

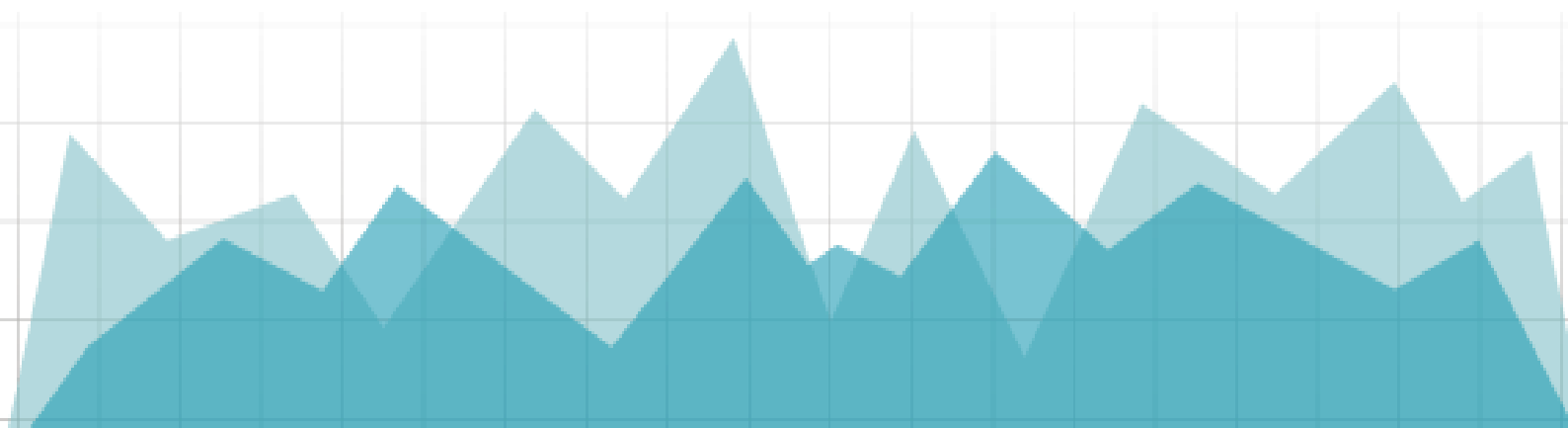


Food and Agriculture Organization
of the United Nations

Statistical Standard Series

Observation Status Code List
Version 3

Endorsed by the Technical Data Coordination Group (DCG-T)
on 26 May 2023



This standard provides the list of recommended observation status codes (flags) to be used when disseminating FAO statistics to external data users. It includes the definitions and guidance on how to use the codes.

CONTENTS

BACKGROUND	5
DEFINITIONS	5
GENERAL RECOMMENDATIONS	6
TECHNICAL RECOMMENDATIONS	6
<i>Flags for primary data</i>	6
<i>Flags for secondary data</i>	7
<i>Flags for secondary data – Implementation guidelines</i>	8
<i>Flags for derived variables</i>	13
<i>Flags for derived variables – Implementation guidelines</i>	13
<i>Flags for aggregates</i>	16
<i>Flags for aggregates – Implementation guidelines</i>	16
GOVERNANCE PROCEDURES	18
Annex 1: FAO flag hierarchy	19
Annex 2: Mapping between the code list for observation status established by FAO and the guidelines of the SDMX Statistical Working Group	20
Annex 3: Document history.....	21

BACKGROUND

FAO statistical activities aim at producing and disseminating statistics on food, nutrition, agriculture, forestry and fishery that permit comparisons over time and between geographical regions. A large part of FAO statistics is based on secondary data, i.e. data supplied by national statistical authorities or by other international organizations. These data are disseminated by FAO either in their initial form or after some transformation, reclassification or other manipulation, including further aggregation to derive regional and global aggregates. FAO also collects microdata (primary data sources) which serve as basis for the calculation of estimates on relevant phenomena. In other cases, different types of data are integrated and become the input of complex models aimed at producing nowcasts and forecasts.

In 2016, to increase the accessibility and clarity of the FAO statistics, it was decided to establish a unified and standardized system of codes (also named flags) to be associated to data disseminated externally by FAO, which would provide to users a clear indication of the source of the figures (or missing figures); the system was inspired by the SDMX “Code list for Observation Status”.

