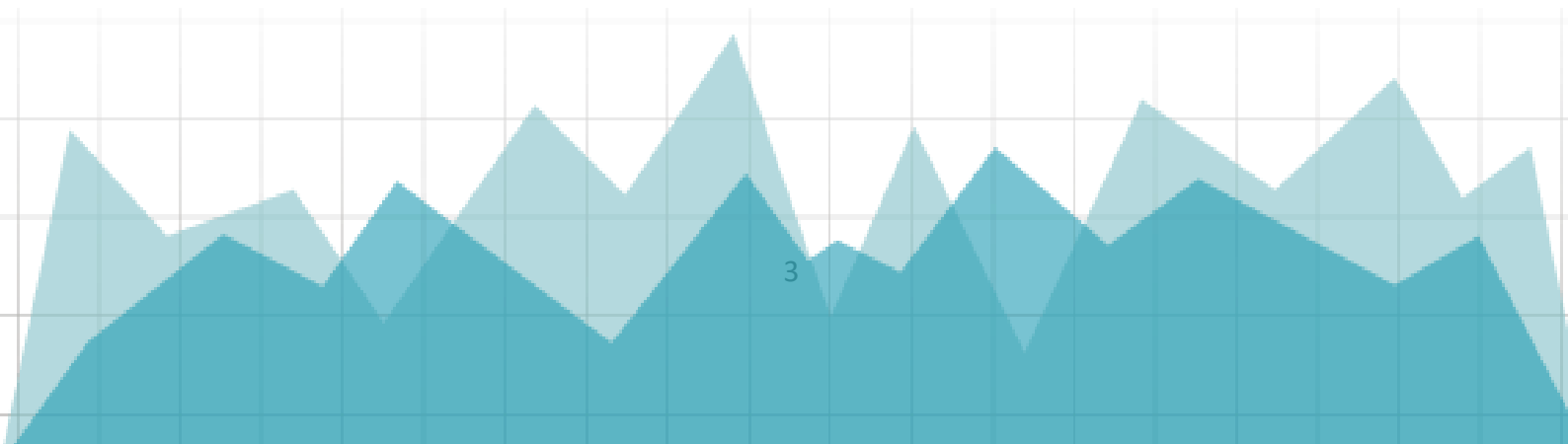
Food and Agriculture Organization of the United Nations

Statistical Standard Series

Quality Indicators for External Users
Version 3.0

Endorsed by the Technical Task Force of the Data Coordination Group
3 November 2023



This standard provides the list of recommended quality indicators to be disseminated to FAO external data users. It includes key definitions, guidance on how to compile the indicators, general and technical recommendations on their compilation and dissemination as well as some governance considerations.

The version 3.0 of the document was endorsed by the Technical Task Force of the Data Coordination Group (DCG-T) on 3 November 2023.

CONTENTS

Quality indicators for external users

BACKGROUND

FAO is committed to produce and disseminate high quality statistics, where quality of statistical outputs is measured in terms of (i) relevance; (ii) accuracy and reliability; (iii) timeliness and punctuality; (iv) coherence and comparability; and (v) accessibility and clarity, as indicated in the FAO Statistics and Data Quality Assurance Framework (SDQAF)[1]. Measuring the quality of FAO statistical outputs is crucial to establish a roadmap for quality improvements. In addition, disseminating information on quality indicators facilitates a correct interpretation and use of the statistical outputs, contributing directly to improve their accessibility and clarity.

However, measuring quality is not straightforward. For some quality dimensions, like timeliness, a quantitative assessment can be easily obtained, while for other dimensions, this type of assessment may encounter several methodological difficulties and require non-negligible efforts. For instance, assessing the relevance of a statistical output could involve carrying out a users' satisfaction survey. Accuracy itself, which implies the quantitative assessment of how close an estimate is to the corresponding true (unknown) value, would require the estimation of the *Mean Square Error* (MSE), which is often unfeasible. For these reasons, the common approach to the measurement of the quality of statistical outputs consists in the compilation of a series of indicators that are directly or indirectly related to the quality dimensions. The indicators only indirectly measuring statistics quality, usually focus on the critical phases of the production process of the statistical outputs (performance indicators), on the assumption that the less errors are done in the whole process, the higher is the quality of the final statistical outputs.

