacities for data management.
7. The best practice to ensure data quality as outlined in this document is twofold:
 - (a) Proactively¹ preventing potential risks that could cause quality deterioration, with a well-designed data management system, well-trained personnel and the culture of data quality; and
 - (b) Identifying and formulating data problems and implementing corrective actions, through regular reviews and continuous improvement processes.

2. Scope, applicability, and entry into force

2.1. Scope

8. The “Guidelines for the Quality Assurance and Quality Control of Data used in the Establishment of Standardized Baselines” (hereafter referred to as the QA/QC Guidelines) include the quality control (QC) procedures for ensuring the quality of datasets compiling the required datasets and the quality assurance (QA) procedures for assessing that the QC system is designed and implemented to meet the data quality objectives. ensuring the overall quality of the datasets by assessing the conformity and the effectiveness of the QC system, based on data quality objectives and general provisions. The QA/QC Guidelines also elaborate provisions for the documentation to be submitted in accordance with paragraph 8 of the “Procedure for Submission and Consideration of Standardized Baselines” “Procedure for development, revision, clarification and update of standardized baselines”.
9. The QA/QC procedures/activities and the best practices presented in this document are not exhaustive and it may not always be feasible for DNAs to operationalize them fully. DNAs could adapt the procedures to their own circumstances and apply other good practices as long as best practices are applied it can be demonstrated that the data quality objectives are met.

¹ It is the best practice to check the quality of data as early as possible. DNAs should avoid low-quality data being compiled in the datasets through pre- and post-submission quality checking.

2.2. Applicability

10. The QA/QC Guidelines are applicable to those entities involved in the collection, processing, compilation and reporting of data needed for the establishment of **sector-specific** standardized baselines.² Such entities include:
- (a) DNAs³ that develop, validate, and/or own the datasets used for the establishment of standardized baselines;
 - (b) Designated **Operational Entities (DOEs)** that conduct review procedures to assess the quality of the **data management-QC** system used to establish the standardized baseline;
 - (c) Project participants or other entities⁴ that develop standardized baselines.

2.3. Entry into force

11. This document shall enter into force on [xx 2014].

3. Normative references

12. The guidelines are related to, and hence should be implemented in conjunction with the latest version of:
- (a) The “Guidelines for development of standardized baselines”;
 - (b) The “Guideline for the establishment of standardized baselines for afforestation and reforestation project activities under the CDM”;
 - (c) The “Procedure for development, revision, clarification and update of standardized baselines”;
 - (d) The “Standard for determining coverage of data and validity of standardized baselines”.

4. Definitions ~~Key concepts and data quality objectives~~

13. The definitions contained in the Glossary of CDM terms apply.
14. For the purpose of this document, the following key concepts apply:
- (a) Data quality: data quality is a multi-dimensional concept and is commonly described as the degree to which data are “fit for use”.⁵ Data quality is ensured

² Standardized baselines can be developed using baseline and monitoring methodologies, methodological tools and/or approved guidelines and standards adopted by the CDM Executive Board.

³ DNAs should ensure that the QA/QC Guidelines are applied by all involved entities when the collection and management of data is outsourced.

⁴ The CMP at its sixth session decided that Parties, project participants, as well as international industry organizations or admitted observer organizations through the host country’s DNA may submit proposals for standardized baselines. According to the “Procedure for development, revision, clarification and update of standardized baselines”, the Board may also decide to develop standardized baselines in agreement with the DNA(s) of a Party(ies) through top-down approach.

when it can be demonstrated that the datasets are relevant, complete, consistent, **credible/reliable**, current, accurate and objective. In addition, processing data to derive standardized baselines should be conservative, secure, transparent and traceable;

- (b) Quality control (QC): QC is a system of routine technical activities to be conducted by a DNA to assess and maintain the quality of the datasets **as the data are being compiled**. It begins with pre-submission QC activities **described in paragraph 273 below**, followed by post-submission QC activities, internal review and a summary of the QC implementation (QC report);
- (c) Quality assurance (QA): QA is a system developed by a DNA to ensure that the QC system is designed to meet the data quality objectives below **and the provisions specified in paragraphs 12 to 21** and that **it is** implemented effectively. The conformity and the effectiveness of the QC system are reviewed and the review activities/results are reported (assessment report) by DOEs.⁶ **that are not directly involved in the compilation/development process of the datasets. Reviews are performed on the QC systems/processes used to establish the datasets and verify that objectives and all the provisions specified in the QA/QC Guidelines were met**

5. Data quality objectives and general provisions

16. Data quality objectives

15. The following data quality objectives are intended to guide the implementation of the QA/QC procedures **activities specified in paragraphs 22 to 36**:
 - (a) Relevance: **collect select data and other information appropriate required to for the establishment of sector-specific standardized baselines, which includes These are** mainly activity data and information applicable for the determination of the baseline emissions in a sector;
 - (b) Completeness: include all relevant activity data and information to produce **“true and fair” representative standardized baselines covering the relevant target population (installations/facilities/companies) that contribute to the production of the output of the sector in a region, country or group of countries, based on the scope of the standardized baselines and the level of aggregation chosen. Efforts should be undertaken It requires procedures to avoid, identify and handle missing data (e.g. relevant companies that have been excluded or incomplete data entry);**
 - (c) Consistency: present the **all same** data in the same **definition/scope/format** and make the datasets compatible with other related data; **which enables meaningful comparisons if necessary;**

⁵ Knight, S. & Burn, J. (2005) Developing a framework for assessing information quality on the world wide web. *Information Science Journal* vol. 8. pp.159–172.