The system of flags is intended to be used by all potential users of FAO statistics, including international organizations which are going to use FAO data as input for their statistical production. Version 3 of the Standard updates the 2016 standard to account for the specific characteristics of the FAO integrated statistical system that includes the Statistical Working System (SWS) and the Statistical Data Warehouse (SDW), FAO new integrated disseminated portal; in addition this new version of the standard considered the recent updates to SDMX Code list for Observation Status (2019; version 2.2¹). The updates introduced in this document increases the alignment of FAO observation status codes with the corresponding SDMX list, although some few additional flags are still considered.

This version of FAO list of Observation Status codes is tailored to be used in the dissemination of FAO statistical outputs to the external users, and in data exchange with other international organizations.

DEFINITIONS

Observation status code (flag): “attribute of a cell in a dataset representing qualitative information on the value of that cell”², it provides information on the quality of a value or unusual or missing values. In SDMX, flags are translated into codes that describe the *observation status*, where an *observation* is the “value of a variable”.

Primary data: is commonly referred to a source of data (primary data source) collected for statistical purposes, i.e. a set of data referred to statistical units (individuals, households, enterprises, farms, etc.) that represent the basis for the production of statistical outputs (secondary data) about one or more phenomenon for the considered target population. Popular examples of primary data sources are the data collected in a Census or a sample survey carried out by a national statistical agency.

Secondary data: in this document it is referred to estimates, more in general statistical outputs, compiled usually by national statistical agencies based on primary data (see previous definition) in order to get insights on some social or economic phenomena. The terms “secondary” reflect the fact that FAO usually collects these estimates from national institutions, rather than collecting directly the primary data on which the statistical outputs are based on. Note that the adopted definition of “secondary data” should not be confused with that of “secondary data source” that was recently introduced to identify all

¹ https://sdmx.org/wp-content/uploads/CL_OBS_STATUS_v2_2.docx

² OECD Glossary of Statistical Terms <https://stats.oecd.org/glossary/detail.asp?ID=5168>

other data sources other than data collected for statistical purposes (primary data sources)³.

Derived variable: is a variable that is created as a function of one or more other variables. The creation of a derived variable can involve a simple transformation of a variable (e.g. units of measure conversion), ratio of variables, sums and more complex aggregations expressions.

Regional and global aggregates: statistical outputs referred to geographical sub-regions, regions or to the whole World obtained by aggregating the statistical outputs referred to the countries belonging to each target geographical area.

GENERAL RECOMMENDATIONS

- Each statistical output disseminated externally by FAO should come along with the corresponding flag as indicated in the next Section (technical recommendations). Furthermore, missing statistical outputs should have an associated flag explaining the reason for the missing values.
- An estimated or missing statistical output should be associated with only one single flag. As the proposed flags are not always mutually exclusive (e.g. a missing value can generate a break in time series, an estimated value can be of low reliability, data provided by other international organizations are often official country estimates, etc.), in case of doubts the flag having a higher hierarchy should be adopted. The flags hierarchy is indicated in the Annex 1.

TECHNICAL RECOMMENDATIONS

Flags for primary data

It is not necessary to attach an observation flag to microdata disseminated by FAO as the nature of these data (primary data collected by FAO or another organization) would be described in the metadata.

The Flag “E” (Estimated value) should always be used to accompany statistical outputs produced by FAO on the basis of primary data collected or obtained by the Organization (e.g., national estimates produced on the basis of a nationally-representative survey collected by FAO, statistics calculated from a microdata file publically available or shared with FAO by a national statistical agency) without going through a collection of secondary data.

³ UNECE (2011) Using Administrative and Secondary Sources for Official Statistics - A Handbook of Principles and Practices, United Nations Economic Commission for Europe
https://unece.org/DAM/stats/publications/Using_Administrative_Sources_Final_for_web.pdf

Flags for secondary data

Table 1 provides the list of flags that can be associated to secondary data collected by FAO and disseminated externally as well as to missing secondary data.

