

**RESULTS OF THE UNECE 2018 PILOT STUDY OF DATA FLOWS
FROM COUNTRIES TO CUSTODIAN AGENCIES
RESPONSIBLE FOR SDG INDICATORS**

PREPARED BY THE TASK TEAM ON DATA FLOWS
OF THE
UNECE STEERING GROUP ON STATISTICS FOR SUSTAINABLE DEVELOPMENT GOALS

The report presents the results of the second UNECE pilot study of data flows from countries to custodian agencies responsible for SDG indicators.

The aim of the pilot studies was to explore methods to facilitate understanding and agreement between national statistical offices and custodian agencies regarding harmonized, global statistics for SDGs. The pilots were undertaken at the request of the 2017 and 2018 UNECE Expert Meetings on Statistics for the Sustainable Development Goals (SDGs). A Task Team carried out the pilots, co-chaired by France and Turkey, consisting of Denmark, Kazakhstan, Russian Federation, Sweden, United Kingdom and United States.

The first pilot of data flows was undertaken in summer 2017. Five countries and six custodian agencies participated in the pilot. Data flows for eleven SDG indicators were analysed. The findings from the first pilot contributed to the guidance prepared by the Interagency and Expert Group on SDG indicators (IAEG-SDGs) to improve the transparency and efficiency of global data flows for SDGs.

The second pilot was carried out in summer 2018. It focused mainly on the difficulty of validating data not produced by the national statistical system, such as data produced by custodian agencies through modelling, geospatial information, or other data sources generally outside of the purview of national official statistics. Thirty-seven countries and five custodian agencies participated in the pilot, analysing data flows for ten SDG indicators.

The two pilot studies resulted in concrete recommendations to improve the data flows (see Table I). The outcome of the second pilot study contributed to the principles and best practices of data flows prepared by IAEG-SDGs and the Committee for Coordination of Statistical Activities (CCSA) that was endorsed by the 50th UN Statistical Commission. The outcome also contributed to greater clarity regarding data transmission and validation on a number of indicators and to improving the coordination of data provision for the global SDG indicators. Following the two pilot studies, a number of countries in Africa and in ESCAP region carried out similar exercises.

RESULTS OF THE UNECE 2018 PILOT STUDY OF DATA FLOWS FROM COUNTRIES TO CUSTODIAN AGENCIES RESPONSIBLE FOR SDG INDICATORS

EXECUTIVE SUMMARY

1. The scope and ambition of the 2030 Agenda for Sustainable Development calls for national statistical offices and international organizations to strengthen their coordination in producing harmonized, global statistics to monitor progress in the Sustainable Development Goals. The 49th session of the UN Statistical Commission (UNSC) requested IAEG-SDG and the Coordinating Committee of Statistical Activities (CCSA) to prepare implementation guidance for SDG data flows.

2. Recalling the request of UNSC, the interest expressed at the UNECE Expert Meeting on Statistics for SDGs in 2018, and in consultation with IAEG-SDG and the UN Statistical Division, the UNECE Task Team on Data Flows¹ launched a pilot study to examine country and agency practical experiences with SDG data flows, with a particular focus on the difficulty of validating data not produced by the national statistical office (NSO)/national statistical system (NSS).

3. The study was designed to explore methods to facilitate understanding and agreement between NSOs and custodian agencies regarding harmonized, global statistics produced by custodian agencies through modelling, geospatial information, or other data sources generally outside of the purview of NSOs. It also examined plans for automating data flows from reporting platforms.

4. Thirty-seven countries responded during the summer 2018. The high number of countries who participated in this voluntary pilot shows the importance that they attach to the issue of data flows and validation of data that is published for their country in international databases. Custodian agencies for selected indicators were also invited to participate. The main findings of the pilot and possible solutions are summarized in the table I below and in the concluding chapter. The main points include:

a) Some country focal points are still not known by agencies. Agency contacts are sometimes not known to countries (e.g. for 9.1.2, the precise agency focal point contact information is lacking). Countries sometimes do not know the previously established focal points with well-established data flows for some SDGs indicators, and identified new ones.

b) Some metadata are incomplete and misclassified (17.3.1), difficult to understand or have open questions (6.4.2, 15.4.2). Countries sometimes disagree with metadata which impose a data source (15.4.2) or a method of estimation done by agencies (3.9.1). Metadata for Tier I indicators are not currently within the purview of IAEG-SDGs. There is no process to initiate a review or update of metadata in case issues are identified with it.

c) Sometimes national focal points are not asked to validate data associated with their country published in the UN global data base and they disagree with the data (e.g. 9.1.2 on road transport). Data validation is sometimes complex and requires sufficient time for

¹ Organized under the UNECE Steering Group on Statistics for SDGs.

countries and agencies. Countries require clarity about who is responsible for the data being published in the global database.

d) Data or related data are already collected through an existing reporting mechanism (e.g. 6.4.1, 6.4.2) and countries would like to avoid duplicate reporting even if an adaptation of the questionnaire is required.

5. Overall, we found broad consensus in the issues identified and the solutions proposed. Many appear to require a low level of effort. However, it is clear that all actors in SDG data flows have a role to play in strengthening the transparency and efficiency of data flows, and improving the process requires close collaboration.

6. Some of the possible solutions suggested by participants in the pilot (table 1) were taken into account in the IAEG-SDGs and CCSA guidelines². The discussion with custodian agencies allowed to clarify many questions and to improve the process and data quality. Metadata were improved for some indicators (9.1.2, 15.4.2 for instance) or are going to be improved (17.3.1). Other regions were carrying out similar exercises.

² *Criteria for the implementation of the guidelines on data flows and global data reporting for SDGs* (annex I of the IAEG-SDGs report for the 50th Statistical Commission (<https://unstats.un.org/unsd/statcom/50th-session/documents/2019-2-IAEG-SDG-E.pdf>))

Table I. Possible solutions suggested by participants in the pilot

Proposed lead	Possible solutions
NSOs (with UNSD)	1. National country focal point information should be added to the UNSD global database as a resource for agency focal points (at the country's discretion) and regularly updated.
	2. National contacts provided by NSOs should be coordinated at the national levels and consistent with the mandate received by the custodian agencies from country governments. ³ Such coordination is also essential for reporting non-statistical indicators.
	3. The list of national country focal points should include any other relevant national contacts for extant data flows related to SDG indicators in the UNSD global database (at the country discretion).
	4. NSOs should inform agencies how they want to receive data requests from agencies (i.e., to the focal point, or with the focal point in copy)
	5. Countries should be able to ask IAEG-SDGs to review Tier I or Tier II indicators and metadata when they note an issue
IAEG-SDGs (with UNSD)	1. IAEG-SDGs does not have a governance mechanism or mandate to review metadata for Tier I and II indicators. Countries felt this is a clear need. ⁴
	2. Some indicators have more than one metadata file. This can be confusing. The files should be integrated by custodian agencies. ⁴
	3. Major changes in metadata files should be reviewed by IAEG-SDGs. IAEG-SDGs should receive feedback from countries on metadata and follow up with custodian agencies as needed.
	4. It would be helpful to denote in the UNSD global database and/or official indicator list which global indicators are non-statistical in nature, and those that apply to a subset of countries only (for example, LDCs, SIDS, LLCs, DAC members) and those which rely on a common model.
	5. In cases where country validation is difficult, a note should be added indicating that the data are calculated/modelled by the custodian agency and the status of validation by the country (validated by country/country can't validate/not validated by country/pending review). Thus, it would be easier to publish globally harmonized national data and statistics in the UNSD database with the acceptance of the country. This was recommended as update to the 2018 UNSC data flow guidance.
Custodian Agencies (with UNSD)	1. The data sources used to calculate harmonized statistics should be referenced in the UNSD global database, including data taken from other databases. (Unanimous opinion of respondents.)
	2. The UNSD calendar should indicate agencies' data collection and validation time frames. When exact dates are unknown, tentative dates should be used so NSO work programs can plan ahead.
	3. Agency contact information should be updated/completed for indicators.
	4. The following indicators seem to have missing metadata or require clarification, based on the data flow pilot include: 9.1.2, 15.4.2 and 17.3.1. ⁵
	5. Work plans for some Tier III indicators are outdated and need to be updated.
UNSD (with countries and agencies)	1. A dashboard is needed to support communication between countries and agencies. This dashboard should integrate contact information for country and agency focal points and data collection schedules. The information for the dashboard should be provided by both countries and agencies, should be up to date, and easily accessible/shared.
	2. Any updates made in the metadata files should allow notification and tracking of these updates by national focal points. This will help NSOs respond to agency requests in an efficient manner.
	3. The validation status should be displayed in the UNSD global database, and updates (such as new postings by custodian agencies or revisions) should be communicated to national and agency focal points. For example, such changes could generate an automatic email to focal points. (Unanimous opinion of respondents)

³ This coordination is especially necessary for mandatory data flows, such as DAC flows, where the established national contact is nominated by the government's foreign ministry or UN national permanent mission).

⁴ Countries observed, through the course of their other activities, that multiple metadata files were posted for the following indicators (not the focus of the data flow pilot): 1.1.1, 1.3.1, 2.2.2, 5.5.1, 6.6.1.

⁵ Countries also observed that metadata were missing/needed clarification for the following SDG indicators (not the focus of the data flow pilot): 1.2.2, 1.a.2, 3.3.4, 4.4.1, 3.5.2, 8.9.1, 10.c.2, and 11.1.1.

Table of Contents

EXECUTIVE SUMMARY	II
I. PURPOSE, APPROACH, AND LIMITATIONS OF THE STUDY	1
A. Background	1
B. Objectives.....	1
C. Methods.....	1
D. Findings.....	3
E. Limitations of the study.....	3
F. Acknowledgements	4
II. NATIONAL COORDINATION OF SDG DATA FLOWS.....	5
III. ANALYSIS OF DATA FLOWS BY INDICATOR	6
A. Indicators modelled by agency (3.9.1, 11.6.2)	6
1. Description of the indicator(s) and rationale for study.....	6
2. Country response.....	6
3. Agency response	9
4. Findings.....	9
B. National data produced by agencies using geospatial information (15.4.2).....	9
1. Description of the indicator and rationale for study	10
2. Country response.....	11
3. Agency response	13
4. Findings.....	14
C. Non-statistical indicator based on a survey whose responses may be adjusted by the agency (5.1.1).....	15
1. Description of the indicator and rationale for study	15
2. Country response.....	15
3. Agency response	16
4. Findings.....	16
D. Indicators produced by agencies that use data transmitted by well-established process (17.2.1, 17.3.1)	17
1. Description of the indicator and rationale for study	17
2. Country response.....	18
3. Agency response	18
4. Findings.....	19
E. National data are transmitted already to an intermediary international database (6.4.1, 6.4.2).....	19
1. Description of the indicator and rationale for study	19
2. Country response.....	19
3. Agency response	20
4. Findings.....	21
F. Indicators calculated by the agency without a validation of the nationally nominated SDG focal point (3.6.1, 9.1.2).....	21
1. Description of the indicator and rationale for study	21
2. Country response.....	22
3. Agency response	24
4. Findings.....	25
5. Post survey agencies' development.....	26
IV. SELF-ASSESSMENT BY COUNTRIES AND CUSTODIAN AGENCIES.....	27
A. General description of approach.....	27
B. Suggestions from countries	27
C. Suggestions from agencies	29
D. Findings.....	29
ANNEXES	32
Annex 1. Country and agency surveys for the second data flow pilot.....	32
Annex 2. Country and agency participants to pilot flow survey	33
Annex 3. References	35
Annex 4. Initial findings of the second data flow pilot for IAEG-SDG and CCSA	36

I. PURPOSE, APPROACH, AND LIMITATIONS OF THE STUDY

A. BACKGROUND

7. The scope and ambition of the 2030 Agenda for Sustainable Development calls for national statistical offices and international organizations to strengthen their coordination in producing harmonized, global statistics to monitor progress in the Sustainable Development Goals. With the contributions of many, including the UNECE SDG Steering Group's Task Team on Data Flows 2017 Pilot Study, the Interagency Expert Group on SDGs (IAEG-SDGs) and the Committee for the Coordination of Statistical Activities (CCSA) prepared guidance to improve the transparency and efficiency of global statistical production for SDGs. The 49th session of the UN Statistical Commission (UNSC) welcomed this initial guidance and requested IAEG-SDG and CCSA to prepare implementation guidance for discussion at the 50th session of the UNSC. Recalling the request of the UNSC, the interest expressed at the UNECE Expert Meeting on SDG Statistics, and in consultation with IAEG-SDG and UN Statistical Division the Steering Group proposed, at its Second Expert Meeting on Sustainable Development Goals in April 2018, that its Task Team on Data Flows launch a second pilot study to further inform the development of the coordination of global data flows. The pilot was also to aim to inform IAEG-SDG's and CCSA's preparations of the implementation guidance. This study would build upon the success of the first pilot, conducted in 2017, which described country and agency practical experiences with SDG data flows.