⁶ According to the **“Procedure for Submission and Consideration of Standardized Baselines” “Procedure for development, revision, clarification and update of standardized baselines”**, DOEs will prepare an assessment report. In the case where **an eligible** Party (with 10 or fewer registered CDM project activities as of 31 December 2010) does not provide an assessment report in the submission **(allowed only once)**, the secretariat should prepare the assessment **report in accordance with the procedure**.

- (d) Credibility: identify and utilize authoritative data sources. Collected data/information should always reference their sources. Various data sources can may be available, but DNAs should highlight the importance of primary data collection use the most credible data source;

The following indicates the order of priority of data sources:

Data collected from data providers mainly to establish standardized baselines (primary data);

Data officially obtained by DNAs for other purposes;

Data collected by other government authorities;

Data collected by peer-reviewed international statistics documents;

Data collected by research institutes, individual academic research works or United Nations institutions;

Data established based on technology supplier information;

Tip: Best practices to identify and use credible data sources

DNAs may identify which data sources are available and determine which data sources are more accessible and more authoritative through consultation with experts or representatives of the particular sector as well as relevant authorities and institutions collecting data. DNAs may conduct a comparison of available data for a best fit for standardized baselines. The primary data that are collected from data providers to establish standardized baselines can be the best option because DNAs have the ownership for maintaining the data management system, which can provide flexibility in its operation and facilitate in meeting the data quality objectives in an efficient and effective way. When secondary data sources are used, DNAs will need to predetermine procedures to: (i) ensure the consistency of data derived from multiple data sources (e.g. data coverage, definitions and formats); (ii) evaluate the credibility of each data source (e.g. checking how QA/QC activities have been conducted by data providers); and (iii) maintain the data system (e.g. addressing how to update the database or how to incorporate changes in secondary data sources).

The following secondary data sources can be used:

- Data already collected/maintained by DNAs for purposes other than SBs;
- Data collected by other government authorities;
- Data collected by peer-reviewed international statistics documents;
- Data collected by research institutes, individual academic research or United Nations institutions;
- Data collected from relevant manufacturers or suppliers.

- (e) Currentness: utilize the most recent data available in a sector in order to reflect the current economic and technological practices. The currentness of data should be as per the provisions of the “Standard for data coverage and validity of standardized baselines”;
- (f) Accuracy: avoid and reduce errors. DNAs should identify and address uncertainties as far as it is practical and cost-effective. The QC procedures should be well designed to ensure the accuracy, which should focus on the procedures to avoid potential duplications and errors. When a sampling approach is employed to derive the required data, DNAs should apply statistically sound sampling approaches, comply with the sampling requirement for reliability (95/10

confidence/precision for industrial sectors⁷ and 90/10 confidence/precision for other sectors) and select sample sizes in accordance with the latest approved “Standard for sampling and surveys for CDM project activities and programme of activities”;

- (g) Objectivity: avoid biased, prejudiced and partial information. There should be little room for assumptions or differing interpretations in managing the data⁸;
- (h) Conservativeness: while processing data, ensure that any deviation from the QA/QC Guidelines that may lead to an overestimation of the baseline emissions in a sector should be addressed by taking a conservative approach. Wherever other data quality objectives specified above in this document could not be met, a conservative approach (e.g. use of predetermined conservative default values) should be applied to substitute for missing, incomplete, invalid, old or incorrect data. For example, in the case of a lack of accuracy, the average values of the top 20%, based on performance data in the sector could be used;

Tip: Best practices to apply conservative approaches

Various approaches may be used by DNAs to ensure conservativeness as long as they are justified. DNAs may consider the following practices to compensate for missing, incomplete, invalid, old or incorrect data with justification:

- 1) Default values pre-established by government agencies or international organizations (such as IPCC);
- 2) Average values of the top 20% best performing facilities;
- 3) Average values of facilities installed in the last five years;
- 4) Applicable design values provided by manufacturers;
- 5) Best performers' data recognized by sector experts;
- 6) Instead of applying conservative default values, DNAs may provide evidence-based justifications confirming that any deviations (in terms of missing, incomplete, invalid, old or incorrect data) do not lead to an overestimation of the baseline emissions.

- (i) Security: develop procedures for restricted access to the datasets and maintain the security of the datasets. The procedures should include how to identify, process, present and manage confidential data. The data should be open to the public in an anonymous aggregate form;
- (j) Transparency: disclose sufficient and appropriate data and processes to allow monitoring of the quality of the compiled datasets and the generated outcomes. Through public engagement, reviewers may make decisions with reasonable confidence. DNAs should prepare a report that includes: (i) the objectives of public consultation; (ii) processes followed for public consultation; (iii) participants who attended the public consultation; (iv) a summary of the comments provided by stakeholders/experts; and (v) how comments were taken into account;

⁷ Based on the category of sectoral scopes in the CDM, industrial sectors include energy industries, manufacturing industries, chemical industries, mining/mineral production and metal production.

⁸ Assumptions or different interpretations could be applied only when a conservative approach is taken and transparently reported.

