

QUALITY ASSESSMENT FORM

Title of the evaluation: Evaluation of the Forest Law Enforcement, Governance and Trade (FLEGT) Action Plan

DG/Unit: DEVCO C2

Evaluator/contractor: TEREA

Assessment carried out by: Evaluation Steering Committee

Date of the Quality Assessment: 16.04.2016

(1) RELEVANCE

Does the evaluation respond to information needs, in particular as expressed in the terms of references?

SCORING	Poor	Satisfactory	Good	Very Good	Excellent
		X			

Arguments for scoring:

The evaluation report responds to all the evaluation questions foreseen in the Terms of Reference and further specified during the inception phase. It covers all Action Areas of the FLEGT Action Plan.

(2) APPROPRIATE DESIGN

Is the design of the evaluation adequate for obtaining the results needed to answer the evaluation questions?

SCORING	Poor	Satisfactory	Good	Very Good	Excellent
		X			

Arguments for scoring:

The design of the methodology was adequate, building, among others, on literature and trade data review, surveys of EU Member States and all Voluntary Partnership Agreement (VPA) countries, an open public consultation, and visits to VPA and other key timber producer / consumer countries. Challenges faced with data collection are mentioned in the report. Balanced attention was given to the perspectives of the different groups of stakeholders. Specific attention was given to emergent outcomes, both positive and negative, and the underlying causal pathways. The methodology is properly described in the report.

At the same time, the overall assignment could have been better managed. Excessive resources have been allocated to data collection and initial consultations, leaving insufficient time for the analytical part of the evaluation and further consultations on preliminary findings.

(3) RELIABLE DATA

Are data collected adequate for their intended use and have their reliability been ascertained?

SCORING	Poor	Satisfactory	Good	Very Good	Excellent
		X			

Arguments for scoring:

The evaluation properly built on data and monitoring systems available. Some of the difficulties faced to collect data, in particular on levels of illegal logging –for which there is no official figure- are acknowledged.

There is a lack of data available on some of the FLEGT achievements (e.g. progress on governance). This led the evaluation to rely more –for various aspects of the Action Plan- on stakeholder views and inputs than on hard evidence.

Data collection on financial allocations specifically is assessed as weak. While the report acknowledges some difficulties, more relevant data (e.g. on breakdown of FLEGT-related Official Development Assistance) could have been collected.

Data on the individual actions of the different groups of FLEGT actors is also assessed as insufficient. The questionnaires used for the surveys could have been better tailored to capture the actions and views of the different groups of actors consulted.

(4) SOUND ANALYSIS

Are data systematically analysed to answer evaluation questions and cover other information needs in a valid manner?

SCORING	Poor	Satisfactory	Good	Very Good	Excellent
		X			

Arguments for scoring:

The analysis builds on the inputs and data collected, whenever such data is available. The report makes proper links between findings presented and the inputs collected (e.g. reference to the survey results to substantiate claims).

The analysis related to the FLEGT Voluntary Partnership Agreements is assessed as weaker, possibly as a result of the different views from stakeholders on the priority objectives that these VPAs aim at.

(5) CREDIBLE FINDINGS

Do findings follow logically from and are justified by, the data/information analysis and interpretations based on pre-established criteria and rational?

SCORING	Poor	Satisfactory	Good	Very Good	Excellent
		X			

Arguments for scoring:

The overall findings of the evaluation are considered fair and balanced, reflecting an acceptable compromise of the perceptions of stakeholders.

Details provided in the analysis are sometimes weak however. The report includes a lot of generalisations and extrapolations which sometimes lack nuance. On VPAs for example, better consideration of political contexts could have improved the quality of the findings.

(6) VALID CONCLUSIONS

Are conclusions non-biased and fully based on findings?

SCORING	Poor	Satisfactory	Good	Very Good	Excellent
		X			

Arguments for scoring:

The evaluation conclusions are clearly presented in the report, and cover each key question foreseen. They reflect the findings detailed in the report.

(7) HELPFUL RECOMENDATIONS

Are areas needing improvements identified in coherence with the conclusions? Are the suggested options realistic and impartial?

SCORING	Poor	Satisfactory	Good	Very Good	Excellent
		X			

Arguments for scoring:

The recommendations are broadly coherent with the evaluation conclusions and provide useful indications of possible directions to follow to improve the Action Plan. However the practical implications of these recommendations are not sufficiently covered, leading to some recommendations being probably unrealistic.

(8) CLARITY

Is the report well structured, balanced and written in an understandable manner?