B. OBJECTIVES

8. Like the first pilot, the second pilot was designed to examine data flows for global SDG indicators. The aim of the pilot was to propose simplified procedures to limit reporting burden for both countries and agencies while maintaining transparency of process, concordant with the *Fundamental Principles of Official Statistics*.

9. Based on the main concerns expressed during the second Expert meeting, the specific focus of the second data flow pilot was on the difficulty of validating data not produced by the national statistical office (NSO) or within the national statistical system (NSS). It was designed to explore methods to facilitate understanding and agreement between NSOs and the custodian agencies responsible for SDG indicators regarding harmonized, global statistics produced by custodian agencies through modelling, geospatial information, or other data sources generally outside of the purview of NSOs. It also examined plans for automating data flows from reporting platforms.

C. METHODS

10. The examination of experience focused on validation issues represented by a small set of global SDG indicators. The indicators were selected to represent the different cases where data/statistics are not sent directly to agencies by countries, either because they are calculated, modelled or estimated by the agencies; or because they are already transmitted to agencies via another already existing, well-established process. In addition, the choice of indicators represented the different categories of indicators present in the global indicator framework: statistical indicators, non-statistical indicators, indicators produced by NSOs and indicators not produced by NSOs. All indicators except one were classified as Tier I in the global indicator framework. Indicator 5.1.1 is classified as Tier II.

11. Pilot participants included key actors in the SDG data flow process: UN member countries (in this case, mostly UNECE members but also a number of countries from outside the region) and custodian agencies responsible for indicators examined in this pilot study. Country participants were invited from the list of attendees to the UNECE Second Expert Meeting on Statistics for SDGs. Other volunteer countries were also welcomed.

12. Countries participating in the pilot were asked to describe how they validate data for the indicators selected for this pilot (that is, when their data are already provided to agencies by other means). Based on these experiences, countries were then invited to consider how to simplify the data validation process, identifying cases where the process of validation could be lightened. In cases for which countries’ data transmission for SDG purposes was not necessary, countries were invited to describe how this would be reflected in their national reporting platform (NRP), if applicable.

13. Custodian agencies of selected indicators were also invited to make proposals to lighten the burden of data validation. They were also invited to clarify the role they would like their agency reporting platform (ARP) to play in data transmission.

14. Table 1 lists the indicators selected for the pilot study and their associated agencies. Members of the Task Team were asked to summarise pilot results concerning each issue type.

Table 1. Scope of Work for the Second Data Flow Pilot

	Issue Type	Indicator	Custodian Agency
A	Indicators modelled by agency	3.9.1, 11.6.2	WHO
B	National data are produced by agencies using geospatial information	15.4.2	FAO
C	Non-statistical indicator based on a survey whose responses may be adjusted by the agency	5.1.1	UN Women World Bank Group OECD Development Centre
D	Indicators produced by agencies that use data, transmitted by a well-established process	17.2.1, 17.3.1	OECD
E	National data are transmitted already to an intermediary international database	6.4.1, 6.4.2	FAO
F	Indicators calculated by the agency without a validation of the nationally nominated SDGs focal point	3,6,1 9.1.2	WHO ICAO and ITF-OECD

15. A survey for country participants, and separate, customized surveys for each invited agency (FAO, ICAO, ITF, OECD, UN Women, and WHO) were prepared (see Annex 1). The surveys could be completed online using Survey Monkey, as a fillable pdf, or as a fillable Word document. Participant countries were free to choose the indicators they would like to pilot from the list, although they were encouraged to respond to questions for all indicators on the above list. The surveys were posted on the public wiki; any country was free to participate, although the message asking to participate in the pilot was sent to the participants of the UNECE Expert Meeting on Statistics for SDGs. Upon confirmation, country participants were asked to provide survey focal point contact information (if possible, at least two names).

16. Forty-three responses were received at the close of the survey collection period (3 July to 10 August, with some received after the deadline). Of these, 38 were countries and 5⁶ were agencies. Although most responses were contributed by countries within UNECE (31 of 56 UNECE member countries), 7 responses were also received from countries in other regions (ESCAP and ECLAC). Annex 2 lists the contacts for responses received from countries and agencies as of 28 August. The respondents to the survey in countries were the focal points for SDG data in NSOs.⁷

17. As separate activities initiated by IAEG-SDG countries, the survey also informed examinations of SDG data flows in the ECA (Africa) and ESCAP (Asia and the Pacific) regions. Members of IAEG-SDG from the ECA region summarized their experiences in an October 2018 note to IAEG-SDG. The survey also motivated a side event at the October 2018 ESCAP Commission, wherein a panel of ESCAP country representatives to IAEG-SDGs discussed their SDG data flow experiences.

D. FINDINGS

18. Chapter 2 summarises findings regarding participants' coordination of SDG data flows, use of national reporting platforms and attitudes toward SDMX. Chapter 3 presents country and agency experiences for each indicator studied, grouped by potential validation challenges. Common challenges and potential solutions (often offered by pilot participants and custodian agencies) are identified for consideration. Chapter 4 presents the findings of the self-analysis portion of the surveys. The report concludes with an overall summary of findings and practical suggestions to support transparent and efficient global data flows.

19. Initial findings were shared with the co-chairs of CCSA and IAEG-SDGs in late August 2018, and a more synoptic review was presented at the IAEG-SDG meeting in November 2018. The draft report was distributed for comments to pilot participants (countries and agencies), and subsequently, to the full Steering Group before finalization in early May 2019. The first results of this pilot have informed discussions of CCSA and IAEG-SDG in preparation of the data flow implementation guidance requested by the UN Statistical Commission (presented in March 2019).

E. LIMITATIONS OF THE STUDY

20. There are several limitations to this study that may affect the generalization of its findings to other country and agency experiences regarding SDG data flows. These limitations should be taken into account when evaluating the findings and suggestions of this report.

21. This pilot was fielded to describe experiences of countries in the UNECE region. Countries within the same region have more similar data production processes than countries of different regions. Many countries involved in the pilot study are members of OECD and/or participate in Eurostat data collection, another indicator of strong similarities across national statistical systems. Therefore, the pilot results may not reflect circumstances for countries where

⁶ Of these five agencies, one agency (FAO) provided a response for three indicators, another agency (OECD) provided data for two indicators.

⁷ Some agencies have underlined that as the respondents to the survey in countries were the national SDG data focal point in NSOs, the respondents may not have been aware of the data collection process and of the validation process for existing and well-established data flows.

variations in the process (e.g. diverse data sources) can affect the data flows and the comparability of data.

22. The relative maturity of statistical systems varied among participants. However, countries of the UNECE region have less need of statistical capacity building than in other parts of the world. Therefore, pilot findings may represent data flows in countries where statistical system coordination is likely to be strongest. For example, half of country respondents indicated they were DAC donors (19 of 27 respondents).

23. A third limitation of the pilot relates to the selection of indicators for study. The focus of the pilot was on indicators that may be particularly difficult for countries to validate given the nature of the data involved (e.g., data from non-traditional data sources, data modelled by agencies, data for non-statistical indicators). Even so, the indicators selected for this pilot were all Tier I (Tier II in one case). Data flows for Tier I indicators describe a “best case scenario” for indicators that may be difficult to validate because the tier designation indicates the indicator is conceptually clear, has an internationally established methodology and standards, and is produced regularly for at least 50 percent of countries and 50 percent of the population in every region where the indicator is relevant. Data validation for indicators using data from non-traditional data sources that are designated Tier II, for example, may be even more challenging.

24. Despite these limitations, the pilot was conducted with wide support from participants and custodian agencies. There was good collaboration with CCSA and IAEG-SDGs. The high number of countries who participated in this voluntary pilot shows the importance that countries attach to the issue of data flows and validating the data that is published for their country in international databases. It is clear that a strongly felt commitment to robust data flow relationships for successful global reporting on SDGs is widely shared.

F. ACKNOWLEDGEMENTS

25. The UNECE Steering Group (SG) is grateful to the 37 countries and 5 agencies who participated in the second pilot data flows. Although participants were advised that partial responses would be accepted, almost all participants responded to all of the survey questions. Further, to meet tight schedule deadlines, the pilot was conducted during the summer months which required contributions from participants and task team members alike during their holiday period.

26. The SG would especially like to thank the agencies involved which took time from their already demanding schedules to respond to the survey and follow up questions from task team members. Their dedication to ensuring transparency and improving understanding among data providers is highly valued. Without their involvement, this pilot study could not have been implemented.

27. The Task Team on Data Flows is co-chaired by France and Turkey. In addition, Denmark, Sweden, Kazakhstan, Russian Federation, United Kingdom and United States of America contributed sections to the report.

28. Observations and findings are based on replies received from participants of the second data flow pilot during July and August 2018. The task team is aware that conditions may have changed between that time and the finalization of this report, especially regarding new data validation requests and improvements of metadata.

II. NATIONAL COORDINATION OF SDG DATA FLOWS

29. Thirty-eight countries and five agencies participated in the second data flow pilot survey. In most participating countries (28 out of 38), the NSO has some formal coordination role. Six countries stated that all reporting is routed through the NSO, or that NSO is responsible for all the reporting.

30. Almost all participants currently report SDG statistics to custodian agencies along traditional lines, wherein 1) requests are sent to countries from custodian agencies and the responsible data producer transmits the statistics back to the custodian agency or 2) another process is already in place (e.g., statistics are pulled from an already existing database). Only 2 countries (i.e., Poland and USA) said they will direct the custodian agency to a national reporting platform (NRP) where they can find the statistics required⁸.

31. Half of the country respondents reported having a platform or web-portal for SDG statistics. An additional 12 countries stated they are planning to develop a platform/web portal for SDG statistics. Of the responding countries who currently have a platform, very few are using it for the transmission of data for global indicators to custodian agencies. However, 7 countries stated explicitly that they plan to develop the transmission process vis-à-vis custodian agencies.

32. Countries appeared to be generally open to, and willing, to use SDMX or other forms of standardized data transmission. Out of the 33 responding countries 25 replied that they are planning to use SDMX or that they are open to the possibility. Three of these have stated that they will need training. Six countries are undecided. One country stated that they want a unique DSD and MDS and database to avoid having to customize their transmission forms to the different custodian agencies. Only one country declined interest.

33. In addition, 20 countries supported the use of push and/or pull technology for indicators on their NRP or potential NRP.

34. Nineteen respondent countries indicated they are posting, or would post, statistics for global indicators calculated by custodian agencies on their national platform. In eight of these cases, the country indicated that this would require the country to validate the statistics before posting. Only 3 countries disagreed with this approach. Eight countries were undecided, but most of these state that they would consider it.

⁸ However, and as custodian agencies have stressed, a review of countries' data hubs highlights the need for a clear differentiation between the official SDG indicator (needed by the custodian agency for global reporting) and national proxies on the national reporting platforms.

III. ANALYSIS OF DATA FLOWS BY INDICATOR

35. The first part of the survey asked questions regarding specific data flows. The questions in this section were organized according to six themes, each relating to an aspect that may make validation of SDG statistics for the indicator challenging. In general, two indicators were selected as examples of each theme. A description of each of these thematic areas, the indicators selected, and country and agency responses are provided below. Findings are summarized at the end of each section.

A. INDICATORS MODELLED BY AGENCY (3.9.1, 11.6.2)

36. The second data pilot focused on indicators that may be especially difficult for countries to validate given the nature of the data involved. One such case are indicators where the data used to produce the indicator value is modelled by the custodian agency rather than collected directly. Two indicators sharing this data characteristic were selected for the second pilot study: 3.9.1 *Mortality rate attributed to household and ambient air pollution*) and 11.6.2 *Annual mean levels of fine particulate matter (e.g., PM2.5 and PM10) in cities*). The World Health Organization is the custodian agency for these indicators.