Table 1 – The code list for Observation Status to be associated to secondary data (including missing values) disseminated by FAO

ID	NAME	DESCRIPTION
A	Official value	Value provided as official when the source agency assigns sufficient confidence that it is not expected to be dramatically revised.
B	Time series break	Observations are characterized as such when different content exists or a different methodology has been applied to this observation as compared with the preceding one (the one given for the previous period).
E	Estimated value	Observation obtained through an estimation methodology or based on the use of a limited amount of data (e.g. to produce a value at an early stage of the production stage while not all data are available). If needed, additional information can be provided through free text using the COMMENT_OBS attribute at the observation level or at a higher level (in SDMX-compliant environment). This code is also to be used when the estimation is done by a sender agency (and flagged as such). 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 "Value imputed by a receiving agency".
F	Forecast value	Value deemed to assess the magnitude which a quantity will assume at some future point of time (as distinct from "estimated value" which attempts to assess the magnitude of an already existent quantity).
G	Experimental value	Data collected on the basis of definitions or (alternative) collection methods under development. Data not of guaranteed quality as normally expected from provider.
I	Value imputed by a receiving agency	Observation imputed by a receiving agency to replace or fill gaps in reported data series. This code is intended to cover all cases where a receiving agency publishes data about a sending agency that do not come from an official source in the sender agency's reporting framework. When the estimation is done by the sender agency, the flag to use is E "Estimated value".
L	Missing value; data exist but were not collected	Used, for example, when some data are not reported/disseminated because they are below a certain threshold.
M	Missing value; data cannot exist	Used to denote empty cells resulting from the impossibility to collect a statistical value (e.g. a particular education level or type of institution may be not applicable to a given country's education system).
N	Not significant	Used to indicate a value which is not a "real" zero (e.g. a result of 0.0004 rounded to zero).
O	Missing value	This code is to be used when the reasons why data are missing cannot be determined.
P	Provisional value	An observation is characterized as "provisional" when the source agency – while it bases its calculations on its standard production methodology – considers that the data, almost certainly, are expected to be revised.

Q	Missing value; suppressed	Used, for example, when data are suppressed due to statistical confidentiality considerations.
S	Strike and other special events	Special circumstances (e.g. strike) affecting the observation or causing a missing value.
U	Low reliability	This indicates existing observations, but for which the user should also be aware of the low quality assigned.
V	Unvalidated value	Observation as received from the respondent without further evaluation of data quality.
X	Value from international/mandated organization	Observation from an international or a supranational organization that does not use any flagging system in data sharing

Flags for secondary data – Implementation guidelines

Flag A – Official value

The flag “A” is the default flag assigned to an official figure that is originally reported to FAO by a statistical agency being part of the national statistical system.

If the consulted national statistical agency does not indicate a flag or an annotation indicating special circumstances in association to an observation, then the default code “A - official value” can be assumed.

The flag “A” applies also to statistical outputs taken from official country publications or websites.

If a national statistical agency explicitly provides a data point but states that it cannot be considered official, the flag to use should be derived on the basis of the flag of the sender (e.g. “P” flagged values should be flagged “P”) or if no flag is provided, the default flag to use is “E” (estimated value).

Official data may be collected by *external organizations*; two cases can be distinguished:

- i. Data falling under FAO mandate but that are collected by other international/regional bodies having a formal agreement (MoU, etc.) with FAO (outsourced data collection);
- ii. Data collected from national agencies by another international organization (usually being members of the CCSA) because the data fall under the mandate of that organization (not under FAO mandate)

Official data collected by bodies listed under item (i) should be flagged as “A - official value” unless otherwise stated; i.e. a value imputed by the organization collecting data on behalf of FAO should take the flag “E – estimated value”.

More in general, when collecting data from international/mandated organizations already adopting the SDMX Code list for Observation Status, a data point flagged as “A” by the sender should be flagged as “A” by FAO. If the data come with no flagging system or without information on source and quality of value then the flag “X - Value from international/mandated organizations” should be used.

Flag B – Time series break

The flag “B” can be associated to both a missing and non-missing disseminated statistical output.

If a statistic disseminated comes along with the flag “B”, it indicates that this statistic cannot be compared with the value preceding it in the time series.

A missing value associated with the “B” flag indicates that the given target value either cannot be

produced at this time or will no longer be disseminated (e.g. because of a change in concepts or definitions).

Flag E – Estimated value

The flag “E” must be used when a national statistical agency provides a data point but explicitly states that it cannot be considered “official” because, for instance, obtained as the result of a non-standard process.