SCORING	Poor	Satisfactory	Good	Very Good	Excellent
		X			

Arguments for scoring:

The report clarity has initially been problematic, with various sections overlapping. The writing quality of the first draft has also been assessed as poor. These aspects have been improved in the final version of the report, where the key messages are summarised and highlighted. The length of the report has remained excessive. The report also lacks a proper focus of most relevant issues

OVERALL ASSESSMENT

OF THE FINAL EVALUATION REPORT

Is the overall quality of the report adequate, in particular:

- Does the evaluation fulfil contractual conditions?

Yes

- Are the findings and conclusions of the report reliable, and are there any specific limitations to their validity and completeness?

The findings appear broadly reliable. They will serve as basis for further dialogue with key stakeholders of the Action Plan, which will provide opportunities to validate these findings and further define the recommendations

- Is the information in the report potentially useful for designing intervention, setting priorities, allocating resources or improving interventions?

The report provides an assessment of the past 11 years of the FLEGT Action Plan implementation. It documents findings and views which had not been sufficiently documented in the past. This provides a useful basis for further reflection on the Action Plan. The recommendations remain very general. They give indications of possible directions to pursue, but their details and feasibility will require further consultation and assessments.

GUIDE ON SCORING THE CRITERIA

This list of indicators aims at helping to score each criterion and it can also assist in the process of developing the argumentation underpinning the score.

The indicators may be adapted according to the specificities of each evaluation and some indicators may be omitted and others added when appropriate.

The indicators are, roughly speaking, presented in order of importance (i.e., those at the start of the list are crucial even for a moderate score while the concurrent accomplishment of those at the end of the list may suggest a higher score).

1. - RELEVANCE

Does the evaluation respond to information needs, in particular as expressed in the terms of references?

This criterion concerns how well the evaluation responds to the terms of references.

- The evaluation report deals with and responds to the evaluation questions
- A justification was provided if any evaluation questions wasn't answered
- The scope covers the requested periods of time, geographical areas, target groups, parts of budget, regulations, etc
- Limitations in scope are discussed and justified
- Effects on other policies, programs, groups, areas etc are considered
- Unintended effects are identified
- The evolution of the intervention is taken into account possible changes in the problems and needs compared to the situation at the start of the intervention have been addressed
- The evaluation broaden the scope or enlighten the approaches in the policy cycle
- The evaluation add value to existing policy knowledge
- Other

2. - APPROPRIATE DESIGN

Is the design of the evaluation adequate for obtaining the results needed to answer the evaluation questions?

This criterion concerns the inception phase. It operationalises and possibly complements the terms of reference. In some cases, because of unforeseen events, it may also relate to a subsequent reorientation of parts of the evaluation work.

- The rationale of the program, cause-effect relations, outcomes, policy context, stakeholder interests, etc have been studied and taken into account in design the evaluation
- The evaluation method chosen is coherent with evaluation needs and requests
- The method is clearly and adequately described, in enough detail for the quality to be judged. It is described to the extent that the evaluation can be replicated
- Information sources and analysis tools are adequate for answering the evaluation questions
- Judgement criteria to help answer the evaluation question were pre-defined
- Weaknesses of the selected method are pointed out along with potential risks
- Other methodological alternatives are considered; their pros and cons are explained
- Research design has been validated with experts or relevant stakeholders if appropriate (e.g. experts on related policies, specific evaluation know-how)
- Ethical issues are properly considered (confidentiality of sources of information, potential harms or difficulties of participation of stakeholders, etc)
- Other

3. - RELIABLE DATA

Are data collected adequate for their intended use and have their reliability been ascertained?

This criterion concerns the relevance and correctness of both primary and secondary data.

- Available information and sources are well identified
- Relevant literature and previous studies have been sufficiently reviewed

- Existing monitoring systems were used
- Data and information are free from factual or logic errors; data gathered are correct and sufficient
- Data collection rationale is explained; and it is coherent with the design of the study
- The quality of existing or collected data was checked and ascertained
- The amount of qualitative information and quantitative data is balanced and appropriate for a valid and reliable analysis
- Tools and means used to collect and process data (e.g. surveys, case studies, expert groups, etc...) were: selected in relation to criteria specified in the inception phase; complete and suitable for answering the evaluative questions; adequately used as to guarantee the reliability and validity of results
- Tools and data collection limitations (missing coverage, non-participation or non-attendance of selected cases) are discussed and explained.
- Correcting measures have been taken to avoid any potential bias and/or their implications
- Other

4. - SOUND ANALYSIS

Are data systematically analysed to answer evaluation questions and cover other information needs in a valid manner?