This document provides a list of recommended quality indicators intended to be disseminated to external users, jointly with the statistical outputs. The quality indicators are presented according to the dimensions of quality they refer to.

[1] FAO (2023). The FAO Statistics and Data Quality Assurance Framework. Available at: https://www.fao.org/3/cc6683en/cc6683en.pdf.

DEFINITIONS

Relevance

Relevance is the degree to which the statistics produced meet current and potential users' needs (see FAO SDQAF). For a given statistical production process or, more generally, statistical domain, an indicator of relevance is **the average proportion of satisfied users,** estimated from the corresponding users' satisfaction survey. As reported in the Statistical Standard on user consultation[2], the indicator is derived by adding the responses "Strongly Agree" or "Agree" of the five-items Likert scale (see Annex 1) for each of the SDQAF Principles related to quality of statistical outputs (P12-P17) and dividing this value by the total number of responses. Details on the calculation are reported in Annex 1.

A component of relevance is *completeness*, defined as the degree to which all statistical outputs needed by the users are available. An indicator of completeness for an international organization like FAO is the **geographical completeness**, i.e. the percentage of countries with relevant data (see Annex 1) in a given statistical domain. This is also an indicator related to non-observation errors, which are commonly assessed under the accuracy dimension.

Accuracy and reliability

Accuracy is the closeness of an estimate to the true value of what is measured. In this case, we are not interested in the accuracy of the incoming country data, but rather in assessing solely FAO's contribution to the overall accuracy of the final statistical outputs calculated and disseminated by FAO (usually regional or global aggregates). This component of the overall accuracy of the statistics published by FAO will depend on the errors that may happen in the FAO statistical production process: the higher the number/size of errors, the lower the overall accuracy. The best strategy to improve accuracy is to prevent errors from happening and correct the ones discovered (before the dissemination of the final statistical outputs).

The main errors that can affect the FAO statistical production processes are:

- **Non-observation errors**: a data provider does not report to FAO part or any of the required data.
- **Measurement errors** (in a broad sense): the final value for a given variable does not correspond to the one disseminated at the national level because of errors in the FAO statistical process (in the data collection, data treatment and data processing). The most frequent reasons for this to happen may be: wrong instructions or definitions provided in the questionnaire; classification/coding errors in the collection or treatment of incoming data; errors in converting the unit of measure during the data collection phase or in the subsequent treatment of incoming data; errors in the data validation phase (a non-erroneous value is identified as suspicious and modified when data are checked by FAO).
- **Estimation errors**: errors introduced by FAO during the calculation of the final aggregates. Typically, they correspond to errors in the software codes or in the model/assumptions

[2] See Statistical Standard on User Consultations at https://www.fao.org/3/cb9340en/cb9340en.pdf

underlying the aggregation procedure of the collected data needed to derive complex indicators (e.g. use of a simple average instead of a weighted average, etc.).

Non-observation errors are critical, since just a small fraction of units not reporting their data may hinder the calculation of regional/global aggregates[3]. The preferred solution to overcome this problem consists in imputing[4] the missing values and then calculating the final aggregates using imputed values as if they were really observed (jointly with the actually observed values). The alternative option of calculating the final statistical outputs by processing solely the observed values (i.e. discarding missing values) may need ad hoc assumptions and/or methods; for instance, the bias (underestimation) introduced by calculating the total amount of a variable by summing up just the observed values can be considered negligible if the missing values are assumed to be very small compared to the observed ones[5]. In some cases, mixed strategies can be adopted (e.g. imputation only of the most influential non-reported values).

When the data collection involves the dispatch of questionnaires, a first indicator related to non-observation error is the **questionnaire reporting rate**, i.e. the number of returned filled-in questionnaires (with complete or partial valid information) divided by the number of dispatched questionnaires. Since, in some cases, the official country data can be collected without using a questionnaire (e.g. official country data collected from regional/international organizations), the questionnaire reporting rate should come along with the **percentage of observed data**, i.e. the number of observed data points compared to the entire set of data points to be used for calculating the final statistical outputs (see Annex 1 for the calculation details).

When imputation is used to compensate for non-observation errors or for erroneous reported values (typically values identified as errors are deleted and replaced with valid plausible imputed values), the **percentage of imputed data**, i.e. the ratio of the imputed data over the total number of data items to be used for calculating the final statistical output, should be calculated (see Annex 1 for the calculation details).