In statistics, the concept of “**estimation**” is quite general and denotes any “inference about the numerical value of unknown population values from incomplete data such as a sample”⁴; although following this definition also an official value is the outcome of an estimation process, in this flagging system the term “estimation” is used to denote a value obtained applying an estimation methodology different from the one normally used to produce the official figures.

When collecting data from external organizations already adopting the SDMX Code list for Observation Status, a data point flagged as “I” (imputed) or “E” (estimate) by the sender will get the flag “E” by FAO.

Nowcasting cannot be seen as an imputation of a missing value. As such, the “E” flag should also be used to denote **nowcasts** produced by FAO or provided to FAO by another international organization.

Backcasting⁵ is a statistical technique being employed to ensure the coherence of the time series across time. When used by FAO to compensate missing values causing breaks in time series, backcasts should be considered as imputed values and flagged as “I” (note that this is handled in a different manner in SDMX CL_OBS_STATUS v. 2.2, which instead suggests using the flag “E”). However, if backcasts are done by national reporting agencies or by international/supranational organizations providing data to FAO, they should be flagged as “E”.

A missing value replaced by FAO with a non-missing statistical output should be flagged by “I” (imputed value).

Flag F – Forecast value

The flag “F” should be used to denote **forecasts**⁶ done by FAO, or forecasts produced by national statistical agencies and disseminated by FAO.

When collecting data from other international organizations already adopting the SDMX Code list for Observation Status, a data point flagged as “F” (forecast) should also get the flag “F” in FAO database.

Flag G – Experimental value

Experimental statistics are data obtained using new data sources and methods under development different from that normally used to produce the official values. Being procedures under development, in general it may not be guaranteed quality as normally expected from provider, although the experimental method may be designed to better respond to our users' needs in a timely manner.

When using experimental methods/procedures that have not reached full maturity in terms of harmonization, coverage or methodology, etc., it is suggested to flag the outcoming values with “G –

⁴ OECD glossary of statistical terms; “making an educated guess about a quantity that is unknown, but concerning which some information is available” under “Estimation, classical” in *Encyclopedia of statistical sciences 2nd edition*, edited by N. Balakrishnan et al., John Wiley & Sons, 2006.

⁵ Backcasting refers to forecasting backward in time; the term has also been used for extrapolation. This is done by applying the forecasting method to a series starting from the end and going to the beginning of the data.
<https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Glossary:Backcasting>

⁶ Estimate of a quantity referred to a time reference beyond the dissemination date.

Experimental value”.

Flag I – Imputed value

The flag “I” should be used in all the cases where FAO:

- i) imputes a missing value, i.e. a quantity that exists but could not be collected; or
- ii) replaces a collected value not considered reliable with a value estimated by FAO.

The flag “I” should be used when the imputed missing values are taken from an external data set compiled by NGOs, private companies/associations or other non-official/non-commissioned data sources (as this practice corresponds to cold-deck imputation, as mentioned in the FAO Statistical Standard Series on Imputation; version 1.2, 15 November 2019). To better explain the source of the imputed value it is suggested to accompany the flag “I” with an annotation providing information on the source of the data.

When FAO imputes a missing value with data collected 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, then the imputed values should be flagged as “X”.

When the value imputed by FAO (in replacement of a missing value or an unreliable collected value) is NOT ratified/validated by national agency that was expected to produce it, FAO can disseminate the imputed value flagging it as “I” and adding an annotation explaining that it is not yet ratified/validated by national agency that had to compile and send it to FAO.

In case of a missing value replaced by FAO with a 0 because the phenomenon is assumed negligible for the considered unit, the flag to use is “I” (imputed) and NOT “N – not significant”, usually associated to a 0 obtained as a result of rounding (not being a “real zero”).

The flag “I” should also be used to denote backcasts done by FAO to replace missing values back in the past of the considered time series or to replace past values in time series considered not comparable with the latest ones because of break in the series (due to a change in the definitions/methodology).

When a missing value is related to a phenomenon that does not exist in the country (the measurement cannot be done), the missing value must NOT be replaced by any value (including 0). It should be flagged as “M”.

Flag L – Missing value; data exist but were not collected

The flag “L” can only be associated to a missing value when the target quantity exists but it is not possible to collect/estimate it.