1. Description of the indicator(s) and rationale for study

37. Indicator 3.9.1 requires data on the national-level of human exposure to household and ambient air pollution and the burden of disease as expressed in deaths (mortality by cause of death). Indicator 11.6.2 requires data on the mean annual concentration of fine particulate matter. Both are estimated by WHO.

2. Country response

Indicator 3.9.1, Mortality rate attributed to household and ambient air pollution

38. Thirty-four countries responded to questions about indicator 3.9.1. Most countries (26 of 34) reported they do not produce this indicator at the national level. Of these, three are considering a methodology and four are planning to calculate the estimate at the national level in the near future.

39. Eight countries produced estimates for this indicator; one of which a national proxy indicator (i.e., number of deaths caused by registered socio-domestic intoxication cases). Among these 8, 3 countries reported routine (i.e., regular or annual) calculation of this indicator, while the remainder reported infrequent reporting (2 countries) or no response.

40. Ten countries indicated that the data producers for this indicator were official statistical institutions outside of NSO. The focus of these institutions related to hygiene, health or environmental issues.

41. Most respondents (21 of 29) indicated that they received a request for a national focal point from UNSD in 2017. In most cases, NSO was named as the national focal point for coordination purposes. However, two thirds of respondents (18 of 29) reported they did not receive a data validation request from the custodian agency. In these cases, the request may have been sent to another institution, such as the Ministry of Health, and was not shared with

NSO. As a result, most respondents indicated that NSOs did not validate data from the custodian agency because they were not aware of the request to validate.

42. Six responding countries indicated that they agreed with the estimate prepared by the custodian agency. Three of these countries explained the basis for this support. One noted they trusted the custodian agency and the expert responsible for the model. One indicated that the values produced by the custodian agency and the nationally generated estimate were similar. One indicated that the metadata were clear and reasonable. Nonetheless, most responding NSOs (21) could not validate the modelled data, most of these (15) citing that they were not requested to validate this indicator prior to publishing.

43. Twenty-two responding countries agreed with the proposal to allow publication of indicators without prior validation by the country if it is indicated that data are calculated by WHO and the status of validation by country (validated/not validated, pending validation). In most cases (15), countries explained that their support of the proposal still required their review of the process used to provide the statistics (such as the national input data used to model the statistics, the model used, and the output of the model). An additional five countries supported the proposal and endorsed a simple process where agencies only share the national input data used to model statistics and the output of the model for country awareness and identification of obvious error. Four responding countries did not support the proposal to allow publication of indicators without prior country validation.

44. Overall, responding countries prefer that if national data are disseminated in the global SDG database, the national focal point must review the model used to provide these global statistics. However, five countries indicated that national data modelled by the agency should not be published in the global database when the national focal point cannot validate the data (or underlying model) even if the data source and the status of validation by the country is indicated.

45. Eight responding countries indicated that they would include the indicator estimated by the custodian agency in their national reporting platform, either because it is an indicator used to monitor their national SDG strategy and no other national indicator is available, or due to communication purposes with the general public. Ten responding countries indicated they would not include the global estimated indicator in their NRP; seven of these countries indicated that they only publish on their NRP the data produced or validated by their country and 3 indicated that they produce a national indicator which differs from the globally harmonized indicator produced by the custodian agency.

Indicator 11.6.2, Annual mean levels of fine particulate matter in cities (population weighted)

46. Thirty-five countries responded to questions regarding indicator 11.6.2. Most of these countries (23) reported producing national data for this indicator, generally on an annual basis. Twenty-eight reporting countries indicated that national data producers were other official statistical institutions or non-statisticians, generally related to hydrometeor or environmental issues. Two countries indicated that they produce proxy indicators, which are very similar to the global indicator. Some countries reported that they are unable to produce this indicator because the state level data are not centralized in a manner that would allow for national estimates.

47. Twenty-five responding countries indicated that they had received a request from UNSD for national focal point information in 2017. In most of these responses, NSO was named as the national focal point for coordination purposes. However, most responding countries (20) indicated they did not receive a request to validate the data. Further, 17 countries could not validate the data modelled by the custodian agency, some citing as the reason that they were not aware of such a request.

48. Four responding countries indicated that they agreed with the estimate prepared by the custodian agency. Two noted they trusted the custodian agency and the expert responsible for the model. One indicated that the values produced by the custodian agency and the nationally generated estimate were similar. One indicated that the metadata were clear and reasonable. Nonetheless, most responding NSOs (19) could not validate the modelled data, most of these (11) citing that they were not requested to validate this indicator prior to publishing.

49. Nineteen responding countries agreed with the proposal to allow publication of indicators without prior validation by the country if it is indicated that data are calculated by WHO and the status of validation by country (validated/not validated, pending validation). Four of these countries agreed with the proposal, indicating that their responsibility was clear and the global process is transparent. However, in most of these cases (12), countries explained that their support of the proposal still required their review of the process used to provide the statistics (such as the national input data used to model the statistics, the model used, and the output of the model). An additional three countries supported the proposal and endorsed a simple process where agencies only share the national input data used to model statistics and the output of the model for country awareness and identification of obvious error. Six responding countries did not support the proposal to allow publication of indicators without prior country validation.

50. Overall, responding countries prefer that if national data are disseminated in the global SDG database, the national focal point must review the model used to provide these global statistics (that is, the national input data used to model the statistics, the model used, and the output of the model). However, six countries indicated that national data modelled by the agency should not be published in the global data base when the national focal point cannot validate the data (or underlying model) even if the data source and the status of validation by the country is indicated.

51. Seven responding countries indicated that they would include the global estimated indicator in their national reporting platform, either because it is an indicator used to monitor their national SDG strategy and no other national indicator is available, or due to communication purposes with the general public. However, 13 responding countries indicated they would not include the global estimated indicator in their NRP; 7 of these countries indicated that they only publish on their NRP the data produced or validated by their country and 4 indicated that they produce a national indicator which differs from the globally harmonized indicator produced by the custodian agency.

3. Agency response

52. WHO indicated that the quality and coverage of information about human exposure to urban ambient air pollution have improved significantly in recent reporting periods; more than 1,000 cities and towns have been added to the WHO Global Urban Ambient Air Pollution database since 2016. The improvement was facilitated, in part, through the WHO-led Global Platform on Air Quality and Health, which has enabled greater cooperation among partners (including academia) on data assessment and modelling methods.

53. Accordingly, the tracking of progress using indicator 11.6.2 is expected to improve. The database has been updated, and the updated version will be released in May 2018, featuring more than 4000 human settlements publicly reporting their ambient air quality levels.

4. Findings

54. Overall, there were similarities in responses for both indicator 3.9.1 and 11.6.2 regarding data availability, the validation process, and publication and dissemination. Most responding NSOs currently do not produce national data for these indicators. There are various reasons cited for this, chief among them being the absence of a national-level data source for the global indicator and the availability of proxy indicators very similar to the global indicator.

55. Most respondents provided contact information on the national SDG statistics focal point, and in general, NSO were the appointed national focal point with a coordination role. Nonetheless, most NSOs could not validate the modelled data from the custodian agency. The reason cited most often was that NSOs were not aware of such a request (perhaps the request was sent to another agency). When asked to validate the data for their country in the global database, most NSOs could not do so. There were two main reasons for this. First, other national institutions were responsible/had the required expertise to validate and there was not sufficient time to contact them. Second, there was insufficient information in the global metadata to ascertain why national and global estimates differed, or no national data were available.

56. With regard to the publication and dissemination of global estimates, if national data are modelled by a custodian agency and included in the global database, this should be noted in the database as modelled by the custodian agency and the calculations used should be reviewed by the national focal point prior to posting. If the national focal point cannot validate these data, 6 countries required the national data not to be published in the global database. Further, most NSOs would not publish national data modelled by custodian agencies on their NRP if the data were not produced or validated by the country/national focal point.

B. NATIONAL DATA PRODUCED BY AGENCIES USING GEOSPATIAL INFORMATION (15.4.2)

57. The second data pilot focused on indicators that may be especially difficult for countries to validate given the nature of the data involved. Another such case are indicators where the data used to produce the indicator value is generated from geospatial information⁹. One indicator with this data characteristic was selected for the second pilot study: 15.4.2 *Mountain*

⁹ Reviewing this report, FAO stressed that geospatial information is used to estimate national figures for this indicator only when no alternative data has been provided from the countries. Knowing that a very limited number of countries are able to produce data for SDG 15.4.2, FAO has produced country estimates using geospatial data, has requested countries to validate these estimates and invited countries to provide alternative data through national sources.

Green Cover Index (MGCI)). The Food and Agricultural Organization is the custodian agency for this indicator.

1. Description of the indicator and rationale for study

58. Indicator 15.4.2 is one of two indicators in place to measure the target 15.4 “by 2030, ensure the conservation of mountain ecosystems, including their biodiversity, to enhance their capacity to provide benefits that are essential for sustainable development.” MGCI is meant to measure the changes of the green vegetation in mountain areas and provides an indication of the status of the conservation of mountain environments. A direct correlation between green coverage of mountain areas and their state of health is recognized by the scientific community. Green mountains and especially trees fulfil various ecosystem services and thus prevent, for example, soil erosion and mudslides.

59. The Food and Agriculture Organization of the United Nations (FAO) has been designated as the custodian agency for this indicator, being responsible, amongst other things, for collecting data from countries, producing regional and global estimates, and providing statistical capacity development support to countries for their regular reporting. UNEP is listed as partner agency in this context.

60. Given that most countries are not in a position to produce data for 15.4.2 and in order to establish a global baseline based on comparable and harmonized national estimates, FAO relied on Collect Earth, a remote sensing application, to determine the coverage of green vegetation in mountain areas. As these estimates, however, were not based on national official data, FAO proceeded to validate the data with all countries and offered countries the possibility of providing alternative national data if available.

61. IAEG-SDGs designated MGCI in 2016 as a Tier II indicator, meaning that the measure was considered to be conceptually clear and based on an established methodology, although not routinely measured in all regions of the world where relevant. In the case of MGCI, the tier designation was given although there were no metadata available at the time of classification. In 2017, a global baseline was prepared using a dataset of 500,000 sampling points which were collected at regional scale by circa 30 scientific and government institutions and coordinated by FAO. At the end of November 2017, MGCI was reclassified to Tier I because of its data availability.

62. On 20 December 2017, FAO contacted either the NSO SDG focal points (for the 90 countries that had provided them at the time) or alternatively the NSO Head to validate the national estimates of SDG indicator 15.4.2 calculated by FAO. Concerns or other feedback as well as country data could be sent until 18 January 2018. National data were attached to the email by FAO (a pdf file) and countries were invited to introduce alternative national data in a separate Excel table if such data was available.

63. After this first contact, FAO shared a detailed document on the data collection methodology, definitions and the methods to assess the indicator with all countries. FAO also provided to the raw data used for calculating the SDG indicator 15.4.2.

64. The UNECE Data Flow Pilot for indicator 15.4.2 consisted of three questions focusing on (national) data availability/production, three more questions on the data validation process in 2017 including the national SDG statistics focal point request by UNSD and two final

questions on the country opinion of publication of non-country and/or non- validated data (eight questions in total, numbers 24 to 31).

2. Country response

65. Thirty-two countries responded to pilot questions regarding indicator 15.4.2.¹⁰ At present, 26 of 32 responding countries (81%) indicated they do not produce data for this indicator. Two of these countries stated that the indicator is not relevant for their country as there is no mountain area within national borders. Five countries (16%) responded that the indicator is under development. One country confirmed that it now produces the indicator fully aligned with the global metadata. In two of the five countries presently developing the indicator, a periodicity of three years is foreseen; while the others could not provide any information on this. Only the country already producing MGCI provides the data annually, starting from 2017. Four countries answered only the first question and skipped the rest. Therefore, 28 countries are considered as total in the subsequent analysis.

66. Fourteen countries responded to questions regarding possible national data providers. A range of possible data providers were mentioned, including NSO, ministries for forest and agriculture or nature protection as well as national institutions/agencies responsible for topics like agriculture, cadastre/mapping/cartography, environment or forests.