When the final statistics are obtained as a sum of values (total amount), an important quality indicator to be calculated is the **contribution of the imputed values to the final sums** (for calculation details see Annex 1). This indicator may also prove useful when the disseminated outputs are obtained as a function of one or more sums (e.g. ratio of sums).

The values of the quality indicators related to imputation should be provided together with summary information on the methods applied to perform the imputation procedure.

Reliability indicates how close the initial estimates are to the subsequent or final estimates. Assessing reliability is preferred to measuring accuracy when dealing with a complex statistical process that uses multiple data sources being updated at different times, possibly, in some cases, with only provisional data. In this context, a common practice is to produce provisional estimates that are subsequently

[3] See Statistical Standard on Data Aggregation at https://www.fao.org/3/cc2918en/cc2918en.pdf

[4] The term Imputation is used here in its broader meaning, including also the replacement of missing values with values from nonofficial sources, historical data, etc. For further details, see the corresponding Statistical Standard on Imputation at https://www.fao.org/3/cb9339en/cb9339en.pdf

[5] For example, the sum of the missing values contributes to 1% or 2% of the total amount.

revised[6] when the underlying data are updated. Revised estimates should be accompanied by at least two revision indicators: the **mean revision** (MR) and the **relative mean absolute revision** (RMAR). The MR provides an indication of the direction of the revision, while the RMAR is used to assess the relative size of the revisions (calculation details are reported in Annex 1).

Timeliness and Punctuality

Timeliness is the lapse of time between the end of the reference period (or the reference date) and the dissemination of the statistical outputs. The **overall time lapse** can be split in two parts: the timeliness of respondents in submitting the data requested by FAO and the time required for FAO to produce and disseminate the statistical output. Timeliness should be measured in months (days for monthly data): the **timeliness of the "data provider"** is calculated as the number of months (days) between the reference date of the statistical output and the last day of the FAO data collection period. The **timeliness of the FAO process** is the number of months (days) between the first day after closing the data collection and the date of dissemination of the FAO statistical outputs (key variables for regional and global aggregates). Details on the calculation of these indicators are reported in Annex 1.

Punctuality refers to the possible time lag (in months) between the actual delivery date of FAO statistical outputs and the target delivery date. In practice, punctuality can be measured only when a target date of dissemination is set by FAO (usually in the official dissemination calendar).

Coherence and comparability

Coherence indicates how adequate the statistical outputs are to be meaningfully combined with other statistics in different ways and for various uses. Generally speaking, the coherence refers to the extent to which statistics on the same phenomenon can be compared or combined. Coherence can be assessed at different levels: (i) in the same statistical domain, when comparing provisional with final estimates of the same aggregate; (ii) across statistical domains, when comparing similar statistics disseminated by different units (domains) within the same agency; and (iii) across agencies, when comparing statistics on the same topic produced by different agencies.
For FAO purposes, it may be worth calculating coherence indicators when the same aggregates are estimated by different units within FAO, or by other organizations beyond FAO. For numerical variables, the **difference between the estimates** of the same aggregate produced by different units or different organizations (sometimes expressed in relative terms, as shown in the Annex 1) should be calculated.

Comparability refers to the extent to which differences in estimates between geographical areas, non-geographical domains, or over time, can be attributed to real differences of the variable being measured. Comparability is a stricter concept than coherence, as it assumes that the statistics being compared are produced by processes sharing the same concepts, definitions, classifications, methods, etc. A quality indicator related to comparability over time is the **number of comparable data points in a time series** since the last structural break. A break in a time series may occur when the definitions, the classifications or the procedures used in the statistical process are changed. When a break in a time series is introduced, then the indicator of comparability over time should be provided together

[6] For more details on recommended practices for data revisions, see the Statistical Standard on Data Revision at https://www.fao.org/3/cb9311en/cb9311en.pdf

with information on the main reasons for the break.

Accessibility and clarity

Accessibility is the set of conditions and modalities that determine how easy it is for the users to obtain data. A commonly indicator of accessibility is the **number of visits of a given web page** or the **number of data downloads**. These indicators can be calculated, for instance, using Google Analytics facilities.

GENERAL RECOMMENDATIONS FOR IMPLEMENTATION

- The data owner should evaluate which of the quality indicators presented in this document are applicable to the statistical process under his/her responsibility. All relevant quality indicators should be calculated and disseminated to external users jointly with the disseminated statistical outputs. A brief text should be provided to explain how to interpret and use the disseminated quality indicators.