When the missing value provided to FAO is caused by a special event (strike, extreme climatic event, war, etc.) and the missing value is not replaced (imputed) by FAO with a non-missing value then the flag to use is “S – Strike and other special events”.

Flag M – Missing value; data cannot exist

The flag “M” can only be associated to a missing value referring to a target quantity that cannot exist (e.g. the planting area or the production of a crop that cannot be harvested in a given country). In doubt, if it is unknown whether the data is missing because it exists but was not collected (Flag L) or because it is the result of a target quantity that cannot exist, the missing value should be flagged as “O” (see below)

In case of a quantity that cannot exist, the dissemination of a “0” value associated with the Flag “M” or

any other flag (as per hierarchy rules) should be avoided. The value should be left missing.

Flag N – Not significant

The Flag “N” is used together with a “0” value to indicate a value that is not a "real" zero, but instead the result of a rounding operation.

Flag O – Missing value

The flag “O” can only be associated to a missing value. It must be used when there is no additional information to ascertain why it is missing (e.g. when flags L and M cannot be used with certainty).

Flag P – Provisional value

An observation is characterized as “provisional” (flagged as “P”) when the source agency (including FAO) considers that it, almost certainly, is expected to be revised.

Flag “P” should be used for data reported to FAO as provisional by a sender agency that is either a national statistical agency or an international or supranational organization that adopts SDMX in flagging the provided value or clearly states that a given value has to be considered as provisional.

Flag Q – Missing value; suppressed

The flag “Q” can only be associated to a missing value when the target quantity exists but it is suppressed:

- for confidentiality reasons;
- because the value provided to FAO is deemed of poor quality and consequently it is deleted and not replaced by any imputed value;
- the imputation done by FAO to replace a missing value or a collected value found wrong or of poor quality is NOT validated⁷ by the national statistical agency that provided or that had to provide it and FAO decides not disseminate any value (note that FAO may also decide to disseminate the imputed values flagging it with “I” with the corresponding annotation or to maintain the country value flagging it as “U - low reliability”).

Flag S – Strike and other special events

The lag “S” can be used in connection to both observed and missing values:

- It should be associated to a non-missing value provided to FAO by a National statistical agency to highlight an extreme value that actually occurred in the phenomenon being observed because of a special circumstance (e.g. strike, extreme climatic event, war, etc.).
- It can also be associated to a missing value provided to FAO only when the missingness is caused by a special event and the missing value is not replaced (imputed) by FAO with a non-missing value.

⁷ Although it would be advisable to ask to national statistical agencies to validate all FAO imputations to ensure country ownership and avoid open disagreements on the data published by FAO, this practice would be too burdensome if applied extensively and therefore it is preferable to apply validation procedures mainly in case of very visible statistical outputs (e.g. SDG indicators).

Flag U – Low reliability

The flag “U” should be used to denote data received by FAO by a sender agency which is characterized by poor quality.

In the event of a dispute between FAO and national institutions concerning the quality of the country figure, the FAO may disseminate the data submitted by the national agency with the flag “U” that indicates the low quality and non-compliance with quality standards.

The flag “U” should not be associated to provisional values, even if by nature they could be characterized as of low accuracy. However, as they are expected to be revised, they should be flagged as “P”.

Flag V – Unvalidated Value

The Flag “V” should be used to denote observations received from the national reporting agency or international agency and used/re-disseminated by FAO without carrying out checks to assess whether they are reliable or not.

Flag X – Value from international/mandated organization

The flag “X” should be 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 by 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). To better explain the source of the imputed value it is suggested to accompany the flag “I” with an annotation providing information on the source of the data.

Flags for derived variables

FAO disseminates the collected secondary data either in their initial form or after some transformation, reclassification or other manipulations, which can involve different data items, and result in so-called **derived variables**. Occurring data manipulations are:

- Simple transformations (e.g. unit measure transformation; recoding/re-classification, etc.);
- Calculation of ratios (e.g. $\text{yield} = \text{production} / \text{area_harvested}$ of a given product in a country);
- Calculation of sums or averages (e.g. total land area in a country);
- Calculation of proportions (a part in relation to the total, e.g. proportion of forest area in country)

In other cases, the input data items referred to a country are used for deriving complex indicators, as in the case of some SDG indicators (e.g. the country's Agriculture Orientation Index; the Food Loss/Waste Index, etc.) and more in general variables that cannot be directly measured.