67. In 2017, UNSD requested countries to provide them with the contact information for their national SDG statistics focal point. Twenty-one responding countries confirmed that they provided this information. In most of these cases, NSO was designated as the national focal point. Five countries¹¹ skipped the question, and two said they had not provided this information.

68. As the custodian agency for indicator 15.4.2, FAO sent a data validation request to the national SDG focal points in December 2017. This request reached 19 of the 28 countries (68%). Six countries were not aware of such a data validation request, and three countries skipped the question. One reason could be the missing nomination of an SDG national focal point (see above). However, four of the six countries that said they were not aware of the request affirmed in the prior question the provision of a focal point contact information to UNSD. Other possible reasons could be wrong email addresses, as well as changes in staff assignments or general organisational/technical issues.

69. In summary, 7 countries¹² validated the data, while 19 (68%) did not. The only country that already produces MGCI in compliance with the metadata, however, did not validate the data prepared by the custodian agency due to differences between FAO's estimation and their national values. The main reason for the different value was identified to be the lower number of observation points in FAO's calculation. Additionally, two countries did also not validate the custodian agency's data because of inconsistent values. Other reasons for non-validation (16 countries) were missing country data for the assessment, technical issues and/or missing responses of other involved national stakeholders, missing instruments, non-relevance of the

¹⁰ 33 countries responded. But one of these countries had not received the relevant information from the responsible ministry to answer the questions sufficiently within the period of the pilot study and therefore did not answer further questions.

¹¹ This excludes the four countries that only answered the first question on national data availability and skipped the following questions.

¹² One of the two countries that reported the indicator as not being relevant still validated the data.

indicator, concerns on the methodology or that FAO's request did not reach the person in charge. One country replied they could not validate the data because of an issue concerning the interpretation of the trend of the indicator related to climate change. Two countries skipped the question.¹³

70. The initial SDG data flow guidance noted by the UN Statistical Commission in 2018 recommends the use of globally harmonized national data only to calculate the global and regional aggregate when national data or statistics provided by the custodian agencies are not validated by the country. But the globally harmonized national data may not be published in the SDG global database. As alternative to this proposal, the pilot examined the opinions of the participating countries on another possible procedure. In cases where data validation might be difficult for countries as they do not produce the indicator, an alternative guidance could be a note including "both that the data are calculated by the custodian agency and the status of validation by the country (e.g. validated by country, country can't validate, not validated by country, pending review)". Of the 22 countries that responded to this proposal 9 did not favour it while 3 countries fully agreed to it, 7 wanted to check the custodian agency's calculation before their release and 3 preferred pre-release information for country awareness and identification of obvious errors. Six countries skipped the question

71. The final question examined if countries would publish the indicator calculated by the custodian agency on their national reporting platform (NRP). Three out of twenty-three countries (13%) approved because of the national relevance of the indicator, while four more countries limited their "yes" to communication purposes. One country answered with "yes" and "no" and mentioned that it would prefer to produce the indicator in accordance to the internationally accepted methodology on its own. Seven more countries responded they prefer country data. Two countries would accept and publish indicators calculated by custodian agencies if the method has been accepted and data have been assessed as high quality by the country, which means that data have been validated. Two countries indicated they would publish only indicators with national relevance on their NRP. One country skipped this question as they had no NRP implemented yet.

72. One country drew attention to the error of $\pm 1\%$ at global level that increases to $\pm 2\%$ at regional level and is expected to rise further at the national level. Therefore, it was suggested to publish FAO's data with the corresponding accuracy.

73. Regarding the publication of non-validated data, two countries emphasised in their comments that the decision on the data publication should be left to the countries following the principle of subsidiarity. One of these countries underlined that it would not be seen as non-transparent or in contrary to the Fundamental Principles of Official Statistics if national data produced by agencies are only used to calculate the global and regional aggregates. Another country underlined a footnote is not sufficient as users of the data may gloss over it and think it is the country response.

74. Three countries indicated methodological concerns with the indicator. For the first country, an increase of the green cover of mountains would not only be perceived as positive as it may be a sign of climate change (e.g. melting of snow and ice, trees growing higher up).

¹³As these two countries also skipped the prior questions on providing the SDG focal point as well as receiving the FAO's request, it is unclear whether or not the validation request of the FAO reached them and what the result was. According to the global database, one of them has validated the data, while the validation is pending for the other one.

With regard to that, the country proposed to add a caveat in the metadata concerning its relevance and the interpretation of its evolution. The second country had a general problem with the meaningfulness of the data and saw a great uncertainty as measuring the change over 3 to 5 years will most likely not provide meaningful information. The third country disagrees with the FAO methodology in the national context.

3. Agency response

75. FAO is at present the (sole) custodian agency for 19 SDG indicators and together with other agencies responsible for three more SDG indicators. Three indicators were part of the second UNECE data flow pilot, namely 6.4.1, 6.4.2 and 15.4.2.

76. For indicator 15.4.2 FAO contacted all national SDG focal points collected by UNSD in 2017 by email. A short note was provided, with background information on the indicator, the classification used for defining mountains and the methodology used for obtaining the data at country level. The recipients were requested to validate the data. FAO estimates that 7 percent of the countries refused to validate the data, 12 percent validated the data, 20-25 percent started a validation process and 56-61 percent did not respond to the validation request. FAO supported countries that could not validate the data by many email exchanges and Skype discussions to describe the methodology more in detail and to encourage them to validate the data. Two capacity development workshops have been organized (December 2017; November 2018) for all countries that requested a training. The workshops' goal was to enhance countries capacities so as to enable them to validate the MGCI data collected by FAO and to intensify the sampling points at national level.

77. FAO plans to update the metadata with missing information, e.g. a schedule for data validation, guidance on how to interpret an increase of the green cover of mountains in cases where it should not be regarded as positive (e.g. losses of permanent snow cover area), and date of release. A better understanding of the methodology would enhance country validation.

78. FAO was also interviewed on their opinion on the alternative proposed guidance to publish, additionally to regional and global values, the national data or statistics calculated by FAO in the UN global database as well as the validation status by country (see 12). FAO agreed to this proposal because the process of validation will be simplified, but FAO will still offer countries training to make them able to implement the methods for their country and to check the data calculated by FAO¹⁴.

¹⁴ In addition, FAO is promoting land monitoring and assessments at national level with the same methodological approach in more than 50 countries. Twenty-two countries have already officially submitted data collected by using the FAO tool, Collect Earth, under the reporting schemes of the Climate Change Convention (UNFCCC REDD+ Web Platform <https://redd.unfccc.int/submissions.html>). It is expected that many of these countries will elaborate the MGCI using their national database

4. Findings

79. As many as 33 of the 38 participating countries of the second UNECE data flow pilot answered the questions on indicator 15.4.2, indicating a high importance of this topic for the countries.

80. The indicator is not relevant for all countries as there are countries with neither mountains nor local elevation range higher than 300 meters. But even so, for some of these countries, FAO's geospatial tool produced results¹⁵.

81. One country declared they already produce the indicator in compliance to the metadata, but with a different result than FAO's estimation. However, no country sent data to substitute FAO's estimates.

82. FAO contacted the national SDG statistics focal points in NSOs, using the list provided by UNSD, to validate its estimates for SDG indicator 15.4.2 between December 2017 and January 2018. However, as many as up to 61 percent of countries did not react to FAO's validation request. Possible reasons could be wrong email addresses as well as changes in responsibilities and contact persons or general organisational/technical issues. The time of year could also be a factor.

83. Inadequate national coordination may play a part in the failure to validate the estimates. Often times FAO experienced that the SDG national focal points did not share the data with the relevant ministry/technical unit therefore hampering the validation process. It could be helpful if countries identified specific national focal points for each SDG indicator, besides the generic NSO SDG focal point. Custodian agencies can support countries by providing names of persons or organisations in the countries that they are in contact with regarding similar issues or that have previously been appointed focal points for specific pre-existing data flows.

84. Nineteen of the countries responding to question nr. 29 in the data flow pilot questionnaire had not validated the data calculated by FAO for different reasons, e.g. inconsistent values, missing country information or instruments for the assessment, non-relevance of the indicator or disagreement with the metadata – in particular due to a different definition of mountain areas.

85. FAO published its estimates also in cases where the data had not been validated because of no country feedback. Some countries were critical of this process. Footnotes itself were considered as not sufficient as users may gloss over it and regard the data as country approved or country data. However, this issue has now been resolved by the final IAEG-SDG Guidelines on Global Reporting that allow for custodian agencies to publish data in case of non-response by countries, provided that a reasonable amount of time has elapsed.

86. Most countries prefer to publish their own calculated results on their national reporting platform, but would also publish national data produced by agencies, if they considered the data quality sufficiently good. One country recommended publishing data estimated by agencies with the corresponding precision rate.

¹⁵ The indicator value is very low and a footnote indicates in the UN SDGs Global data base, these data should be removed because of no mountains in the country. FAO took notice of this issue and is now in the process of removing the data for these countries from the database

C. NON-STATISTICAL INDICATOR BASED ON A SURVEY WHOSE RESPONSES MAY BE ADJUSTED BY THE AGENCY (5.1.1)

1. Description of the indicator and rationale for study

87. In late 2017/early 2018, custodian agencies of SDG 5.1.1, *Whether or not legal frameworks are in place to promote, enforce and monitor equality and non-discrimination on the basis of sex* (UN Women, World Bank's Women, Business and the Law initiative and the OECD Development Centre's Social Institutions and Gender Index) sent UN member countries an email requesting review and validation of a questionnaire compiled by their legal experts on the legal frameworks for gender equality. They asked the countries to review and validate a pre-filled survey consisting of 45 "Yes/No" questions on national gender legislation. This process reduced country respondent burden, while giving them the opportunity to engage and contribute, including adding quality checks via validation.

88. At the IAEG-SDGs meeting in April 2018, the tier status of SDG 5.1.1 was raised to a Tier II (from a Tier III), increasing the incentives to measure this indicator. The 2018 UNECE Data Flow Pilot survey attempts to track data flows on a "non-statistical indicator based on a survey whose responses may be adjusted by the agency" such as SDG 5.1.1.

2. Country response

89. Out of 33 countries participating in the survey, almost two-thirds have identified a specific data provider for indicator 5.1.1. Less than half of the national focal points identified were "NSO Statistician." The rest of the national focal points were from "relevant policy-making institutions."

90. Out of the 20 countries transmitting information on a national NSO focal point to UNSD in the fall of 2017, 9 received the questionnaire sent by UN Women. Among them, 5 received the request indirectly. Four reported that they "were not aware of such a request."

91. Among the countries who replied to this question in the data flow questionnaire, less than one-third (6 countries) completed and/or validated the questionnaire. Only a few survey respondents gave specific information regarding negotiating discrepancies with the custodian agency; there does not seem to be any major disagreements over the negotiation process except one question with one country concerning the interpretation relative to gender discriminations of national safety measures which protect women –and not all workers- from hard work dangerous for their health. Less than half (42%) of the country respondents agreed to letting IAEG help settle any pending disputes, many countries expressed a preference to communicate directly with the custodian agency.

92. Half of the countries indicated that they would post the indicator (generated by the custodian agency) on their NRP. The other half reported that they won't post and will not retain this indicator for their national implementation of SDGs.

3. Agency response

93. The custodian agencies for indicator 5.1.1 reported just over half of the countries (50 of 89 or 56%) identified a national focal point, of which only one-fourth (26%) were NSO Statisticians. Although the country survey results show a somewhat higher percent of national focal points as NSO Statisticians (36%), the implications are the same—since this is not a statistical indicator, most of the national focal points have a policy-related (not statistical) background.

94. The custodian agencies estimated a global response rate of just over 50 percent (46 of 89), which includes completed questionnaires as well as surveys undergoing validation. By region, the response rates range from a high of 59 percent (Europe and North America) to a low of 24 percent (Sub-Saharan Africa). Based on the pilot, discrepancies were relatively mild and were resolved directly with the countries. Countries were contacted and given another chance to intervene before data is “finalized” by the custodian agency. Data would not be released without this validation process, which was approximately 3-4 months long. According to the guidelines presented by IAEG-SDGs at the 49th UNSC, when globally-harmonized national data provided by the custodian agency are not validated by the country, the globally-harmonized national data can be used to calculate the global aggregate, but, cannot be published in the SDG global database. But to increase transparency and be able to publish data used to calculate the aggregate even when they are not validated by the country, the custodian agencies would support the questionnaire suggestion to include a footnote indicating the answer was not validated or was pending validation/review, as specified by the country concerned. However, during the validation process, they would continue to offer countries training to better understand their assessment framework.