Tip: Best practices to ensure transparency

Despite a DNA's efforts to meet the quality objectives, the uncertainty of data quality may remain significant. Through effective public engagement, data quality reviewers as well as DNAs can make decisions with reasonable confidence. By establishing collaborative processes with stakeholders, DNAs can improve the overall quality of data. DNAs may include stakeholders and sector experts from the planning stage for standardized baselines in order to better understand their perspectives and interests and/or encourage their participation in data collection. If applicable, DNAs may promote the involvement of stakeholders/experts during the overall QA/QC processes and regularly take into account their feedback on the data quality. DNAs should call for public consultation in a way that facilitates comments from relevant stakeholders (e.g. CDM/sector/quality experts, industry representatives, general public representatives, potential project participants, representatives of relevant ministries, etc.) and allows a reasonable time for comments on the established standardized baselines and relevant procedures.

- (k) Traceability: document all data sources as well as measurement, calculation and estimation methods, which will enable the reproduction or review of the data used for the development of the standardized baselines by a third party. All documentation provisions for the review ~~will be~~ are addressed in section 8VIII.

17. General provisions

16. To meet the above data quality objectives, DNAs should develop a QA/QC system that outlines QA/QC activities, processes, schedule and responsibilities of the personnel involved as well as the institutional arrangement. ~~The goal of this system is to ensure that the above data quality objectives and the provisions given below are met.~~ DNAs may face a challenge in achieving the data quality objectives due to different priorities, perspectives, interests or circumstances. To address such a challenge DNAs may apply a more feasible and simplified approach to meet these objectives.

~~6. Sector-specific data templates: To ensure the consistency of the datasets, DNAs should use sector-specific data templates provided by the secretariat. The data templates include standardized data formats and lists of relevant data, based on sector-specific situations and mitigation measures.~~

~~7. DNAs may request a revision to these data templates or propose a new data template in accordance with the Procedures. If requested, further clarifications on the use of the data templates will be provided by the secretariat.~~

~~8. Data vintage: DNAs should complete the data templates, based on the most recent data available using in accordance with the vintage of data that is sector-specific and predefined by the Board. For a Party with fewer than 10 registered CDM project activities as of 31 December 2010, the Board may establish a different data vintage for certain sectors if necessary.~~

~~9. Update frequency Validity: DNAs should update the datasets used for the establishment of standardized baselines at a frequency that is sector-specific and predefined by the Board. When updating the datasets, DNAs should apply a consistent approach. For example, data delivery protocol, compiling methods, assumptions and relevant calculations should be documented and comparable over time.~~

~~10. Completeness: DNAs should make efforts to include all relevant installations/facilities/companies (target population) that contribute to the production of~~

the sector output in a country or a group of countries, based on the scope of a sector defined in the data templates. Where some data (e.g. for companies) are dropped out due to non-response, DNAs should determine the level of population incompleteness by calculating the ratio of the number of dropped-out companies to the total number of companies and also by estimating the rate of the historical contribution of these dropped-out companies to the total production of the sector output. DNAs should apply a conservative approach to address the issue of this incompleteness (e.g. by assuming that top-performing companies are dropped) and provide justification for the conservative approach.

11. DNAs should provide all required data specified in the data templates. Incomplete data entries will be filled automatically by conservative default values if the values are predetermined in the data templates. Otherwise, DNAs should apply the average values of the top 20% performance data as conservative default values. If other default values are applied, DNAs should demonstrate that those values are conservative.

12. Accuracy: DNAs should implement measures to avoid errors and duplications. Sources of uncertainties should be identified and reduced as far as practicable. However, where a high level of accuracy cannot be achieved, a conservative approach should be taken and a justification should also be provided.

13. When a sampling approach is employed to derive activity data, DNAs should apply statistically sound sampling approaches and comply with the sampling requirement for reliability (95/105 confidence/precision for industrial sectors and 90/10 confidence/precision for the others) and select sample sizes in accordance with the "Standard for sampling and surveys for CDM project activities and programmes of activities".

14. Transparency: Where applicable, DNAs should invite stakeholders to provide comments on the established sector-specific standardized baseline and the process used to establish it. The consultation should be done in an open and transparent manner, in a way that facilitates comments from stakeholders and allows a reasonable time for comments to be submitted. DNAs should prepare a public consultation report that includes a summary of the comments provided by stakeholders and how they considered all the comments received.

6. Quality control

17. As part of the QA/QC system, DNAs should develop QC procedures that address how to ensure the data quality from pre-submission QC activities to the finalization of a QC report. The QC procedures may be revised where necessary.

18. Where data are to be collected, It is very important to facilitate the collection of high-quality data, so DNAs should develop a “data template”⁹ along with a ~~should conduct a pre-submission quality check by establishing a~~ “data delivery protocol”¹⁰ for data providers that describes specific rules and procedures for the collection and delivery of the requested data in order to ensure pre-submission data quality. DNAs ~~could may~~ include the following components in the data delivery protocol:
- (a) Purpose of data collection: DNAs could set multiple purposes in addition to the establishment of standardized baselines for a certain sector;
 - (b) Data types: DNAs should specify the lists, units and formats of data requested, based on ~~the sector-specific pre-established data templates which will help to achieve a high level of consistency and relevance~~. Clear definitions or descriptions of the data ~~provisions~~ related to accuracy and/or conservativeness should be provided in an objective manner;
 - (c) Data acquisition procedures: DNAs should provide clear guidance on data compilation and aggregation, for example how specific unit-level (equipment/facility) data should be consistently aggregated into entity-level (company) data. The procedures should ~~indicate highlight~~ that data should be collected from valid data sources and only the current data should be compiled;
 - (d) Traceability: All the data acquisition procedures should be documented which essentially include data sources, references and, ~~if possible~~, the persons responsible for different functions. ~~If feasible~~, it is recommended that data providers ~~prepare a summary report explaining~~ how the data was collected and how the quality of the data was ensured. ~~The report should include a declaration of conformity (e.g. data providers declare that they provide the data in accordance with the data delivery protocol).~~ This explanation may be provided in a summary report which ~~should may~~ include ~~all~~ identified issues related to data