When the output of such transformation/manipulation done by FAO is disseminated to external users then it should come along with proper observation status that allows its correct use and interpretation. The flag to use depends on the flags associated to the contributing values as well as the type of aggregation formula/method. The main rules are provided in the following section.

Table 2 – Flags that can be associated to the derived variables compiled by FAO

ID	NAME
A	Official value
B	Time series break
E	Estimated value
F	Forecast value
M	Missing value; data cannot exist
O	Missing value
P	Provisional value
X	Value from international/mandated organization

Flags for derived variables – Implementation guidelines

Flag A – Official value

With the exceptions of the particular cases listed below, the flag “A” should be associated to values of derived variable compiled by FAO when:

- all contributing items are based on official values; or
- at least 2/3 of contributing data items are official values, and the remaining items can be estimated/imputed with reasonable confidence; or
- the derived variable (e.g. a total) is provided to FAO directly from the data provider (as official data) without however providing the contributing data items (in such case, if FAO imputes the contributing data points, they should be flagged as “I”; if some of the contributing data items are assumed

negligible and therefore close to 0, FAO should explicitly disseminate a 0 value with the associated flag “I”).

Exceptions/particular cases:

In case of ***simple transformations*** of a single official value, if the transformation follows international standard (e.g. standard unit of measures) then the transformed values should maintain the flag “A”. However, if the transformation is based on rules/coefficients decided by the FAO technical unit and NOT being part of international standards then the corresponding transformed value should be flagged as “E – estimated value”

In case of ***proportions***, the flag “A” should be used only if the figures at numerator (part of the total amount) and denominator (total amount) are both flagged as “A”.

More generally, for ***ratios***, the resulting figure should be flagged as “A” when both the contributing quantities are flagged as official.

In case of ***sums or simple averages*** where not all contributing items are official, if the values of the non-official items cannot be estimated with a reasonable amount of certainty or if they contribute for more than 1/3 of the derived value, the derived variable should be flagged as “E”.

When the derivation of outputs involves ***aggregation*** formulas with coefficients set by FAO according to some assumptions (e.g. weighted sums/averages with weights set by FAO according to some assumptions that are not defined in an international standard), the corresponding statistical output should be flagged as “E” even if all the contributing data entries are all official.

Flag B – Time series break

The flag “B” is associated to both a missing and non-missing derived variable value.

A numerical estimate of a derived variable disseminated with the flag “B” indicates that it cannot be compared with the value preceding it in the time series because of different methods in calculation or in the definition of one or more of the input data points needed to calculate the derived variable. In other words, the flag should be used when there is at least one of the contributing values flagged as “B”

A missing value associated with the “B” flag indicates that the estimates referred to given target quantity that either cannot be produced at this time or will no longer be disseminated (e.g. because of a change in concepts or definitions).

Flag E – Estimated value

The flag “E” has to be used when:

- the estimation of the derived variable is done using a complex calculation involving a series of assumptions or coefficients decided by FAO; or
- the final estimate is based on input data points that are flagged in different manners and the flag cannot be derived based on the guidelines related to other possible flags (other than E) for derived variables.

The flag “E” should also be used to denote nowcast or backcasts of derived variable values compiled by FAO.

If needed, additional information can be provided through free text using the COMMENT_OBS attribute at the observation level or at a higher level (SDMX-compliant environment).

Flag F – Forecast value

The flag “F” should be used to denote a forecast (a quantity referred to a time reference beyond the dissemination date) for the derived variable produced by FAO. The flag is used for all forecasting exercises, no matter the source of the contributing values (official, imputed, estimated, etc.).

Flag M – Missing value data cannot exist

The flag “M” can only be associated to a missing value when it is impossible to derive an estimate for the targeted derived variable because it cannot exist (e.g. the yield of a product that cannot be harvested in a country).

Flag O – Missing value

The flag “O” can only be associated to a missing value. It must to be used when it is not possible to compile the derived variable value for different reasons; e.g. all of the input values necessary to compile the derived variable are missing due to various reasons; or not all input values are missing, but the combined result cannot be calculated (e.g. missing value for the denominator in a ratio).

Flag P – Provisional value

The flag “P” should be associated to the value of a derived variable compiled by FAO that is completely based on values flagged as “P”.