95. While the custodian agencies think the current process works well, they will support IAEG reconciliation for any potential on-going issues.

96. The custodian agencies plan to release data for SDG 5.1.1 every other year, beginning in the first quarter of 2019. At present, they have no plans in place to automate the exchange of SDG data and metadata.

4. Findings

97. The custodian agencies engaged with this indicator report that this process is inclusive and productive. They reported planning to release SDG 5.1.1 during the first quarter of 2019. After the survey was closed, UN-Women provided some clarification on the process. Some of the countries had not been approached at all in the first cycle of data validation initiated by UN-Women. This was the case for 7 of the countries replying to the data flow pilot questionnaire, which would explain some of the disconnect regarding points of contact. Another reason could be due to job changes since the last collection of NSO focal points made by the UNSD. Still, some 40 percent of the NSOs from countries that participated in the data flow pilot, *and* were approached by UN-Women for this request, were not aware of the existence of the questionnaire. This gap needs to be bridged.

98. The issue could possibly be resolved via asking the UNSD and UN Women to compare their contact lists for this indicator, and then sending it out to the country focal points for verification. Another approach may be to split the question to identify a “national focal point”

and the “subject matter expert” for the specific indicator. The former is the facilitator. The latter is the data provider and negotiator. Of course, it is possible for one person to fill both roles.

99. UN-Women confirms that it makes all efforts to ensure correspondence reaches the appropriate country counterpart, using the more recent UNSD contact lists, complemented by any updates it receives directly from countries.

100. The 5.1.1 indicator is not a statistical indicator. When there is disagreement between agencies, the primary information provider can negotiate with the custodian agencies. If an agreement cannot be met after a reasonable effort, custodian agencies would support the suggestion to publish national data with their status of validation (not validated/pending validation /review as specified by the country). Some countries support the idea of publishing the status of validation of any indicator published in the UN database, but some are still reluctant to allow the custodian agency to publish data not validated even with this notation provided.

D. INDICATORS PRODUCED BY AGENCIES THAT USE DATA TRANSMITTED BY WELL-ESTABLISHED PROCESS (17.2.1, 17.3.1)

1. Description of the indicator and rationale for study

101. The collection of ODA related statistics is part of a long-standing OECD statistical practice, governed by the OECD/DAC Working Party on development finance statistics (WP-STAT). The data flow pilot used indicators 17.2.1, *Net official development assistance (ODA) as proportion of the OECD Development Assistance Committee donor's gross national income* and 17.3.1, *Foreign direct investment (FDI), official development assistance, and South-South Cooperation as a proportion of total domestic budget* as examples of indicators produced by agencies that use data already transmitted to the international organisation by the countries through an already established process. Both indicators are classified as Tier I in the SDG global indicator set, but only 17.2.1 is produced. Indicator 17.3.1 is not produced due to both methodological issues and lack of data.

102. DAC statistics are often produced and transmitted to OECD by a national Development Cooperation Agency or by the Ministry of Foreign Affairs. National SDG focal points have not been requested to validate the indicators before publication in the global SDG indicator database.

103. The reasons for choosing to investigate the two indicators are twofold. Firstly, the task team wanted to get an indication on whether countries would agree to the indicators being posted in the global database without the (re-)validation of the indicator by NSOs, given the fact that the statistics have already gone through a validation process when it is first transmitted to OECD/DAC. Secondly, the task team was interested to hear countries and OECDs views on the metadata that have been communicated for the two indicators and on the tier classification. The task team's starting point was that indicator 17.2.1 was classified correctly but that indicator 17.3.1 in all but actual classification is a Tier III indicator.

2. Country response

104. To the task team's surprise, 13 out of 25 countries that replied to the question about the metadata for the indicators said that the metadata provided on the UNSD website was sufficient for a Tier I indicator. Eleven countries found the metadata not sufficient. Eight of these added that this response applied for indicator 17.3.1 but that the metadata for 17.2.1 was fine.

105. Following up with a few countries the task team concluded that respondents have generally examined the metadata for ODA and found that sufficient. One reason for this could be that the indicator has often been identified as not relevant for their country which has resulted in a not very thorough investigation of the metadata. Another could be that as a general rule NSO does not produce the statistics that go into the calculation of the indicator and hence NSO does not have a clear idea about what would constitute sufficient information for the indicator.

106. A majority (16 of 17) of countries stated they agreed with the current process of data collection from country to OECD, and the calculation of ODA by OECD. The one country that disagreed did so mostly from the point of collection calendars. A majority (16 out of 21) also supported the guiding principle for the indicator that was proposed in the questionnaire:

When data for an indicator are transmitted to agencies and validated according to a well-established process by a national data provider designated by the country, countries also should designate the same data provider as an SDG focal point for the indicator to avoid duplication and confusion.

107. However, some countries stated that their reply only applied to the two investigated indicators and that any other indicators of the same kind (separate, but well-established collection and validation process) needed to be investigated on a case by case basis.

108. Four countries stated that they wanted a process where the indicators are (re-)validated by NSO before they are published in the global database.

109. Fifteen out of twenty countries responded that they would be open to including the ODA indicator calculated by OECD in their national SDG indicator set. Some of these also said they would vet and validate them first. Four countries stated that they only post indicators produced by national producers on their SDG data platform or portal.

3. Agency response

110. OECD stated that the current process regarding data collection for ODA works well and that the WP-STAT (under OECD/DAC) continuously engages in improvements to data and reporting mechanisms. It seems no plans for changes to the data collection relative to the ODA indicator is known to the rapporteur. OECD sees no need for additional validation by the national SDG focal points since the data are already validated by the country in another process.

111. OECD also stated that the tier classification is a matter for IAEG and that OECD (and the IMF) have raised some serious methodological issues with indicator 17.3.1 in IAEG-SDGs. OECD mentions several options to improve the indicator, for example, the data reported by OECD under indicator 10. b.1. on total resource flows for development could be used in the numerator, and the denominator could be changed to GDP or population, which would make the data under this indicator comparable over time and between countries. It could also be an

option to use figures on Total Official Support for Sustainable Development (TOSSD) when these become available. Any change – including interim use of proxy indicators while a final resolution is sought – would however be up to IAEG SDGs to discuss and decide.

4. Findings

112. As stated, the motive behind choosing these indicators to be investigated in the data flow pilot was twofold. The task force concludes that given the sample of countries that replied to the questionnaire, most countries supports the current process for collection of ODA related statistics. However, the national coordination would be facilitated if OECD would communicate to NSO SDG focal points the contact information for the national data provider.

113. OECD could also consider establishing a process where NSOs, as a courtesy, are given the opportunity to examine the indicator(s) and consult with the national data provider, some time before they are published in the global database.

114. Regarding the issue of metadata and tier classifications, the task team found, after drilling down into the subject, that there are some serious methodological issues with indicator 17.3.1 as it stands. It would recommend that the IAEG-SDGs either consult the OECD about using a readily available substitute indicator (such as 10.b.1 as a share of GDP), or downgrade the present indicator to Tier III and start the process to find a new indicator for this target.

E. NATIONAL DATA ARE TRANSMITTED ALREADY TO AN INTERMEDIARY INTERNATIONAL DATABASE (6.4.1, 6.4.2)

1. Description of the indicator and rationale for study

115. The indicators investigated were 6.4.1 *Change in water efficiency over time* and 6.4.2 *Level of water stress: freshwater withdrawal as a proportion of available freshwater resources*. Twenty-nine countries provided responses/comments to the section on those two indicators.

116. National data and information on water resources, water use and agricultural water management are key to monitor the water-related Sustainable Development Goals (SDG) indicators and to provide policy makers with comprehensive information on the state of water management in their country in support of national policy formulation and assessment.

117. Regarding these two particular indicators, it is not clear how the data flows will work. OECD/Eurostat/UNSD collect data on water biannually and the data can be partly used for monitoring of indicators 6.4.1 and 6.4.2. The next data collection, at the time the pilot was conducted, was foreseen to for 2018. However, FAO sent a questionnaire on data on water for SDG indicators in spring 2018. This creates an apparent situation of double-reporting. FAO comments that data collected by OECD/EUROSTAT do not provide enough information in order to produce water related SDG indicators, and that a discussion has been initiated in order to formally merge the two questionnaires.

2. Country response

118. Twenty-nine countries responded to the pilot questions regarding indicators 6.4.1 and 6.4.2. To the question whether metadata posted on the UNSD website provide sufficient information needed to produce national statistics for the two indicators, 13 countries responded

‘yes,’ 6 countries ‘no,’ and 10 countries provided another answer. One country commented that metadata for indicator 6.4.1 was sufficient but it was not the case for the indicator 6.4.2.

119. On the provision of a national SDG statistics focal point, 22 countries responded that they had provided such information, three countries had not and one country responded that it was not contacted.

120. Regarding the question on reception of a request from FAO on nomination of a contact person, 8 countries responded that they received the questionnaire indirectly with the focal point in copy and 11 countries responded that they had received the questionnaire directly to the focal point nominated. Six countries were not aware of such a request.

121. Nine countries answered that a focal point dedicated to FAO's questionnaire had been identified, and seven countries that they have an established focal points for reporting the data to Eurostat. The other answers were evenly distributed between other possibilities.

122. Fourteen countries replied that they had responded to FAO's questionnaire, 10 countries hadn't. Replies to the question about the reason for not replying to the FAO questionnaire were sparse. Six countries replied that they did not want to duplicate data transmission. Twenty-two countries were positive towards encouraging FAO to coordinate their data collection process on water with existing data flows.

123. The data flow pilot introduced a proposal for lightening the burden by allowing custodian agencies to publish data drawn from already available databases furnished by national, or nationally verified, data for the purpose of SDG global reporting. Eighteen countries agreed to this but six countries preferred to calculate their own national statistic and transmit it to FAO in order to more closely monitor country statistics for global SDG indicators.

124. Finally, on the question on whether a country would include data calculated by the custodian agency in their national data platform, 13 countries replied ‘no,’ and 8 replied ‘yes.’

3. Agency response

125. Globally, 95 countries (51%) have nominated AQUASTAT national correspondent and 60 countries (32%) have sent back the questionnaire (as of the time of filling in the data flow questionnaire sent by UNECE). FAO states that the 2018 data collection allowed collecting more data than previously collected through the former AQUASTAT questionnaire.

126. As the custodian agency for this indicator, FAO is responsible for collecting the necessary data from countries. However, as some of the data items required are already collected by other agencies through different instruments, FAO has established a dialogue with OECD/EUROSTAT and UNSD to harmonize the definitions used in their respective questionnaires, and better coordinate the data collection process. However, this means that existing OECD/EUROSTAT or UNSD data collections have to be adapted (i.e. include more information) to the needs of the SDG data collection. Furthermore, some of the existing questionnaires seem too complex for developing countries and the metadata is not detailed enough. However, FAO is providing support to the countries through the Integrated Monitoring Initiative for SDG6, reinforcing and complementing the information available in the metadata. Countries will also be given the opportunity to calculate their own statistics for the SDG a

127. FAO clarified that they are waiting for how to proceed on situations when nationally published data conflicts with internationally published data. Nationally published indicators would only be useful in situations when not only the results, but all the variables of the indicators subject to calculation are available to them.

128. FAO also noted that, as water statistics are cross-sectoral and irrigation statistics are cross-thematic, it was difficult to identify the most appropriate AQUASTAT national correspondents.

129. Finally, FAO agreed that the validation status of globally harmonized national statistics should be noted with a flag (official data as opposed to estimated/unofficial data), and that IAEG-SDG should serve as a referee in case of disagreements between countries and agencies.

4. Findings

130. Most responding countries seem to agree on the necessity of encouraging FAO to coordinate data collection with existing data flows already implemented (22 positive replies). Furthermore, most responding countries have also provided information on national SDG statistics focal points (22). Eighteen countries supported the proposal on coordination of international data flows. The other replies are more evenly distributed between countries. The comments provided by the countries indicate that there is a lack of clarity on the process as well as lack of understanding for why data from their NRP are not used as well as why countries are asked to duplicate reporting.