- The quality indicators should be calculated at both global and regional levels (when relevant) in accordance to the suggestions provided in the following Table.

Quality dimension	Quality indicator	Geographical level	Disseminated outputs
Relevance	Average proportion of satisfied users (M)		Referred to the key variables/statistical outputs
	Geographical completeness (GC)	Global and Regional	Referred to the key variables/statistical outputs
Accuracy	Questionnaire Reporting Rate (QRR)	Global and Regional	
	Weighted Questionnaire Reporting Rate (WQRR)	Global and Regional	
	Percentage of missing data (PM)	Global and Regional	Referred to the key variables/statistical outputs or to the whole domain
	Percentage of observed data (PO)	Global and Regional	Referred to the key variables/statistical outputs or to the whole domain
	Percentage of Imputed values (PI)	Global and Regional	Referred to the key variables/statistical outputs or to the whole domain
	Contribution of Imputed Values to Totals (CIVT)	Global and Regional	Referred key variable(s) contributing to total(s) or to a function of totals
Reliability	Mean Revision (MR)	Global and Regional	Referred to the key statistical outputs that are revised
	Mean Absolute Revision (MAR) (or RMAR)	Global and Regional	Referred to the key statistical outputs that are revised
Timeliness	Overall Timeliness	Global and regional (if regional are disseminated later)	Referred to the key variables/statistical outputs
	Timeliness of incoming data	Global and regional (if regional are disseminated later)	Referred to the key variables/statistical outputs
	Timeliness of FAO statistical process	Global and regional (if regional are disseminated later)	Referred to the key variables/statistical outputs
Punctuality	Delay in publication (*if a dissemination calendar exists*)	Global and regional (if regional are disseminated later)	Referred to the key variables/statistical outputs
Coherence	Difference or relative difference between the same statistical outputs produced by different units	Global and regional	Referred to the key statistical outputs
Comparability	Number of comparable data items in a time series	Global and regional	Referred to the key statistical outputs
Accessibility	Number of web pages visits		
	Number of data downloads		

- Quality indicators different from those suggested in this document can be provided to users if they are considered relevant for the statistical outputs being disseminated.

TECHNICAL RECOMMENDATIONS FOR IMPLEMENTATION

- Quality indicators should be easily accessible from the web pages where the statistical outputs they refer to are displayed. Electronic publications disseminating statistical data should include an Annex with the pertinent quality indicators.

- The quality indicators should be disseminated along with the **reference metadata**[7]. For complex statistical processes it may be worth preparing a **quality report**, i.e. a summary document providing the most important information about the quality of the process and the corresponding statistical outputs, which obviously should include the relevant quality indicators.

- A contact or a link to additional information about quality indicators (underlying definitions, calculation formula, main use, etc.) should be provided for users interested in getting more insights.

GOVERNANCE PROCEDURES

- All technical units in charge of producing statistics in FAO are accountable for the implementation of this Standard.
- The list of quality indicators may be revised and/or updated in the future (e.g. with the inclusion of additional indicators considered relevant for existing or new statistical production processes).
- This standard on quality indicators may also need to be updated in case of changes to the standard related to the FAO observation status and flags, and evolving best practices.
- Changes to the list of quality indicators may need to be reflected in the standard on metadata dissemination.
- The Statistics Division is responsible for updating this standard and seek the endorsement of the revised/updated list of quality indicators by the Technical task force of the Data Coordination Group.

[7] See the Statistical Standard on Metadata Dissemination for FAO statistical databases at https://www.fao.org/3/cb9292en/cb9292en.pdf.

Annex 1: Details on indicators

Relevance – average proportion of satisfied users

A users' satisfaction survey should allow to derive the following table, where the satisfaction regarding each of the quality dimensions is measured through a five-items Likert Scale.