⁹ According to the “Procedure for development, revision, clarification and update of standardized baselines”, DNAs (proponents of SBs) shall use a data template published by the secretariat or propose a new data template for its publication. The published data templates will help DNAs meet the data quality objectives (in particular, relevance, completeness, consistency, objectivity, accuracy, currentness and traceability) as well as prevent potential risks that could cause quality deterioration. That is, this provision for data templates may increase transaction costs at an early stage but help identify/address issues at an early stage with support from the secretariat/experts and avoid/reduce additional steps to solve problems at a later stage, which will decrease the overall transaction costs. As presented in appendix 2 (data template sampler), the data template should include the lists of data requested for establishing SBs with clear definitions and consistent formats. The data template can be developed for data collection, processing, compilation and/or verification. In cases where SBs are to be developed by using a methodological approach in the approved CDM methodologies/tools, DNAs may use their own or other available data templates in accordance with the monitoring provisions of the approved methodologies/tools (e.g. grid tool template).

¹⁰ The data delivery protocol is the guidance for data providers, explaining what data are to be collected, why and how, and when data are to be delivered to DNAs. As presented in appendix 1 (data delivery protocol sampler), it can be in a comprehensive format. In cases where data templates are to be filled directly by data providers, the data delivery protocol can be a simple instruction on how to fill the data template. ~~will help data providers better understand the requested data and ensure consistency, credibility, currentness and in particular accuracy and objectivity of the data delivered by the data providers.~~

quality (e.g. intrinsic uncertainty, limited data availability or inconsistent data system). If a conservative approach has been used to address the data quality objectives that could not be achieved, this conservative approach should be described in the summary report;

- (e) Delivery schedule requirements: DNAs may should specify a scheduled time frame and types of the deliveries. DNAs should ensure that all mandatory data are delivered. DNAs could request periodic reports if necessary;
- (f) Confidentiality issues: DNAs should describe how they will address issues related to confidentiality, if relevant;
- (g) Contact: DNAs should be responsive to queries from data providers and provide assistance whenever requested, so the protocol should contain contact information and define the deadline for responses to queries. DNAs should establish lines of effective communication and feedback with data providers to identify specific opportunities to improve the data quality of data.

Tip: Best practices to facilitate data delivery from data providers (primary data source)

It is difficult to collect data when data providers do not have incentives or legislative obligations for data delivery. The DNAs may establish a partnership with data providers directly or indirectly through relevant associations.

The selection of target sectors for the establishment of standardized baselines requires taking into account the interests of stakeholders and the mitigation potential of the sector, which can be efficiently identified through such partnership.

If applicable, data templates may be developed in cooperation with stakeholders and sector experts, which will help in establishing a more practical data delivery protocol. Before initiating data collection from data providers, DNAs should identify whether some of the required data are already collected for different purposes by other organizations. The coordinated data collection based on institutional arrangements can achieve multiple purposes in an efficient manner including promoting the consistency of the data.

The data delivery protocol should be distributed to the entire target population of data providers. DNAs should define the scope of the population in a particular sector. DNAs should ensure that there is no duplication or missing entities from the population. For example, a company included in a sector "A" should not be included in another sector "B" unless the data provided by the company for each sector are not interrelated and double counting does not occur.

19. DNAs should conduct a post-submission quality check by assessing the credibility of the data sources and the accuracy of the data and, based on primary data and secondary data as well as documents submitted by data providers. DNAs should also review the summary reports of data providers and assess whether the data were generated in accordance with the data delivery protocol. Where necessary, DNAs mayshould ensure the credibility or accuracy of the data by arranging for a third party to conduct a check of at the facilities/companies. The following should be checked:

- (a) **Completeness**: whether all relevant facilities and all the relevant activities within the facilities have been taken into account;
- (b) **Currentness**: whether the data are current and meet the data vintage coverage provisions as per the "Standard for data coverage and validity of standardized baselines";