131. FAO's answer indicates that existing data collection for water is not sufficient to cover water related SDG indicators. Furthermore, FAO reports an increase in response from countries comparing with the previous AQUASTAT data collections (though without indication of response rate in previous collections).

F. INDICATORS CALCULATED BY THE AGENCY WITHOUT A VALIDATION OF THE NATIONALLY NOMINATED SDGs FOCAL POINT (3.6.1, 9.1.2)

1. Description of the indicator and rationale for study

132. There are several indicators that are calculated by a custodian agency without asking the national SDG focal point for validation. The examples looked at here are indicator 3.6.1, *Death rate due to road traffic accidents*, with World Health Organization (WHO) as the custodian agency, and indicator 9.1.2, *Passenger and freight volumes, by mode of transport*, with the International Civil Aviation Organization (ICAO) and the International Transport Forum (ITF) as the custodian agencies. The Task Team on Data Flows was interested in finding out whether countries were publishing their own data for these indicators, and whether countries agree with the figures included in the global reports. The task team also wanted to find out what level of support there was for the proposal that, even when there are existing processes, national statistical offices should be given the opportunity to validate any new or changed data for SDG indicators calculated by a custodian agency.

2. Country response

Indicator 3.6.1, death rate due to road traffic accidents

133. Information about experiences with this indicator was received from 35 countries. For all except one data for their country is reported in the WHO Global Status Report on Road Safety 2015.¹⁶

134. Most responding countries (84%, 22 countries) confirmed they had not been asked to validate their country's data, although more than two-thirds (69%, 24 countries) confirmed that these data are collected through a process that pre-dates SDGs. Some countries have confirmed that they are having difficulties identifying existing focal points in their country for this indicator and they have not been informed of the data validation process. One country confirmed that an expert, though not the SDGs focal point for that country, had asked them to validate the data.¹⁷ Two-thirds (66%, 23 countries) confirmed they had provided details of their focal points to the UNSD.

135. Countries were also asked whether the custodian agency had provided sufficient information to allow for validation. As only one country had been asked to validate their data, a fifth (20%, 7 countries) said that as no request to validate had been received, no metadata had been received either, and so they could not comment on the quality of the metadata. A further 49 percent (17 countries) said that the information was not adequate, but this could also be because no information was provided directly.¹⁸ Metadata for this indicator are available in the SDG metadata repository and four countries were satisfied with the quality of the metadata available there.

136. Countries were then asked to check their national data against the data available in the global report.¹⁹ Of the 29 countries who were able to check, 9 countries (31%) were not satisfied with the figure presented in the global report, finding it inconsistent with their own data. A further 10 countries were content that the figure was similar, or close enough for international comparisons. It was suggested that the differences between the figures in these cases could be, for example, due to the use of national population projections calculated by the country compared with those calculated by the UN department of population. Another cause of the discrepancy may be different definitions (e.g. country using police data with a death cut-off rate of 30 days and WHO taking into account any death due to road traffic accident regardless the date of death they occur). Two countries said that they trusted the figure produced by the agency.

137. The task team had proposed in the questionnaire that when a custodian agency calculates new data, or adjusts existing data for countries, and then publishes these data in the global database, the agency should always send these new figures to the national SDG focal point for

¹⁶ In our review of the report, Ukraine is the only one of the 26 not included.

¹⁷ It is not clear from the country's (Lithuania) response if this request for validation was specifically for the SDG indicator, or part of the existing process.

¹⁸ This question may have been misinterpreted by respondents. It is not clear whether the eight responding countries said that the information was not adequate because they didn't receive it, or because they think the metadata are not clear enough.

¹⁹ This question may also have been misinterpreted by respondents. More countries than 14 answered this question, but their answers seemed to be answering a slightly different question.

validation, even if an existing method for collecting data and validation is well-established. Of the 33 countries who commented on this proposal, 29 countries (88%) were in favour.

138. Of the 26 countries who responded to whether they would consider publishing data calculated by the custodian agency for their country on their National Reporting Platform, 6 countries said they would, 5 said they might while 15 said that they would not. The undecided countries were either not sure, because they have not yet developed an NRP and so have not defined criteria for publishing data or said that they may consider using data calculated by a custodian agency if national data were not available and the methods and data sources were validated and approved by their focal point.

Indicator 9.1.2, passenger and freight volumes, by mode of transport

139. This indicator covers several different issues, asking for passenger and freight volumes by mode of transport. The metadata provided by custodian agencies addresses only air, road and rail transport mode. For most countries, this information is taken from questionnaires they complete for the ICAO (air) and the ITF (rail and road). Data are modelled for countries who are not members of the ITF.

140. Information was received from 31 countries about this indicator, all of whom said they produce national statistics for this indicator. In total 26 countries (84%) said they produce national statistics for all three elements (air, rail and road), while 6 countries said they produce national statistics for just part of this indicator. Almost half of the countries (17 of 35) produce these data annually, 5 produce data monthly, one produces quarterly, while 7 countries produce data covering different periods for different elements of the indicator, such as quarterly for air and rail, but monthly for road.

141. These data are produced, for the most part by NSOs (26 countries), often with input from other organizations such as the Ministry of Transport or the Aviation Authority. Almost two-thirds (19 countries) said that they provided details of their focal points when asked last year.

142. None of the countries received a request (that they were aware of) to validate their country's data, although all except two, confirmed that they were members of ICAO and ITF. Most of the countries knew that they complete questionnaires for these organizations (19 of 31 respondents for ICAO, 17 of 31 for ITF), but the others did not know, or did not receive a questionnaire.

143. Countries were also asked to look at the data available online and compare it with their own data. Of the 30 countries who responded to this question, only one said that the data were accurate. A few countries were unsure of how accurate the data were, three because they didn't have their own data, one because there was no data for their country in the global database and eight because they needed more information or more time to discuss with topic experts. The biggest problem seemed to be a problem with units²⁰, and it's worth noting that while some countries said that they could not validate the data, this was sometimes due to just one component, not all three.

²⁰ SDG database used to have a unit problem between aviation and road/rail but that's seem no longer the case when this report was finalized.

144. Only four countries would be content to publish these data for their country in their own national reporting platform, but a further five would consider it. The 15 countries who would not publish this figure said that they would only publish their own data, or data they have validated.

145. Of the 30 countries who responded to the proposal that custodian agencies would need to validate data with countries prior to publishing, 28 were in favour.

3. Agency response

Indicator 3.6.1, death rate due to road traffic accidents

146. WHO believe that the global metadata for this indicator are precise, and pointed out their data are calculated from Civil Registration and Vital Statistics data (CRVS), with ITF-IRTAD (International Road Traffic Accidents Database) usually coming from Police records. They reported that they ask countries (the National Data Coordinator²¹ for the Global Status Report on Road Safety) to validate the data they publish and estimate that between 81 and 90 percent of countries comply.

147. WHO said they aim to convince countries to share their CRVS data through ongoing conversations, including, when required, travel to the country to hold face to face discussions. WHO suggest that clear explanations of how CRVS data are collected, and how they differ from Police recorded data, would aid country validation. WHO were in favour of the proposal to clearly identify whether data have been validated and propose ensuring the use of the phrase “SDGs” in the subject line to help facilitate email requests for data validation. WHO also suggest involving the person responsible for validating data earlier in the process to ensure understanding of why WHO use estimated, and not reported, data. WHO do not support the proposal that IAEG-SDGs act as a referee to resolve disagreements between countries and custodian agencies.

148. WHO could not confirm whether they would be developing SDMX capability and do not wish to take part in a further pilot to test data flows using country National Reporting Platforms.

Indicator 9.1.2, passenger and freight volumes, by mode of transport

149. ICAO²² replied with information from its perspective regarding indicator 9.1.2. ITF/OECD did not reply to the main questions raised by the countries involved in that pilot study. The task team had identified that the metadata available are not in the format proposed by IAEG-SDGs, and that some information is missing, or not clear. ICAO was asked if metadata would be revised and ICAO stated they would not.

150. Although most of the countries who responded said they could not validate their data in the global database, the custodian agency, ICAO, confirmed that most of their data is reported data, and very little data are modelled. They also reported that they use country data, and that

²¹ This national data coordinator is appointed through a formal process by the national government. Their responsibility is to validate data before final publication. But most of the time, the national SDGs focal point does not know this contact in his country and is not aware of the data validation process.

²² Only ICAO replied and commented their data. No reply from ITF for which the main concerns come with their data because information failed to get them.

between 91 and 100 percent of countries validate the data. They were confident that coverage and quality of reported data are excellent.

151. The custodian agencies were supportive of the task team's proposals to improve communications through methods such as ensuring all email correspondence contains "SDGs" at the start of the subject, or that countries set up functional mailboxes for SDG work.

152. The agencies would like to automate the exchange of data or metadata for some or all the components of this indicator.

4. Findings

153. The responses from countries and custodian agencies show that communication regarding SDG data flows remains a challenge. Although WHO publishes the name of their country contacts in the annex of the annual report on road accidents, in some cases these names cannot be identified within the country.

154. The national data coordinators for these indicators are appointed through a formal process by the national government, with the responsibility to validate data before final publication. The agencies could inform the National SDGs focal point of their pre-established contact and data provider to facilitate coordination within the country.

155. ITF/OECD also publishes data, but the countries were never invited to validate them. Some data in the global SDGs database are different from the official statistics produced by the ITF themselves (data which come from a common questionnaire shared with UNECE and Eurostat).

156. After the survey was closed, ITF-OECD provided the main reasons why data provided to the SDGs are different from ITF country data. The data ITF collects come from questionnaires sent to countries. But these data contains limitations that do not serve the purpose of the 9.1.2 SDG indicator, i.e. urban transport is not included, road transport is collected for EU countries on the nationality principles (not the territoriality) and ITF data covers ITF members countries only, not the whole world. For these reasons, ITF used data from the ITF freight and passenger transport models. But ITF also confirms they will try in the next SDGs reports to use as much as possible "real" data, coming from questionnaires countries supplied to both the ITF and the UNECE.

157. There also seems to be a need for greater clarity around metadata and the differences between published data in the global databases and country data. The fact that some countries were not able to identify the cause of differences between the two sets of data suggests that information on methods and sources used are not sufficiently transparent. The metadata currently describe how the aviation data are produced, but for road and rail much of the metadata description details the ITF's model and does not describe statistics production. A reference to the Eurostat/ITF/UNECE Glossary for Transport Statistics could also help improve understanding, as would adding country-specific metadata that already exists in the ITF database.

158. With regards to the indicator itself, it is also worth noting that its wording in the official list of SDG indicators is "Passenger and freight volumes, by mode of transport". Therefore, Custodian agencies might be questioned why inland water transport is not included in freight

(a significant proportion of many countries' tonne-km), especially since these data are available through the ITF's statistics database. Similarly, pipeline transport should also be considered in freight transport. Equally, it might be analytically useful to split road passenger transport and private transport.

5. Post survey agencies' development

159. For the 2019 edition of the SDGs report, the custodian agencies have significantly enhanced their internal coordination, and also coordinated with UNECE, ITF-OECD and UNCTAD. For the first time thanks to UNCTAD, data on the maritime transport have been provided in the 2019 report in addition to other modes. Also in collaboration with UNECE and ITF-OECD, the ITF modelled / calculated rail, road country data was replaced with the actual reported data from States so that the variances between the State data and the agency provided data is minimal. ITF also confirms that inland waterway and pipeline transport could be added to the data transmission. In addition UNECE has updated the metadata information on how States report the transport data. This too has been provided by ICAO to the UN for their updates to the metadata repository.

IV. SELF-ASSESSMENT BY COUNTRIES AND CUSTODIAN AGENCIES

160. The second part of the survey asked responding countries and agencies to provide their own assessment of what worked well, what challenges remain, and some suggestions to strengthen data flows for SDGs. A description of the approach, suggestions received from countries and agencies, and overall findings are presented below.