Response Categories (*five-items Likert scale*)	**FAO statistical outputs are:** Relevant	Accurate/ reliable	Timely/ punctual	Coherent/ comparable	Accessible/ clear	Total
Strongly agree	$n_{SA,1}$	$n_{SA,2}$	$n_{SA,3}$	$n_{SA,4}$	$n_{SA,5}$	n_{SA+}
Agree	$n_{A,1}$	$n_{A,2}$	$n_{A,3}$	$n_{A,4}$	$n_{A,5}$	n_{A+}
Neutral	$n_{N,1}$	$n_{N,2}$	$n_{N,3}$	$n_{N,4}$	$n_{N,5}$	n_{N+}
Disagree	$n_{D,1}$	$n_{D,2}$	$n_{D,3}$	$n_{D,4}$	$n_{D,5}$	n_{D+}
Strongly disagree	$n_{SD,1}$	$n_{SD,2}$	$n_{SD,3}$	$n_{SD,4}$	$n_{SD,5}$	n_{SD+}
Total	n_{+1}	n_{+2}	n_{+3}	n_{+4}	n_{+5}	n_{++}

Calculation formula:

$$M = \frac{n_{SA+} + n_{A+}}{n_{++}} \times 100$$

This expression can be viewed as a weighted average of users' satisfaction by principle, i.e.

$$M = \frac{\sum_{j=1}^{5} n_{+j} p_{+j}}{\sum_{j=1}^{5} n_{+j}}$$

Where $p_{+j} = \left(n_{SA,j} + n_{A,j}\right)/n_{+j}$ is the proportion of users satisfied with principle *j*. Note that in the absence of missing values or "Don't know", then $n_{++} = 5 \times m$, being m the number of users participating to the survey (respondents to the survey).

Relevance – Geographical completeness

Calculation formula:

$$GC = \frac{N_{covered}}{N_{needed}} \times 100$$

$N_{covered}$ is the number of countries with some valid data in the last year, while N_{needed} is the number of countries that had to be observed ("in-scope", i.e. excluding those where the phenomenon is not relevant). These indicators require a clear definition of the in-scope countries and a clear reference to the FAO revised M49 standard list.

Relevance – Weighted Geographical completeness

The geographical coverage can also be calculated by assigning to each country a measure of its "importance" with respect to the studied phenomenon (sometimes only the subset of important countries is intentionally observed):

$$WGC = \frac{\sum_{j=1}^{N_{cov}} w_j}{\sum_{j=1}^{N_{need}} w_j} \times 100$$

Being w_j a numerical value expressing the importance of country *j*. Typical variables used to represent the importance of a country are its GDP (when dealing with economic phenomena), its total population (social phenomena), or its land area (environmental phenomena).
The calculation of WGC is optional; when calculated it should always come along with the GC indicator.

Accuracy – Questionnaire reporting rate

Calculation formula:

$$QRR = \frac{Q_R}{Q} \times 100$$

Where Q_R is the number of returned questionnaires completely or partially filled with valid information (a questionnaire partially filled whose information is not used for calculation of the final statistical outputs because not relevant or because being of poor quality should be counted as a non-response) and Q is the number of dispatched questionnaires.

Accuracy – Weighted questionnaire reporting rate

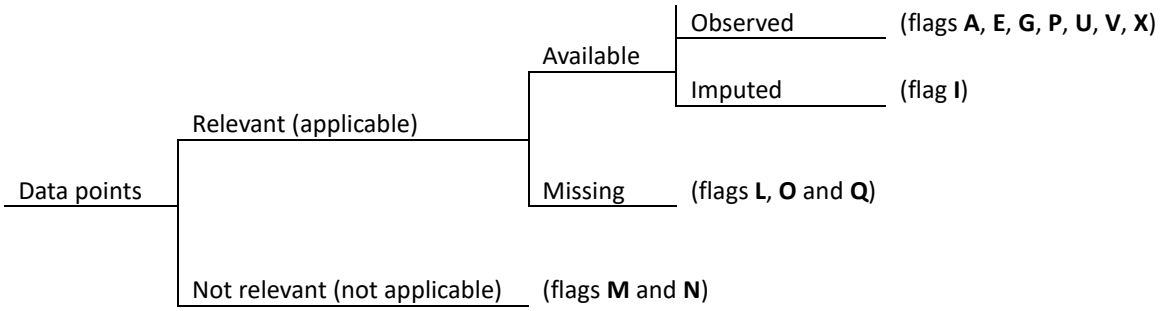
The QRR considers the percentage of countries filling-in the FAO questionnaires but it does not consider the relevance of those countries for the study of the given phenomenon; for this reason, it may be necessary to calculate a *weighted reporting rate*:

$$WQRR = \frac{\sum_{j=1}^{Q_R} w_j}{\sum_{j=1}^{Q} w_j} \times 100$$

Being w_j a numerical value expressing the importance of country *j*.
The calculation of WQRR is optional; when calculated it should always come along with the QRR indicator.