Flag X – Value from international/mandated organization

This flag “X” should be associated to the value of a derived variables compiled by FAO that is completely based on values flagged as “X”, with some exceptions that are listed below.

In case of *simple transformations* of a single value flagged as “X”, if the transformation follows international standard (e.g. standard unit of measures) then the transformed values should maintain the flag “X”. However, if the transformation is based on rules/coefficients decided by the FAO technical unit and NOT being part of international standards then the corresponding transformed value should be flagged as “E – estimated value”.

When the derivation of outputs involves *complex aggregation formulas* with coefficients set by FAO according to some assumptions then the corresponding statistical output should be flagged as “E” even if all contributing data entries are all flagged as “X”.

Flags for aggregates

In producing its statistical outputs, the FAO enriches the collected data also by estimating regional and global aggregates for both the collected data and also for the derived variables. The estimated global and regional statistical outputs disseminated externally should come along with flags of observation status that will be assigned depending on the input data points and on the aggregation formula (see following section for more details).

Table 3 – Flags that can be associated to the global and regional aggregates compiled by FAO

ID	NAME
A	Official value
B	Time series break
E	Estimated value
F	Forecast value
M	Missing value; data cannot exist
O	Missing value
P	Provisional value
X	Value from international/mandated organization

Flags for aggregates – Implementation guidelines

Flag A – Official value

With the exceptions of the particular cases listed below, the flag “A” should be associated to aggregates compiled by FAO when:

- all contributing items are based on official values; or
- at least 2/3 of the contributing data items are official values, and the non-official items can be estimated/imputed with reasonable confidence.

Exceptions/particular cases

In cases where not all contributing items are official, if the values of the non-official items cannot be estimated with a reasonable amount of certainty or if they contribute for more than 1/3 of the aggregate, the derived variable should be flagged as “E”.

If the derivation of regional/global estimates is based on **aggregation methods** decided by FAO and not being established as international standards (e.g. the regional aggregate is obtained by applying the weighted average but with a different weighting system other than that suggested in the corresponding international standards/guideline) or when the regional aggregate is obtained by applying an ad hoc statistical method instead of the standard aggregation formula then the corresponding output value should be flagged as “E” even if all the contributing data entries are all official.

Flag B – Time series break

The flag “B” is associated to both a missing and non-missing values disseminated as global and regional aggregate when one of the contributing values is flagged as “B” or there is a change in the methodology for deriving the global/regional aggregates.

A missing value associated with the “B” flag indicates that the estimates referred to given target quantity that either cannot be produced at this time or will no longer be disseminated (e.g. because of a change in concepts or definitions).

Flag E – Estimated value

The flag “E” has to be used when:

- derivation of regional/global estimates is based on complex aggregation formula involving coefficients set by FAO according to some assumptions (being not established as international standards); or
- the final estimate is based on input data points that are flagged in different manners and the flag cannot be derived based on the criteria listed under the possible flags for aggregates.

The flag “E” should also be used to denote nowcasts or backcasts of values of aggregated variables compiled by FAO.

If needed, additional information can be provided through free text using the COMMENT_OBS attribute at the observation level or at a higher level (SDMX-compliant environment).

Flag F – Forecast value

The flag “F” should be used to denote a forecasted regional/global aggregate.

Flag M – Missing value data cannot exist

The flag “M” can only be associated to a missing value when it is impossible to derive an estimate of the global/regional aggregate because it cannot exist (e.g. the yield of a product that cannot be harvested in a region).

Flag O – Missing value

The flag “O” can only be associated to a missing value. It must to be used when there is no additional information to ascertain the reasons for missingness or when target quantity exists but all or some of the points necessary to compile the global/regional aggregate are missing due to various reasons (no imputation is done).

Flag P – Provisional value

The flag “P” should be associated to the value of a regional/global aggregate compiled by FAO that is completely based on values flagged as “P”.

This flag should also be used when FAO decides to produce aggregates on the basis of a limited number of country values, when data collection still ongoing.

Flag X – Value from international/mandated organization

The flag “X” should be associated to regional/global aggregates compiled by FAO that are completely based on values flagged as “X”.

When the derivation of regional/global aggregates is based on *complex aggregation formulas* with coefficients set by FAO according to some assumptions then the corresponding statistical output should be flagged as “E” even if all the contributing data entries are all flagged as “X”.