A. GENERAL DESCRIPTION OF APPROACH

161. Firstly, countries and custodian agencies were asked to reflect on their experiences interacting with each other to produce harmonized statistics for a small, yet diverse, set of indicators. The Task Team was interested in understanding specific roles of each in the review process, with a view to improving communication and thereby transparency and efficiency of data flows. The Task Team was also interested in the use of tools developed since the time of the first data flow pilot, such as the UNSD indicator focal point dashboard, and how these could be improved.

162. Another interest was the specific ways data were transmitted for selected SDG indicators, and to gauge current use and commitment to use SDMX more broadly for this purpose. Moreover, suggestions were sought from both parties on how to improve the current validation process, make better use of the UNSD database, and improve national coordination in the scope of specific SDG indicators. Among 40 participating countries, 32 responses were received on the self-assessment part of the survey.

B. SUGGESTIONS FROM COUNTRIES

163. Countries were asked about what aspects of their experience with the SDG data flow process worked well and what did not. Fifteen countries gave positive comments on the current process and provided details in this regard. Many of them acknowledged improvements in their own national coordination. Many also noted that national focal points were better identified, better informed, and had more clearly defined and transparent responsibilities. Some countries noted that, in some cases countries suggested more recent data to custodian agencies, and these were agreed upon with agencies and incorporated through the reporting process.

164. Countries also noted some aspects of the current data flow process that remained challenging. Eleven countries identified specific issues that are yet to be resolved. A key underlying issue was **lack of sufficient information provided to NSOs**. Specific examples include:

- when custodian agencies contact NSOs but put other national contacts on blind copy;
- when custodian agencies send requests to a national contact that is not the national focal point, and do not copy the national focal point; and
- when custodian agencies do not describe a process by which countries are invited to validate harmonized data before publication, or if this process is not conveyed to the national focal point.

165. Although existing data flows should be leveraged as much as possible, the way these flows have operated in the past should be adjusted to meet the coordination needs for SDGs. It will not be sufficient to rely on “business as usual.” To address this, **custodian agencies should always copy national focal points when communicating with national contacts, especially**

for extant data flows. This allows NSO to become more familiar with extant data flows and improves communication and coordination overall.

166. Countries were also asked if they were using the agency focal points list provided by UNSD, and whether it was sufficient or not. Twenty-eight of the countries answered, and about half of these explained that they had not known about this contact list earlier or had not used it yet but planned to do so.

167. Seventeen countries indicated that they used the UNSD focal points list and many found it very useful. Of these, 8 countries noted that the list was somehow insufficient and needed improvement. Specifically,

- agency focal points and their contact information were incomplete;
- national focal points were not identified (which could help improve national coordination and awareness of existing data flows); and
- the data publication schedules were not clear on the calendar.

168. Countries were also asked for feedback regarding indicator metadata availability, accessibility and quality: Many countries noted that metadata were available and accessible, and expressed gratitude for the recent improvements in this regard. Twenty-four countries commented favourably on at least one of these three desirable features; and all respondents found the metadata at least partially available and accessible. However, about half of responding countries indicated that metadata quality was not standard, and improvements were needed to address missing information.

169. Countries were asked for feedback regarding how validation status should be described and how and where it should be noted. Of the 31 respondents, 29 supported the validation categories as described in the survey (or similar) for inclusion in the UNSD global database. Specifically, in cases for which country validation might be difficult to achieve, the proposed alternative could be that a note is added indicating that the data are modelled by the custodian agency and the status of validation by the country (e.g., validated by country/country can't validate/not validated by country/pending review) is included. Thus, all globally harmonized national data and statistics could be published in the UNSD global database. This could be recommended to the UNSC as a modification to the 2018 UNSC data flow guidance. Countries also agreed (essentially unanimously) that validation status should be included in SDMX updates for SDG use.

170. Countries were also asked whether validation status should be simplified. There were many different views on this and what the categories could be. Thus, as one country offered, it may be preferable “to wait for gathering enough experience related to data flows and validation” before identifying simplified categories.

171. Most responding countries were very interested in using SDMX for SDG purposes, but are not yet doing so. Of 32 respondents, 5 stated that they are already capable of SDMX transmission for SDG purposes, and they will integrate this into their SDG data flows. Of respondents not yet capable, 10 countries are planning to use SDMX for SDGs indicators in the coming years and 4 are ready to organize SDGs data flows with SDMX. Eight countries confirm that SDMX adjustments are planned after the relevant guidance is provided. Most countries also supported the use of automated “push/pull” data transmissions.

172. Countries were asked if they were willing to incorporate Custodian Agency estimations or calculations in their national reporting platform. Twenty-one of the twenty-six respondent countries found this possibility favourable, but often with a condition that they can validate the agency's calculation or otherwise affirm the data/calculation as sufficient. A few countries were undecided or indicated that only data that are in line with their national SDG Strategy would be posted in NRPs.

C. SUGGESTIONS FROM AGENCIES

173. The agencies participating in the pilot were asked for their views on issues to be addressed and suggestions to improve the transparency and efficiency of SDG data flows. As with country participants, these questions focused on communication with focal points, validation and SDMX related issues.

174. Because agencies were invited to participate only if an indicator they are custodian agency for was selected for the second data pilot, only 7 agencies were eligible to respond. Of these, 4 agencies responded to the survey and only 2 provided feedback on the self-assessment section.

175. The two agencies responded that they appreciate the efforts of countries in the data collection process for their indicators. Both described the need for greater national coordination and communication. One agency agreed with the suggestion to provide more complete contact information for indicator focal points. One agency supported the validation status categories suggested for UNSD global database but did not support simplifying the validation status categories further.

176. Both responding agencies are eager to the make use of SDMX for SDGs as well as automation of data transmissions. One agency volunteered to participate in an SDMX pilot activity, if available.

D. FINDINGS

177. There was broad agreement for the issues identified. These can be described in two groups: metadata quality and focal point coordination. Metadata quality, even for Tier I indicators, was uneven and sometimes incomplete. Updates to metadata were not clearly denoted or communicated to national focal points. Focal point information is also uneven, at times incomplete. Coordination should be improved at both national and agency levels.

178. There also seemed to be broad agreement for some specific, practical, low effort solutions to these issues. Each of the actors involved could contribute to their implementation.

179. Possible solutions that could be implemented by **NSOs (with UNSD)** include:

- National country focal point information should be added to the UNSD global database as a resource for agency focal points (at the country's discretion) and regularly updated.
- National contacts should be coordinated at the national levels by NSOs and consistent with the mandate received by the custodian agencies from country governments. Such coordination is also essential for reporting non-statistical indicators.
- Lists of national country focal points should include any other national contacts for extant data flows related to SDG indicators in the UNSD global database.
- NSOs should inform agencies how they want to receive data requests from agencies (i.e., from the focal point, or with the focal point in copy).

180. Possible solutions that could be implemented by **IAEG-SDG (with UNSD)** include:

- IAEG-SDGs does not have the governance mechanism or mandate to review metadata for Tier I and II indicators. Countries felt this is a clear need.
- Some indicators have more than one metadata file. This can be confusing. The files should be integrated by custodian agencies.
- Major changes in metadata files should be reviewed by IAEG-SDGs. IAEG-SDGs should receive feedback from countries on metadata and follow up with custodian agencies as needed.
- It would be helpful to denote in the UNSD global database and/or official indicator list which global indicators are non-statistical in nature, and those that apply to a subset of countries only (for example, LDCs, SIDS, LLCs, DAC members) and those which rely on a common model.
- In cases where country validation is difficult, a note is added indicating that the data are modelled by the custodian agency and the status of validation by the country (validated by country/country cannot validate/not validated by country/pending review) is included. Thus, all globally harmonized national data and statistics are published in the UNSD database. This could be recommended as update to the 2018 UNSC data flow guidance. (Near unanimous opinion.)

181. Possible solutions that could be implemented by **custodian agencies (with UNSD)** include:

- The data sources used to calculate harmonized statistics should be referenced in the UNSD global database, including data taken from other databases. (Unanimous opinion of respondents.)
- The UNSD calendar should indicate agencies' data collection and validation time frames. When exact dates are unknown, tentative dates should be used so NSO work programs can plan ahead.
- Agency contact information should be updated/completed for indicators.
- The following indicators seem to have missing metadata or require clarification, based on the data flow pilot: 9.1.2, 17.1.2, and 17.3.1.
- Work plans for some Tier III indicators are outdated and need to be updated.

182. Possible solutions that could be implemented by **UNSD (with focal points)** include:

- A dashboard is needed to support communication between countries and agencies. This dashboard should integrate contact information for country and agency focal points and data collection schedules. Information to the dashboard should be provide by both countries and agencies, be up to date, and easily accessible/shared.
- Any updates made in the metadata files should allow notification and tracking of these updates by national focal points. This will help NSOs respond to agency requests in an efficient manner.
- The validation status should be displayed in the UNSD global database, and updates (such as new postings by custodian agencies or revisions) should be communicated to national and agency focal points. For example, such changes could generate an automatic email to focal points. (Unanimous opinion)

183. Looking ahead, there was broad support from countries and agencies to apply SDMX to their SDG data flows. Providing non-technical guidance to NSOs and agencies would assist in the planning process. Providing more opportunities for countries and agencies to pilot SDMX

and its SDG enhancements would also be welcome. It is also anticipated that NRPs may increasingly feature both national statistics and statistics generated by custodian agencies (if validated or otherwise affirmed by the country as appropriate for its national platform).

ANNEXES

ANNEX 1. COUNTRY AND AGENCY SURVEYS FOR THE SECOND DATA FLOW PILOT

The instruments used for the second data flow pilot can be found on the UNECE public wiki (<https://statswiki.unece.org/display/SFSDG/Task+Team+on+Data+Flows+for+SDGs>).

Country Questionnaire (PDF | DOC | Survey Monkey)

Agency Questionnaires

1. FAO (PDF | DOC | Survey Monkey)
2. WHO (PDF | DOC | Survey Monkey)
3. ICAO (PDF | DOC | Survey Monkey)
4. UN Women (PDF | DOC | Survey Monkey)
5. OECD (PDF | DOC | Survey Monkey)

ANNEX 2. COUNTRY AND AGENCY PARTICIPANTS TO PILOT FLOW SURVEY

	Participant	Name of Contacts	Email of Contacts
1	Albania	Alban Cela	Acela@instat.gov.al
2	Argentina	Martín Chojo Hernan Munoz	mchojo@indec.gob.ar hmunoz@indec.gob.ar
3	Armenia	Anahit Safyan	info@armstat.am ; safyan@armstat.am;
4	Australia	Douglas Jacobs	douglas.jacobs@abs.gov.au
5	Azerbaijan	Sabina A. Guliyeva	Sabinag@azstat.org
6	Belarus		sdg@belstat.gov.by
7	Belgium	Frank Verschaeren	Frank.Verschaeren@economie.fgov.be
8	Bosnia and Herzegovina	Ševala Korajčević	sevala.korajcevic@bhas.gov.ba
9	Brazil	Wanda Rodrigues Coelho	wanda.coelho@ibge.gov.br
10	Chile	Paola García Lapaquette Olga Barquero Alpizar	paola.garcia@ine.cl olga.barquero@ine.cl
11	Czech Republic	Eliška Valinová Miloslava Veselá	eliska.valinova@czso.cz cz_miloslava.vesela@czso.cz
12	Denmark	Maciej Truszczynski	MTR@dst.dk
13	Estonia	Kaia Oras	kaia.oras@stat.ee
14	Finland	Ari Tyrkkö	ari.tyrkko@stat.fi
15	France	Functional mailbox Sylvie Eghbal Claire Plateau	coordination-statistique@insee.fr Sylvie.Eghbal@insee.fr claire.plateau@insee.fr
16	Germany	Miriam Blumers Functional mailbox	Miriam.Blumers@destatis.de sdg-indicators@destatis.de
17	Ireland	Kevin McCormack	Kevin.McCormack@cso.ie
18	Israel	Amit Yagur-Kroll Yoel Finkel Sigalit Mazeh	amitk@cbs.gov.il yoel@cbs.gov.il sigalit@cbs.gov.il
19	Italy	Angela Ferruzza	ferruzza@istat.it
20	Japan	Mari Watanabe	dgpp_ss_intl@soumu.go.jp
21	Kazakhstan	Ainur Dossanova	ai.dosanova@economy.gov.kz
22	Latvia	Anna Velika	Anna.Velika@csb.gov.lv
23	Liechtenstein		[Partial response]
24	Lithuania	Audronė Miskiniene	Audrone.Miskiniene@stat.gov.lt