Accuracy - Observation/Non observation indicators

A large part of the accuracy indicators related to observation/non-observation can be derived by applying the following scheme:



The scheme highlights the connection between the various items and the standard flags adopted by FAO[8] and summarized in the Table below.

ID	NAME
A	Official value
B	Time series break
E	Estimated value
F	Forecast value
G	Experimental value
I	Value imputed by a receiving agency
L	Missing value; data exist but were not collected
M	Missing value; data cannot exist
N	Not significant
O	Missing value
P	Provisional value
Q	Missing value; suppressed
S	Strike and other special events
U	Low reliability
V	Unvalidated value
X	Value from international/mandated organization

Please note that:

- the flags “B” and “S” are not included in the scheme since both can be associated to a value or to a missing value[9]; in general, their use in FAO is quite limited.
- The flag “E” is listed under “available observed data” since in the recent 2019 revision of the SDMX flags it is clarified that “E” should be used when “the estimation is done by a sender agency. When the imputation is carried out by a receiver agency in order to replace or fill gaps in reported data series, the flag to use is ‘I’”[10].

[8] https://intranet.fao.org/fileadmin/user_upload/scp/Standards_for_quality_compliance/SSS_Observation_Status_Codes_Flags.pdf

[9] See SDMX Guidelines: https://sdmx.org/wp-content/uploads/OBS_STATUS_Implementation_2_0.docx

[10] https://sdmx.org/wp-content/uploads/CL_OBS_STATUS_v2_2.docx

- The flag “F” is used to denote a forecast and should not be included as imputation is used to fill in missing values that had to be observed and observation cannot be done for the future time periods/occasions.
- The flag “X” is associated with figures collected by FAO from an external organization (i.e. an international organization/supranational organization mandated to collect information within the international system, e.g. CCSA members, or on behalf of FAO or its governing bodies through specific agreements) that does not adopt the SDMX Code list for Observation Status nor alternative flagging system providing information on source and quality of value (see also implementation guidelines of the flag “A”). The use of “mirror data” sourced from the COMTRADE database, when trade data are not available for a specific country, is a special case where the flag “X” can be used to describe the mirror data reported by FAO. Statistical outputs taken from an external data set compiled NGOs, private companies/associations or other non-official/non-commissioned data sources, commonly used to replace missing values should be flagged as “I – imputed value” as this practice corresponds to cold-deck imputation (as mentioned in the FAO Statistical Standard Series on Imputation; version 1.2, 15 November 2019)

The rational under these choices is that the quality indicators are intended to measure, directly or indirectly, solely the FAO’s contribution to the overall accuracy.

Following the general scheme, some of the accuracy indicators that can be calculated are:

Percentage of missing data points:

$$PM = \frac{n_{\text{Missing}}}{n_{\text{Relevant}}} \times 100$$

$100 - PM$ returns the percentage of available data points.

Percentage of observed data points

$$RO = \frac{n_{\text{Observed}}}{n_{\text{Available}}} \times 100$$

Percentage of imputed data points

$$RI = \frac{n_{\text{Imputed}}}{n_{\text{Available}}} \times 100$$

In practice, the final indicators can be easily calculating considering that:

$$n_{\text{Observed}} = n_{\text{A}} + n_E + n_G + n_P + n_U + n_V + n_X$$

$$n_{\text{Imputed}} = n_I$$

$$n_{\text{Available}} = n_{\text{Observed}} + n_{\text{Imputed}}$$

$$n_{\text{Missing}} = n_L + n_O + n_Q$$

$$n_{\text{Relevant}} = n_{\text{Available}} + n_{\text{Missing}}$$

In these expressions, for example n_A refers to all the data points flagged as "A" (official) that are used to calculate a given final statistical output and so on.

When the final statistical output corresponds to the sum of the collected values (at global or regional level) it is important to calculate the **contribution of the imputed values to the final aggregate:**

$$CIVT = \frac{\sum_{j=1}^{n_I} y_{Ij}}{\sum_{j=1}^{n_{\text{available}}} y_{\cdot j}} \times 100$$

In practice, the denominator, $\hat{t}_y = \sum_{j=1}^{n_{\text{available}}} y_{\cdot j}$, is the estimated total of the variable *Y* (at the "World" or regional level) that is disseminated externally and the numerator is the sum of the values that are obtained through imputation (flag "I").