- (c) ~~Whether reliable data systems are established and how the systems are operated and maintained (logs of operation and maintenance of the systems may be required);~~
 - (d) **Traceability**: whether all data sources are documented accurately, and w;
 - (e) Whether the referenced materials are available for traceability;
 - (f) **Credibility**: how the data source quality has been ~~ensured~~assessed (e.g. calibration frequency ~~and~~ the accuracy of measuring instruments);
 - (g) **Consistency**: whether the level of aggregation is appropriate and consistent with the data acquisition procedures specified in the data delivery protocol;
 - (h) ~~Whether the data processing is traceable and appropriate if the data are derived from a different procedure from the procedures described in the data delivery protocol;~~
 - (i) **Accuracy**: whether the data are consistent with official and publicly available statistics ~~or the DNAs' data~~ (cross-checking wherever applicable). Discrepancies do not necessarily imply that the datasets are erroneous. However, such discrepancies should be justified or explained; ~~and whether the sampling requirement is met if the data are derived from sampling including a review of the representativeness of the sample and the appropriateness of the methods used for scale-up.~~
20. When ~~certain problematic data are identified~~the data quality objectives are not met ~~as per paragraph 25 above~~, DNAs should implement the following corrective actions before approval:
- (a) Request data corrections. In order to obtain correct data, DNAs could provide assistance to improve data systems and management practices for data providers;
 - (b) Apply a conservative approach as specified in the QA/QC system;
 - (c) Replace with ~~other reliable secondary data whose credibility has already been confirmed, using conservative approaches. In cases where DNAs utilize activity secondary data collected for other purposes with different systems and approaches, DNAs should confirm that the official national statistics offices or other data compiling agencies have performed reliable QC activities.~~
21. The ~~approved~~ data ~~that meet the quality objectives~~ are the input data for the data template. When compiling the data in the data template, DNAs should identify whether the following risks exist and take appropriate actions to prevent or solve them through internal review:
- (a) Double counting or duplication of records: the data approved per each data provider should be included only once in the data template (one spreadsheet per company);
 - (b) No response ~~and incomplete data~~: ~~DNAs should encourage the active participation of data providers and be responsive to their needs~~for any non-responses ~~or incomplete data~~, a conservative approach should be applied to

replace the missing unavailable data or a justification should be provided confirming that the missing data do not cause an overestimation of the baseline emissions;

- (c) ~~Limited data availability~~ Data of poor quality: the DNA may prefer to apply a conservative approach to handle data which due to poor quality do not meet the quality objectives; data or data compiling systems, some data may not be approved by DNAs. DNAs should apply a conservative approach to unqualified data. DNAs should provide data providers with assistance to improve data systems and management practices for future data collection;
- (d) Incorrect data entry: ~~great care should be taken to avoid~~ typographical errors, erroneous entries (in the wrong column or wrong category) and duplication of entries (within one spreadsheet) should be avoided by regular or automatic data entry checks;
- ~~(e) Incomplete data entry: in order to meet the completeness provision, all mandatory fields in the data templates should be filled in;~~
- (f) ~~Low-quality data~~ Data processing: if further consolidation, calculation or conversion of data is required, DNAs should establish an additional step to double-check the outcomes (e.g. recalculation or reconciliation). ~~Manual processing of data should be avoided and the use of spreadsheets is preferred. The formulae and steps used should be clearly described;~~
- (g) Inconsistency: DNAs should ensure that the datasets in the data template are consistently compiled. Consistent categories, methods, processes and approaches should be applied. ~~When updating, consistency should be checked.~~

Tip: Best practices to reduce risks related to data compiling

An automatic system may be implemented where applicable and cost-effective. When designing data templates, based on predefined algorithms and procedures, an automatic function may be built in as much as possible to avoid typographical errors, omissions or duplications and to enhance consistency and efficiency. The potential for human error leads to a high level of risk, and thus well designed automatic systems reducing manual processing and utilizing predetermined default values may help minimize risks in an effective manner.

- 22. ~~The data templates automatically generate aggregated outcomes for standardized baselines, but DNAs could~~ may cross-check the outcomes (e.g. baseline emission factors or positive lists) of the established standardized baselines against a number of similar international datasets for economic indicators or against primary other relevant data or historical data, if available.
- 23. The level of uncertainties depends on the data availability and knowledge of underlying process and inference methods. A traceable and transparent data system is vital to understand uncertainties, so DNAs should ensure the traceability of the QA/QC system including data sources' conditions (e.g. monitoring equipment or archives). DNAs should may identify key causes of uncertainties, such as a lack of completeness, limited data availability, missing data, misclassifications, non-systematic process of collecting data

~~and misreporting. DNAs should~~ quantify such uncertainties¹¹ and take corrective actions to address them (e.g. by applying a discounting factor).¹²

24. DNAs should ~~document~~ in a QC report ~~document~~ how the QC procedures were implemented and how the data quality objectives ~~and the general provisions~~ were met. The QC report ~~should may~~ specify how all evidence/references to data sources were checked. It ~~may should~~ include justifications on the selected approach for obtaining reliable input data (e.g. measurement, calculation, national statistics, sampling, surveys and/or measurement campaign). Information on the uncertainties associated with activity data and major issues regarding the quality of input data, methods, processing or estimates ~~may should~~ be addressed in this report. If updating, the report ~~may should~~ highlight changes in data inputs or methods, substantial divergences in the datasets and a trend analysis if necessary. The QC report should summarize key findings and present a plan for how to address any identified major issues in the future (e.g. training for data maintenance personnel or automatic data management systems).

7. Quality assurance

25. QA procedures should be based on an approach for assessing the quality of the data management system rather than checking the accuracy of a specific set ~~(or sets)~~ of data. ~~That is, the QA focuses~~ on the system and procedures rather than on the outcomes.
26. As part of the QA/QC system, DNAs should develop QA procedures for the systematic identification, formulation and analysis of risks of not meeting quality objectives/~~provisions~~ for the datasets and for defining and implementing activities that mitigate the identified risks. The QA procedures should include update plans to continuously improve the accuracy of the datasets and the efficiency of the overall data system, ~~if feasible~~. DNAs could develop procedures to secure official organizational approval and support for the datasets, if necessary.
27. As part of the QA system, the DOE contracted by the DNA should check whether the QA/QC system is put in place and assess the QA/QC system against the data quality objectives ~~and general provisions~~ established in this document. It also includes assessing whether the QA/QC system has been implemented as designed.
28. During the review of the QA/QC system by the DOE, it is recommended to include experts in relevant technical fields in the assessment team to check whether the QA/QC system, procedures and approaches as well as the final standardized baselines are reasonable.
29. The elements of the QA/QC system that are to be assessed by the DOE include:
- (a) System availability – identify whether a “standardized” data system (collection, consolidation and maintenance) is currently in place and a procedure for reporting activities conducted as part of the QC system has been developed and implemented;

¹¹ Refer to the 2006 IPCC report for further guidance on uncertainty <http://www.ipcc-nggip.iges.or.jp/public/2006gl/pdf/1_Volume1/V1_3_Ch3_Uncertainties.pdf>.