	Participant	Name of Contacts	Email of Contacts
25	Macedonia, the Former Yugoslav Republic of	Snezana Sipovic	snezana.sipovic@stat.gov.mk
26	Mexico	Pilar García Velázquez Manuel Cuellar Rio	pilar.garcia@inegi.org.mx MANUEL.CUELLAR@inegi.org.mx
27	Mongolia	Ms. D. Batchuluun	international@nso.mn
28	Netherlands	Lieneke Hoeksma	l.hoeksma@cbs.nl
29	Norway	Live Rognerud	Live.Rognerud@ssb.no
30	Poland	Mateusz Wyżykowski Magdalena Zero	m.wyzykowski@stat.gov.pl m.zero@stat.gov.pl
31	Portugal	Conceição VEIGA	conceicao.veiga@ine.pt
32	Russian Federation	Natalia Ignatova	Ignatova@gks.ru
33	Slovak Republic	Claudia Meszarosova Kleinova	Claudia.Kleinova@statistics.sk
34	Sweden	Sara Frankl	sara.frankl@scb.se
35	Turkey	Functional mailbox Övünç Uysal	sdg@tuik.gov.tr OVUNC.UYSAL@tuik.gov.tr
36	Ukraine	Inna Petrichenko	I.Petrichenko@ukrstat.gov.ua ; K.Babak@ukrstat.gov.ua
37	UK	Joanne Evans Functional Mailbox Claudia Wells	joanne.evans@ons.gov.uk sustainabledevelopment@ons.gov.uk Claudia.Wells@ons.gov.uk .
38	USA	Kali Kong	kkong@omb.eop.gov
1	FAO	Jippe Hoogeveen Rosa Laura Romeo	Jippe.Hoogeveen@fao.org RosaLaura.Romeo@fao.org
2	ICAO	Antonin Combes	ACombes@icao.int
3	OECD	Simon Scott	Simon.SCOTT@oecd.org
4	UN Women	Janette Amer Tanya Primiani Lea Fuiet	janette.amer@unwomen.org tprimiani@ifc.org Lea.Fuiet@oecd.org
5	WHO	Emilia Aragon	N/A (no response)

ANNEX 3. REFERENCES

[2017 Pilot Results](https://statswiki.unece.org/download/attachments/128451079/2017-Data-Flow-Report.pdf?version=1&modificationDate=1560760423209&api=v2) <https://statswiki.unece.org/download/attachments/128451079/2017-Data-Flow-Report.pdf?version=1&modificationDate=1560760423209&api=v2>

[African Region Comments on the Guidelines on Data Flows and Global Data Reporting for Sustainable Development Goals, prepared by the United Economic Commission for Africa, 2018](https://unstats.un.org/sdgs/files/meetings/iaeg-sdgs-meeting-08/BEST%20PRACTICE%20FOR%20DATA%20FLOWS%20IN%20AFRICA%20-%20SEPTEMBER%202018.pdf) <https://unstats.un.org/sdgs/files/meetings/iaeg-sdgs-meeting-08/BEST%20PRACTICE%20FOR%20DATA%20FLOWS%20IN%20AFRICA%20-%20SEPTEMBER%202018.pdf>

[Concept Note for the 2017 Data Flow Pilot](https://statswiki.unece.org/download/attachments/128451079/Pilot%20Concept%20Note_2017%20final.doc?version=1&modificationDate=1518203548805&api=v2) (July 2017 - September 2017)
https://statswiki.unece.org/download/attachments/128451079/Pilot%20Concept%20Note_2017%20final.doc?version=1&modificationDate=1518203548805&api=v2

24 August 2018

The initial observations and recommendations below are the result of a second data flow pilot study conducted by the Task Team on Data Flows under the auspices of the UNECE Steering Group on SDG Statistics, which was carried out during July-August 2018. More information regarding the purpose, scope, instruments and findings of the pilot can be found at <https://statswiki.unece.org/display/SFSDG/Task+Team+on+Data+Flows+for+SDGs>.

Finding one

1. Most countries are strongly invested in the transmission of their data to the custodian agencies as shown by the strong participation in this survey on a voluntary basis and well beyond the geographical scope of UNECE.²³ Most countries strive to produce national statistics that are aligned with the UN global indicator metadata, while recognizing that international harmonization of the statistics will be challenging.
2. Most agencies involved in the survey by the selection of indicators ensured on-going dialogue with the task team to provide clarification on their process of collecting national data for global monitoring of SDGs in order to maximize quality, transparency and trust in the SDG global reporting and adherence to the Fundamental Principles of Official Statistics.
3. The following recommendations are intended to contribute to this common aim.

Finding two

4. Most countries noted with appreciation the on-going efforts made by the UNSD to facilitate SDG data flows between agencies and countries and remarked on the continued improvements. They noted with appreciation the release of a dashboard listing contact focal points; and the posting of SDG metadata and global statistics on the UN website's UNSD database. They would like this work to continue and to go further.
5. Regarding information about focal points, custodian agencies regret that some countries have not yet transmitted their national focal point contact information, which has, impeded follow up to resolve data flow issues. In other cases, countries regret that the custodian agencies' focal point contact information is insufficient. For instance, for indicator 9.1.2, a name is provided without an e mail, phone number or physical address.
6. Countries also note that the calendars of agencies' request and of data collection are also too incomplete (and difficult to find on the UN website), limiting the current use of this tool for national data flow coordination.

Recommendation two

7. UNSD could post an online dashboard of focal point contact information for countries and custodian agencies, regularly updated with precise contact information for each indicator and the calendar of the data and validation request which should be agreed by IAEG-SDGs.

²³ As of 24 August, 35 countries and 4 custodian agencies participated in the pilot. Of these, 28 were UNECE countries and 7 were outside the region).

8. Taking into account concerns for confidentiality, this dashboard should be password protected. It should be updated through secure login by the central focal points at each country and custodian agency. Countries would be responsible for maintaining the list of country focal points per indicator for their countries, and custodian agencies would be responsible for maintaining the list of custodian agency contacts per indicator and their calendar of data request and data validation, allowing for planning coordination on the national level. Reminders for updates could be timed with preparations for the SDG annual report to the Secretary General.

9. This dashboard should become a reference document between agencies and countries.

Finding three

10. SDG data transmission doesn't start from scratch in all cases. In fact, some data transmission are already well established with their own focal points and their own national mandate given by the Foreign ministry or another national line ministries. (For instance, the data transmission for indicator 17.1.2 is DAC members' obligation to OECD with their own process of data validation; similar existing data flows are in place for indicator 3.6.1 reporting on rate of death rate due to road traffic accidents and indicator 9.1.2 passengers and freight volume by mode of transport).

11. NSOs have sometimes mapped SDGs indicators with their national statistical system identifying a new data provider and a national focal point, not being aware of these existing data flows. Discrepancies might appear with new data sources and be source of conflict with agencies.

12. In some cases, custodian agencies directed their SDG data request to the previously established focal point without (also) informing the SDGs focal point.²⁴ In several cases, national SDG focal points are not aware of these existing data collections by agencies or recent requests from agencies for some indicators.

Recommendation three

13. To avoid such conflicts, national coordination within countries should be promoted. This will require some support from custodian agencies.

14. In particular, the SDG focal point nominated by the country should be systematically informed by the custodian agencies of existing national focal points for indicators and existing validation processes, where these have been established.

15. To accomplish this, custodian agencies are encouraged to put in copy (not hidden copy) the national SDGs focal point whenever they address any SDG data request to a previously established national data provider, including requests concerning non-statistical indicators. This recommendation applies even if the Custodian agency's request is broader in scope than the SDG indicator request alone (for instance, collection of data on education, of which some SDG indicators are included).

²⁴ This should be indicated in the UN dashboard of focal points.

Finding four

16. The quality of metadata for indicators assigned as Tier 1 or Tier 2 is unequal. While most of that metadata are well structured and have good quality concepts, some metadata are incomplete,²⁵ difficult to understand, and feature open questions and inconsistencies. For instance, metadata for indicator 17.3.1 are incomplete. The definition is limited to the ODA aspect of the indicator. Overall, it seems very similar to the metadata provided for 17.2.1, although it is intended, by its tier designation, to be conceptually clear and to have an internationally established methodology. The same can be said of indicators 6.4.1 and 6.4.2, for which there are many inconsistencies and open questions remain. Also, metadata for 15.4.2 does not take into account, that some countries experience greening of mountain areas due to climate change.

17. Metadata for Tier 1 and Tier 2 indicators are not reviewed by IAEG-SDGs; only Tier 3 indicators are reviewed by IAEG-SDG to determine possible reassignment to Tier 2 or Tier 1 as concepts are clarified and agreed upon methodology is established.

Recommendation four

18. All metadata of Tier 1 and 2 indicators should be systematically reviewed by IAEG-SDGs to ensure that they conform to the quality and the format recommended by IAEG-SDGs. Among these, priority of review should be given to those indicators that are subject to country signature.

19. Furthermore, the task team would like to encourage IAEG-SDGs to review certain indicator's tier classification. It is not evident, for example, why the indicator 17.3.1 is classified as tier 1 even though it is not produced and the custodian agency has not proposed a method for its calculation.

20. A dedicated venue should be allotted for countries (or agencies) to raise their questions on the interpretation of metadata and difficulties they have observed. Custodian agencies should be invited to answer the questions raised and to update accordingly the metadata in track changes (with date of update), which will allow countries to avoid unnecessary review.

Finding five

21. In several cases, NSOs mention that they were not asked to validate of some data associated with their country that is published in the UN global database, and, in some cases, they disagree with the data published.

Recommendation five

22. The national focal point for SDGs should always be informed prior to the publication of data on the UN SDGs global database, even if these data have been validated by another process. It also gives the opportunity for a last check and validation of these national data.

Finding six

23. Some SDGs indicators could be collected with existing reporting mechanisms (for instance 6.4.1 and 6.4.2 Aquastat indicators), with some adaptation or extension of the current process.

²⁵ Or, even lacking altogether (for instance, see indicator 3.3.4).

Recommendation six

24. Agencies are encouraged to minimize the data reporting burden of Member States by utilizing existing reporting mechanisms and adapting them. They are encouraged to work with other international agencies working on these existing data flows to adapt them as appropriate and avoid duplicative reporting and unnecessary burden²⁶.

Finding seven

25. Data validation is sometimes complex, especially when metadata require the indicator to be modelled by agencies or calculated by agencies directly with open data sources (such as geospatial information).

Recommendation seven

26. Data validation requires sufficient time for countries and agencies to conduct their tasks professionally, and full transparency on the status of data which are released in the global database. Accordingly, a transparent validation process that allows maturation of review and agreement should be implemented.

27. Essentially, this is a documentation process. Systematically, the source of data should remain indicated in the UN global data base (i.e., country data, data estimated by the agency) and the status of validation (i.e., validated by country, cannot be validated by country (missing data), pending review) should be included in the data regardless of whether the statistics is published.

28. This modification to current procedures will make the global process more fully transparent. All data used to calculate the aggregate will be noted. Further, reporting on SDGs will be facilitated and potentially take greater advantage of other available data sources that meet acceptable standards of statistical quality. This can also help communicate clearly the responsibility of NSOs in providing national data. Moreover, the progress made in increasing the number of indicators validated by the focal points nominated by the national SDGs coordinator in itself is a measure of progress in country-led monitoring of SDGs.

²⁶ Reviewing the report, FAO comments that the existing reporting mechanism (i.e. OECD/EUROSTAT or UNSD data collections) are largely insufficient for producing the necessary information to monitor indicators 6.4.1 and 6.4.2. Moreover OECD/EUROSTAT or UNSD do not have any plan/responsibility to expand their current questionnaire to start reporting on indicators 6.4.1 and 6.4.2. FAO is the custodian agency and has the responsibility to report on indicators 6.4.1 and 6.4.2. According to the “Criteria for the implementation of the guidelines on data flows and global data reporting for the Sustainable Development Goals”, adopted by the UN Statistical commission in March 2019, “International and supranational statistical agencies will only address data requests to countries regarding a specific Sustainable Development Goal indicator if they are the designated custodian agency. (Paragraph 12)”