This indicator is also relevant when the externally disseminated statistical output is a function of totals of distinct variables (e.g. ratio between totals).

The indicators PM, RO and RI, can also be calculated at level of the whole statistical domain, i.e. by considering all the data points contributing to aggregates of that specific domain (e.g. production of crops).

The calculation of the observation/non-observation indicators in absence of flags should be done following the general scheme after the classification of the data points in the main groups: "observed", "imputed" and "missing".

Reliability – revision indicators

Revision indicators are calculated for aggregates (regional or "World" level) when the disseminated statistical output is subsequently revised, as explicitly foreseen in the revision policy[11].

The **Mean Revision** is an average of the differences between the latest and the previous disseminated value over all terms of the time series:

$$MR = \frac{1}{(t_m - t_0 + 1)} \sum_{t=t_0}^{t_m} (X_{Lt} - X_{Pt})$$

X_{Lt}: latest available estimate of the aggregate (regional or "World" level) for the variable of interest at time *t*;

X_{Pt}: previous available estimate of the aggregate (regional or "World" level) for the variable of interest at time *t*;

11 http://intranet.fao.org/fileadmin/user_upload/scp/Standards_for_quality_compliance/SSS_Data_revision__endorsed_30_January_2019_.pdf

t_0: starting year in the time series: the starting year should not be a year before a break in the time series;
t_m: last year in the time series.

MR provides, on average, the direction of the latest revision vs. the previous one (or the original values). Since revision policies usually foresee that an aggregate is revised more than once, the "P" (previous) estimate considered here is preferably the first released estimate.

The **Mean Absolute Revision** assesses the average size of the revision and can be measured as follows**:**

$$MAR = \frac{1}{(t_m - t_0 + 1)} \sum_{t=t_0}^{t_m} |X_{Lt} - X_{Pt}|$$

Frequently, the MAR is expressed in relative terms, i.e. **Relative MAR**:

$$RMAR_L = \frac{\sum_{t=t_0}^{t_m} |X_{Lt} - X_{Pt}|}{\sum_{t=t_0}^{t_m} |X_{Lt}|}$$

An alternative indicator can be calculated by changing the denominator, i.e. by measuring the relative size of the revision compared to the previous value ("P"), i.e.

$$RMAR_P = \frac{\sum_{t=t_0}^{t_m} |X_{Lt} - X_{Pt}|}{\sum_{t=t_0}^{t_m} |X_{Pt}|}$$

The following example, refers to the case where an aggregate is revised twice (e.g. in some FAO questionnaires every year the data on the latest 3 years are collected, therefore countries have the possibility to report every year not only the data on the latest year, but also revised estimates on the previous two years).

Date of dissemination of aggregates	**Reference year of the data**										
	…	**2008**	**2009**	**2010**	**2011**	**2012**	**2013**	**2014**	**2015**	**2016**	**2017**
2012		L_2008	P2_2009	P1_2010							
2013		L_2008	L_2009	P2_2010	P1_2011						
2014			L_2009	L_2010	P2_2011	P1_2012					
2015				L_2010	L_2011	P2_2012	P1_2013				
2016					L_2011	L_2012	P2_2013	P1_2014			
2017						L_2012	L_2013	P2_2014	P1_2015		
2018							L_2013	L_2014	P2_2015	P1_2016	
2019								L_2014	L_2015	P2_2016	P1_2017

The table shows that the latest revision ("L") of an estimate (final) is in the cell with yellow and green background colors and is obtained after two years from the initial released provisional estimate (P1, cell with red background color). Comparing values in the yellow cells with the corresponding ones in the red ones allows calculating the revision indicators of latest estimate vs the first initial estimate (L vs. P1). In practice, the MR, MAR and RMAR indicators can be calculate only where both L and P1 aggregates are available, i.e. the years from 2010 to 2015.

Timeliness

The **overall timeliness** is calculated as:

$$T_{all} = d_{diss} - d_{ref}$$

It corresponds to the difference between two dates, expressed in months (days for monthly data):
d_{diss} is the date of dissemination of the statistical outputs;
d_{ref} is the last day of the reference period the statistics refer to; e.g. if statistical outputs refer to 2017 then d_{ref} = 2017-Dec-31; with monthly data it is the last day of the reference month.