¹² Such as discounting factors as stated in FCCC/SBSTA/2003/10/Add.2, p. 25.

- (b) Conformity – assess whether the QA/QC system, the procedures and all the approaches to develop the datasets met the data quality objectives ~~and the general provisions specified in paragraphs 12 to 21~~. In particular, DOEs should assess whether a conservative approach has been applied in a consistent manner; ~~whether the data delivery protocol was consistent with the data template if applicable; and whether the transparency was ensured, based on the public consultation report and the QC report~~. DOEs should check whether the QA/QC procedures were: (i) developed in accordance with ~~the provisions set in the QA/QC Guidelines; and (ii) effectively implemented (e.g. met the data quality objectives) within the QA/QC system~~. ~~The data delivery protocol should be assessed to determine whether it was consistent with the sector-specific data template. DOEs should assess whether the transparency provision was met, based on the public consultation report and the QC report;~~
 - (c) Traceability – check whether all data and information relating to the datasets and procedures for standardized baselines were clearly documented;
 - ~~(d) Responsiveness – evaluate whether the data delivery protocol met the provisions of the QA/QC Guidelines and whether communication with data providers was timely and efficient;~~
 - ~~(e) Adaptability – assess whether DNAs addressed the major issues identified and whether the system through its procedures was modified accordingly;~~
 - (f) Security – check whether a security system for data management is in place and has operated effectively. Identify whether any issues related to security occurred;
 - (g) Error tolerance – check whether DNAs planned to minimize errors and established and implemented procedures to identify and correct errors proactively.
30. The results of the QA activities should be documented and included in an assessment report¹³ by the DOE. The DOE should prepare a summary of findings including the key issues identified and provide the overall evaluation. If necessary, DOEs ~~could may~~ provide recommendations to improve the quality of data or the overall data system. The assessment report should be made available to DNAs. DNAs may provide responses to the findings in the assessment report and/or implement appropriate corrective actions that address the findings. The responses as well as the corrective actions implemented should be well documented in the final assessment report, ~~if available~~.

¹³ An assessment report is not required for certain conditions: i) where no intensive data collection/processing is required to establish standardized baselines (e.g. a standardized baseline for landfill methane destruction may require regulation data only on the mandatory destruction level); or ii) where standardized baselines are developed through a top-down approach in accordance with the “Procedure for development, revision, clarification and update of standardized baselines”.

8. Documentation provisions

31. DNAs should document and maintain all data and information relating to the establishment of standardized baselines and submit¹⁴ the datasets compiled in the ~~sector-specific~~ data templates and the following documents in accordance with the “~~Procedure for Submission and Consideration of Standardized Baselines~~ Procedure for development, revision, clarification and update of standardized baselines”:
- ~~(a) QA/QC system – as specified in paragraph;~~
 - (b) Data delivery protocol – as specified in paragraph 18, 23-if applicable;
 - (c) Summary reports – if submitted by the data providers as specified in paragraph 18 23-(d); ~~data providers should prepare these reports. However, if the reports are not submitted, DNAs should produce these reports during approval;~~
 - ~~(d) Raw data and~~ Any supplementary documents, ~~if necessary-submitted by data providers;~~
 - ~~(e) Primary data – wherever used as reference or raw data for the establishment of standardized baselines, the data sources and all the issues related to data quality of the primary data (e.g. the approach of data collection, processing and treatment of the associated uncertainties) should be documented;~~
 - ~~(f) Secondary data – wherever used as reference or raw data for the establishment of standardized baselines, the data sources and all the issues related to data quality of the secondary data should be documented;~~
 - ~~(g) Datasets compiled in the data templates;~~
 - (h) Public consultation report – as specified in paragraph 15 (j) 24;
 - (i) QC report – as specified in paragraph 24-34;
 - (j) Assessment report – as specified in paragraph 30-36 above.
32. ~~Secure archiving of complete datasets is important.~~ DNAs should retain all the data/information for a period of five years after the submission of the standardized baselines. ~~Great care should be taken to ensure confidentiality. DNAs should may ensure data confidentiality and~~ have a secure data maintenance system including code set for access control, strategies for unexpected damage or loss of data and procedures to protect confidential data.

¹⁴ DNAs may submit the required documents in a combined format as long as traceability can be met. For example, the data delivery protocol, summary reports, supplementary documents or public consultation report can be included or combined in a QC report.

Attachment 1. Data delivery protocol (sampler)

Note: this sampler is designed for industrial sectors (e.g. cement sector) when the primary data are to be collected. Samplers in the appendix can be used for reference.

1. This protocol is to promote an effective data delivery from cement industry (data providers) to a DNA by providing general information and specific requirements for data collection and delivery. All principal data providers ~~shall~~ **should** ensure the delivery of the required data and documents in accordance with the ~~requirements—guidance~~ **stated provided** in this protocol.