GOVERNANCE PROCEDURES

- All units processing or disseminating data are responsible for the implementation of this Standard.
- Proposals for extensions and modifications to the standard shall be submitted for approval to the DCG-T and the Chief Statistician.
- The SDMX Statistical Working Group is consulted to align with the cross-domain code list for Observation Status⁸, to the extent possible.

⁸ https://sdmx.org/?sdmx_news=new-version-of-code-list-for-observation-status-version-2-2

Annex 1: FAO flag hierarchy

The proposed flags are not always mutually exclusive (e.g. a missing value can generate a break in time series, an estimated value can be of low reliability, data provided by other international organizations are often official country estimates, etc.). For this reason and because only one flag can be assigned per observation, this hierarchy below facilitates which flag to be used. The flag having the highest priority (appearing in the top of the table) should be picked-up.

Observation status hierarchy <i>(from highest to lowest importance)</i>	Relevant in conjunction with...	
	numeric values	missing values
B / time series break (highest importance)	Yes	Yes
O / missing value (unknown reasons)		Yes
M / missing value; data cannot exist		Yes
L / missing value; data exist but were not collected		Yes
Q / missing value; suppressed		Yes
S / strike and other special events	Yes	Yes
I / imputed value	Yes	
F / forecast value	Yes	
E / estimated value	Yes	
P / provisional value	Yes	
N / not significant	Yes	
U / low reliability	Yes	
V / unvalidated value	Yes	
G / experimental value	Yes	
X / value from international/mandated organizations	Yes	
A / official value	Yes	

Annex 2: Mapping between the code list for observation status established by FAO and the guidelines of the SDMX Statistical Working Group

FAO (2016)		FAO (2023)		SDMX CL_OBS_STATUS (2019)	
B	time series break	B	time series break	B	time series break (highest importance)
O	missing value	O	missing value (unknown reasons)	O	missing value
M	missing value (cannot exist, Not Applicable)	M	missing value; data cannot exist	M	missing value; data cannot exist
M	missing value (cannot exist, Not Applicable)			H	missing value; holiday or weekend
M	missing value (cannot exist, Not Applicable)	L	missing value; data exist but were not collected	L	missing value; data exist but were not collected
Q	missing value; suppressed	Q	missing value; suppressed	Q	missing value; suppressed
				J	derogation
S	exceptional event	S	strike and other special events	S	strike and other special events
				D	definition differs
				K	data included in another category
				W	Includes data from another category
I	imputed value	I	imputed value	I	imputed value
F	forecast value	F	forecast value	F	forecast value
E	estimated value	E	estimated value	E	estimated value
E	estimated value	G	experimental value	G	experimental value
E	estimated value	U	low reliability	U	low reliability
P	provisional value	P	provisional value	P	provisional value
N	not significant (negligible)	N	not significant	N	not significant
		V	unvalidated value	V	unvalidated value
<blank>	official figure	A	Official value	A	normal value
R	revised official figure				
T	unofficial figure				
X	figure from international organization	X	value from international/mandated organization		

Annex 3: Document history

Version	Date	Author	Description of changes/status
REV.1		Corporate FAO Standards Team	Endorsed by SCWG T:\Team_working_folder\F\Standards\Flags\SW S-05 Annex-05 Observation Status Codes.doc
REV. 2	07-08/2016	V.Ramaschiello; J. Karlsson; C. Fabi	Endorsed document ¹² is revised and updated.
	08-10/2016		The draft is circulated for preliminary comments to ESS team leaders and FIPS. The document is revised according to comments received.
	11/2016		The document is shared with the IDWG for comments and endorsement.
	12/2016		The document is endorsed as FAO standard.
REV. 3	04/2022	M. D’Orazio with inputs from by V. Gianfaldoni, J. Montero-Serrano and A. Steel	Revision of the standard to account for suggestions of revisions from ESS and other technical Divisions and 2019 update (Version 2.2) to the SDMX Code list for Observation Status.
	07/2022	M. D’Orazio	Revision of the standard to account for suggestions of revisions from IDWG and other meetings.
	03/2023	M. D’Orazio V. Bizier	Revision of the standard to account for outcomes of ad-hoc meetings.
	05/2023		The document is endorsed through written consultation