The overall timeliness can be split in two components:

1) Timeliness of incoming data (data source), i.e. the number of months (days) from the reference date up to the FAO data collection:

$$T_{coll} = d_{coll} - d_{ref}$$

Where d_{coll} is the **last day** of the FAO data collection period for the data needed for producing the statistics of interest (day of arrival of the last dataset, in case of data provided by other international organizations). In case of a FAO statistical processes based on both data collected from the countries and data provided by other organizations, the date to considerer for d_{coll} is the latest one.

2) Timeliness of FAO statistical process, i.e. the number of months between the collection of the data and the dissemination of the statistical outputs based on them

$$T_{proc} = d_{diss} - d_{coll}$$

The sum of the two components returns the overall timeliness:

$$T_{all} = T_{coll} + T_{proc}$$

Punctuality

Is the delay in disseminating the results, i.e. number of months between the actual dissemination date and the one scheduled (typically indicated in an official dissemination calendar):

$$P = d_{diss} - d_{sched}$$

Coherence

When the same statistical output (Regional or World) is disseminated by different units within FAO (or by FAO and another organization), it may be worth calculating the following coherence indicator:

$$diff_X = X_A - X_B$$

This indicator gives an idea on the closeness of the estimate calculated from unit A with the one produced by unit B, where unit A is the unit reporting this indicator. It can be expressed in relative terms as follows:

$$rd_X = \frac{X_A - X_B}{X_B}$$

Note that in both the expressions the reference estimate for the comparison is X_B, in other words we compare X_A with X_B, i.e. how close is X_A to X_B.

Comparability

Comparability over time is measured on time-series of statistical outputs. It is calculated as the number of comparable data items in the time series since the last structural break.

Annex 2: Calculation level

Quality dimension	Quality indicator	Calculation level/units	Geographical level	Disseminated outputs
Relevance	Average proportion of satisfied users (M)	Responses at the user satisfaction survey on principles P1-P5		Referred to the key variables/statistical outputs
	Geographical completeness (GC)	Countries that should contribute to the global/regional final statistical output	Global and Regional	Referred to the key variables/statistical outputs
Accuracy	Questionnaire Reporting Rate (QRR)	Questionnaires dispatched	Global and Regional	
	Weighted Questionnaire Reporting Rate (WQRR)	Questionnaires dispatched	Global and Regional	
	Percentage of missing data (PM)	Input data points that should be aggregated to derive the final regional/global statistical output	Global and Regional	Referred to key the variables/statistical outputs or to the whole domain
	Percentage of observed data (PO)	Input data points that should be aggregated to derive the final regional/global statistical output	Global and Regional	Referred to the key variables/statistical outputs or to the whole domain
	Percentage of Imputed values (PI)	Input data points that should be aggregated to derive the final regional/global statistical output	Global and Regional	Referred to the key variables/statistical outputs or to the whole domain
	Contribution of Imputed Values to Totals (CIVT)	Input data points that should be aggregated to derive the final regional/global statistical output	Global and Regional	Referred key variable(s) contributing to total(s) or to a function of totals
Reliability	Mean Revision (MR)	Final regional/global aggregates contributing a time series	Global and Regional	Referred to the key statistical outputs that are revised
	Mean Absolute Revision (MAR) (or RMAR)	Final regional/global aggregates contributing a time series	Global and Regional	Referred to the key statistical outputs that are revised
Timeliness	Overall Timeliness	Dates	Global and regional (if regional are disseminated later)	
	Timeliness of incoming data	Dates	Global and regional (if regional are disseminated later)	
	Timeliness of FAO statistical process	Dates	Global and regional (if regional are disseminated later)	
Punctuality	Delay in publication (if a dissemination calendar exists)	Dates	Global and regional (if regional are disseminated later)	
Coherence	Difference or relative difference between the same statistical outputs produced by different units	Final regional/global aggregates disseminated	Global and regional	Referred to the key statistical outputs
Comparability	Number of comparable data items in a time series	Final regional/global aggregates contributing a time series	Global and regional	Referred to the key statistical outputs
Accessibility	Number of web pages visits	Visits		
	Number of data downloads	downloads		