1. General information

2. A. Purpose of data collection: To establish standardized baselines for CDM projects in a cement sector. [If other purposes are incorporated in this data collection, they should be **addressed mentioned**]. All companies of the cement sector located in country “A” are required to provide relevant data and supporting documents. [If there are any incentives/benefits or ~~penalties—legislative obligations~~ involved, they should be addressed].
3. B. Confidentiality: All data and documents collected will be treated as confidential and the data will be open to the public only in an anonymous aggregate form unless consent to disclose certain information is given. Our security system and procedures are developed for a secure data management and will ensure restricted access to the data. [The data will not be used for other purposes without prior permission from data providers].
4. C. Help Desk: ~~We have established a help desk.~~ Please contact the help desk when further clarification or help is needed. ~~We will respond to Any~~ queries and requests **will be responded** to within ~~[x]40~~ days. [Help desk or contact person information should be stated].

2. Requirements for data collection and delivery

5. A. Data types: The following data are to be provided. The accompanying questionnaires [in the data template] will provide more detailed information.

Data	Description (unit/definition)
Amount of cement produced	(tonnes/year) yearly production per type of cement (OPC, PPC.)
Fuel used	(tonnes/year) yearly consumption per type of fuel (Pek, HFO,..)
Clinker kiln technology	(tonnes/year) yearly production of clinker per type of technology used (VSK, Rotary..)

6. B. Data acquisition/aggregation: Data should represent the total production of *cement* by all facilities operated in your company **through aggregation (sum-up of production by**

each facility). Actual measured data should be collected for the last three calendar years before the submission of standardized baselines [2009, 2010 & 2011] [and the average values of the three years should be taken as the best representative data]. When historical records are not available for this period, estimated yearly data from other reliable data sources could be included, which should be accompanied with justification. (why such estimates are selected and how to ensure that they would not lead to an overestimation of the baseline emissions).. All data should be compiled and aggregated consistently. That is, same data units, scopes, definitions and calculations should be applied consistently to all unit-level (equipment/facility) data as well as to the entity-level (company) data. All data should be collected from valid data sources and great care should be taken to avoid typographical errors, duplication of entries and wrong calculations/conversions.

7. C. Traceability: The data acquisition procedures should all be documented including data sources, references and responsible contact persons.
8. A summary report should be produced explaining how the data was collected and how the quality of the data was ensured. This summary report should include all identified issues related to data quality (e.g. intrinsic uncertainty, limited data availability or inconsistent data system).
9. D. Delivery requirements: Please Ensure that all mandatory parts of the questionnaires [in the data template] are completed. Incomplete questionnaires in any of the main parts are likely to be unusable and your efforts will be wasted. The complete dataset, the summary report and a declaration of conformity with the signature of your representative should be delivered by email by [dd/mm/yyyy] 30 June 2012 [within a scheduled time frame].

Date to provide this Protocol

Signature of DNA

Attachment 2. Data template (sampler)

Note: this sampler is designed for industrial sectors (e.g. cement sector). In accordance with the “Procedure for development, revision, clarification and update of standardized baselines”, designated national authorities (DNAs) (proponents of standardized baselines) shall propose a new data template for its publication unless relevant approved data templates are available on the CDM website or available as a part of the monitoring section of an approved baseline and monitoring methodology or methodological tool.

1. **Definition of sector (output):** This data template is to be used by DNAs in establishing sector-specific standardized baselines for the cement sector. The template is applicable only to the sector defined in each data template.
2. **Scope of standardized baselines:** The template is developed based on integrated measures. In order to take into account sector-specific situations, data that are directly or indirectly related to emissions/mitigation will be collected from the cement sector located in country “A”.
3. **Structure of data template:**
 - (a) Introduction: Key information for the data template is clearly described including detailed formulas, approaches and assumptions used to establish standardized baselines. Definitions of key terminologies and instructions on how to fill the data template may be included (e.g. data coverage and validity);
 - (b) Input data: Lists of relevant data categories with consistent format and clear description are provided;
 - (c) Reference: Lists of predetermined defaults and key information/data are already (or to be) used instead of input data or for calculation/processing of input data are provided.
4. **Examples of data categories:** The required data may be listed per data source (primary vs. secondary).
 - (a) Primary data collection: general information (name, capacity, installation/updated date or location of facility) and performance data (technologies, fuel, feedstock and/or production);
 - (b) Secondary data collection: design performance data of technology (from manufacturers/suppliers or research institutions); market trends (from sectoral experts or statistics documents); or policies/regulations (from public documents or government authorities).

Attachment 3. Quality control (QC) report (sampler)

Sector	
Name of DNA	
Primary Person Responsible for QC Procedures	
Contact of the Primary Person Responsible	
Implementation Dates of QC Procedures	
Please describe how your QC procedures were implemented	
<i>Pre-submission QC: data delivery protocol (as attached above)</i>	
<i>Post-submission QC: ...</i>	
<i>Corrective actions:...</i>	
Please specify how the credibility of the data sources was checked.	
Please specify how the accuracy of the data was checked.	
Please specify how the consistency was achieved in particular where multiple secondary data sources were used. and how the data vintage provision was met.	
Please specify how the "Standard for data coverage and validity of standardized baselines" was complied with.	
Please specify how the completeness was achieved.	
Please specify how the transparency was achieved.	
Please specify major issues and uncertainties identified during the QC procedures.	
Please specify major corrective actions taken during the QC procedures.	
Please justify the conservativeness of the approaches taken during the QC procedures.	
Please summarize key findings and present a plan to improve the data quality in the future.	