This criterion refers to the correct interpretation of data and to the adequacy of the method applied.

- There is a clear, solid and coherent deductive analysis (e.g. controlled comparison, experimental research, inferential statistics, etc...)
- The analysis is well focussed on the most relevant cause/effect relations and influences underlying the program logic, and alternative explanations have been considered
- The analysis uses appropriate quantitative or qualitative techniques, suitable to the evaluation context
- Cross checking of findings has taken place. The analysis relies on two or more independent lines of evidence

- Explanatory arguments are explicitly (or implicitly) presented
- The context (historical, socio-economic, etc...) is well taken into account in the analysis
- The report reflects an appropriate range of stakeholders consulted
- Inputs from important stakeholders are used in a balance way
- The limitations of the analysis and exceptions to general explanations or evidences were identified, discussed and transparently presented
- Other

5. - CREDIBLE FINDINGS

Do findings follow logically from and are justified by, the data/information analysis and interpretations based on pre-established criteria and rational?

This criterion concerns the coherence of the findings with the preceding analysis and data.

- Judgements are based on transparent criteria
- Findings are supported by evidence originating from sound analysis
- Generalisations or extrapolations, when made, are justified (e.g., through the sampling or selection of cases)
- Findings corroborate existing knowledge; differences or contradictions with existing know-how are explained
- Stakeholder opinions were considered and reflected when appropriate
- Main findings are replicable
- Limitations on validity are pointed out; trade-offs between internal and external validity are identified and discussed
- Results of the analysis reflect an acceptable compromise of the perceptions of stakeholders and those described by figures and facts observed and estimated
- Other

6. - VALID CONCLUSIONS

Are conclusions non-bias and fully based on findings?

This criterion concerns the extent to which conclusions logically stem from findings and are based on impartial judgement.

- Conclusions are properly addressed to the evaluation questions and other information needs
- Conclusions are coherently and logically substantiated by evaluation findings
- There are no relevant conclusions missing according to the evidences presented
- Conclusions are interpreted in relation to the policy context
- Conclusions are free of personal or partisan considerations; potential influence of values and interests of the evaluation team in the research method and outcome are openly discussed
- Conclusions are orderly presented and related (categorised, ranked, priorities, sequence)
- Controversial issues are presented in a fair and balanced manner
- Other

7. - HELPFUL RECOMMENDATIONS

Are areas needing improvements identified in coherence with the conclusions? Are the suggested options realistic and impartial?

This criterion concerns the soundness and realism of the recommendations

- Recommendations stem logically from conclusions
- Plausible options for improvements are identified
- Recommendations covers all relevant main conclusions
- They are realistic, impartial, and potentially useful
- Relations among recommendations are taken into account (e.g. priority ranking, sequencing, etc)
- Recommendations provide certain guidance for action planning

- Where feasible, cost of recommendations were estimated
- Other

8. - CLARITY

Is the report well structured, balanced, and written in an understandable manner?

This criterion concerns to the clarity of the presentation and the appropriateness of the content of the evaluation.

- The content of the report describes the policy being evaluated, its context, the evaluation purposes, contextual limitations, methodology, findings, etc in a neat and well structured manner
- The report is well structured and signposted to guide and facilitate reading
- Key messages are summarised and highlighted
- There is a clear presentational linked sequence among data, interpretation and conclusions
- The report includes a relevant and concise executive summary, which includes main conclusions and recommendations in a balance and impartial manner
- Specialised concepts were used only when necessary and were they clearly defined
- Tables, graphs, and similar presentational tools are used to facilitate understanding; they are well commented with narrative text
- the length of the report (excluded appendices) is proportionate (good balance of descriptive and analytical information)
- Detailed information and technical analysis are left for the appendix; information overload is avoided in the report
- The report provides a proper focus of truly relevant issues
- Written style and presentation is adapted for the various relevant target readers; the evaluator show awareness of potentially different needs and interests
- Other

OVERALL ASSESSMENT OF THE REPORT

The overall assessment of the evaluation report is not a self-standing criterion. Instead it summarises key elements and consequences of the eight preceding criteria. Moreover, the overall assessment needs to consider the concerns of the potential users of each specific evaluation:

- Does the evaluation fulfil contractual conditions? (*certain internal users*);
- Are the findings and conclusions reliable, and are there any specific limitations to their validity and completeness? (*most internal and external users*);
- Notwithstanding intrinsic weaknesses, is the information in the report -or parts of it- a useful input for designing intervention, setting priorities, allocating resources or improving interventions? (*certain internal users*).