Date to finalize this report

Signature of DNA

Attachment 4. Assessment report (sampler) Checklists for DNAs

Name of DNA			
Name of rReviewer			
Contact details of rReviewer			
Dates rReviewed			
Criterion	Definition	Yes/No	Explanation
System aAvailability and qQuality	Is a 'standardized' data system currently in place? Was QC system (resource/procedure) implemented (including review of its responsiveness, adaptability and error tolerance)?	yes	
	Has a procedure for reporting activities conducted as part of the QC system been developed and implemented? Was a security system for data management in place and has it operated effectively?		
Conformity	Did the QA/QC system/procedures meet the data quality objectives of the QA/QC Guidelines?		
	Did the QA/QC system/procedures meet the general provisions of the QA/QC Guidelines?		
	Were the approaches taken by DNAs conservative? Were the conservative approaches applied consistently?		
Traceability	Were all data and information relating to the datasets and procedures clearly documented?		
	Were QC activities the QC report clearly documented in the QC report accordance with the QA/QC Guidelines?		
	Were all required documents, information and data available for assessment?		
Responsiveness	Did the data delivery protocol meet the provisions of the QA/QC Guidelines?		
	Was the communication of the DNA with data providers timely and efficient?		
Adaptability	Was the system through its procedures modified in order to address the major issues identified?		
	Did the modified system meet the data quality objectives and the provisions of the QA/QC Guidelines?		
Security	Is a security system for data management in place and has it operated effectively?		

	Have any issues related to security occurred?
Error Tolerance	Were there established procedures to minimize errors proactively? Were these procedures implemented effectively?
Summary of Findings assessment including assessment process and findings/resolution	
Responses and Corrective Actions of DNA	
Date of Submission Date to DNA	

Signature by of reviewer _____Signature by of DNA _____**1. Checklist for planning**

QA/QC system is available (procedures, plan and responsibilities of personnel are specified)	<input type="checkbox"/>
Data management system is operated and maintained in a secure and effective manner	<input type="checkbox"/>
Definition and scope of the "sector" are clearly determined	<input type="checkbox"/>
Partnership with the sector is established	<input type="checkbox"/>
Potential data sources are identified	<input type="checkbox"/>
Institutional arrangement with other organizations is established for data collection/management or for developing standardized baselines	<input type="checkbox"/>
Sources of financial, technical or capacity-building support (e.g. through UNFCCC or international organizations) are identified and utilized	<input type="checkbox"/>
Data quality objectives in the QA/QC guidelines and requirements in all other regulatory documents for standardized baselines are understood	<input type="checkbox"/>
Data template is available (if not, develop a new one)	<input type="checkbox"/>

2. Checklist for pre-submission QC activities

Data delivery protocol is developed in consistency with the data template	<input type="checkbox"/>
Key sources of uncertainties related to data sources and data collection are identified and solutions/procedures to address them are established	<input type="checkbox"/>

Key issues related to secondary data are identified and solutions/procedures to address them are established	<input type="checkbox"/>
Partnerships with the sector are fully utilized (in particular, to communicate with the target population)	<input type="checkbox"/>
Incentives or legislative obligations are highlighted to promote data collection	<input type="checkbox"/>
Most feasible method is selected for data collection, taking into account effectiveness and efficiency	<input type="checkbox"/>

3. Checklist for post-submission QC activities

Level of population incompleteness is calculated, based on response rate and data availability	<input type="checkbox"/>
All required data and information are available to review the quality of data collected	<input type="checkbox"/>
By cross-checking, the quality of data collected is evaluated and validated	<input type="checkbox"/>
Solutions/procedures are applied to address the identified issues as predetermined	<input type="checkbox"/>
Through flexible and/or conservative approaches, newly identified issues are addressed in an effective and efficient manner	<input type="checkbox"/>

4. Checklist for compiling and outcomes

Manual processing is minimized	<input type="checkbox"/>
Key sources of uncertainties related to data compiling are identified and solutions/procedures to address them are established	<input type="checkbox"/>
Solutions/procedures are applied to address the identified issues as predetermined	<input type="checkbox"/>
Through flexible and/or conservative approaches, newly identified issues are addressed in an effective and efficient manner	<input type="checkbox"/>
Established standardized baselines are reviewed through transparent consultation processes	<input type="checkbox"/>
If necessary, further modifications (e.g. disaggregation level of output) and additional steps (e.g. additionality demonstration) are applied to establish standardized baselines or positive lists	<input type="checkbox"/>

5. Checklist for QC report

All QC activities are documented	<input type="checkbox"/>
Key identified issues are summarized	<input type="checkbox"/>
The conservativeness of the applied solutions/procedures is justified	<input type="checkbox"/>
An improvement plan to reduce uncertainties and address other key issues is developed	<input type="checkbox"/>

- - - - -

Document information

<i>Version</i>	<i>Date</i>	<i>Description</i>
02.0	14 May 2014	Published within annex 10 to the annotated agenda of EB 79.
01.0	2 March 2012	EB 66, Annex 49 Initial publication.
Decision Class: Regulatory Document Type: Guideline Business Function: Methodology Keywords: quality management, standardized baselines		

DRAFT