



Validation rules for air emissions accounts

2025 data collection

Rules for UNSD/OECD AEA questionnaire for 2025 data collection

Version: 02 June 2025

Contents

1. Introduction.....	3
2. Types of validation results	3
3. Validation rules for AEA	4
3.1 Completeness and structural integrity	4
3.2 Symbols	5
3.3 Internal consistency	5
3.4 Deletion of previously reported data.....	8
3.5 Footnotes	8
3.6 Confidentiality	10
3.7 Plausibility of reported time series	10
3.8 Plausibility of revisions	11
3.9 External consistency (cross-domain plausibility)	11

1. Introduction

Air emissions accounts (AEA) record flows of gaseous and particulate materials (six greenhouse gases including CO₂ and seven air pollutants) in a breakdown by emitting economic activities. Economic activities include production activities (hierarchically classified by the International Standard Industrial Classification (ISIC) Rev.4) and household consumption activities. As the AEA are based on the residency principle, their national total emissions deviate from the totals as defined in national emission inventories that are compiled according to the territory principle. AEA reconciles totals with national inventories through so-called 'bridging items'.

This document presents the validation rules for AEA, as applied by the Organisation for Economic Co-operation and Development (OECD) and the UN Statistics Division (UNSD) in the global data collection for AEA. These rules have been adapted from the AEA validation rules proposed by Eurostat. It can be used as reference when filling the questionnaire. Feedback on this document is welcomed and can be sent to SDD.SEEA@oecd.org and seea@un.org

Many of the validation rules in this document can and should be checked by compilers before sending the data –in particular those checks resulting in ERROR (see next chapter). This will help ensure that national statistical offices run some key validation rules and adjust their data, if needed, before sending the data to OECD and UNSD. This will facilitate the validation process done by the international organizations and ensure a faster dissemination of the data.

The validation rules presented in this document were tested during the 2023 data collection cycle and updated based on adjustments in the AEA questionnaire implemented in subsequent years. Additional questions may be asked beyond the presented rules if any specific observations are made during data validation process.

2. Types of validation results

A validation rule is a logical statement applied to data. Below is a typology of possible validation results which shows how validation rules will be applied.:

OK	This means that the transmitted data passed the validation rule and no specific follow up is required. However, it might be possible that questions will be asked during a later stage of the validation process. In general, validation rules that have 'OK-result' are not listed in the validation report.
----	---

- ERROR** This means that a serious issue related to structure and consistency was detected in the transmitted data. An 'ERROR-result' will usually trigger questions to the sender seeking for an explanation, clarification and correction.
- WARNING** This means that some reported element is 'suspicious'. It might refer to an individual data point, combination of data points or calculations based on reported data points. It highlights an issue of attention and for which a valid explanation might exist. A 'WARNING-result' will usually trigger questions to the sender seeking for an explanation and clarification of the reported values. Countries can accelerate the validation procedures by providing footnotes explaining these warnings.
- INFO** This means that some reported element is simply an issue of attention to be highlighted. A further clarification or explanation is not required.

When a validation rule triggers an ERROR or WARNING, UNSD/OECD will follow up with the compiler. This document addresses ERROR and WARNING rules.

3. Validation rules for AEA

This chapter presents a list of validation rules for AEA.

The validation rules are presented by categories. For each validation rule the type of validation result is indicated in rectangular brackets [ERROR, WARNING].

This document expresses the rules in 'human language' close to the terms of the Excel questionnaire (such as questionnaire cells, symbols, footnotes, etc.). This is the terminology familiar to most compilers.

3.1 Completeness and structural integrity

- Rule 1.** The structure check triggers an [ERROR] if the templates returned by countries deviate from the original template. It controls the first four columns of the worksheets for every gas/pollutant and detects any changes, for instance, in the naming of codes or in the number of rows.
- Rule 2.** The completeness check triggers an [ERROR] if, for a given year and pollutant, there is no value for IND-TOTAL (A_U) but there is data for at least one other item.
- Rule 3.** The Global Warming Potential (GWP) must be chosen in the Instructions worksheet if worksheets with the unit of tonnes of CO₂ equivalent (i.e. GHG, HFC, PFC and SF₆_NF₃) are filled in. Missing GWP triggers an [Error].

3.2 Symbols

Symbols are the alphanumeric characters of a reported value in the data points (cells) of the questionnaire.

Rule 4. Valid symbols are:

- In the bridging items 'other adjustments and statistical discrepancy' and 'land use, land-use change and forestry': positive value, zero value, negative value, colon (:) (i.e. 'not available')
- In any other cell: positive values, zero value, colon (:) (i.e. 'not available')

Invalid symbols trigger an [ERROR] result.

3.3 Internal consistency

This set of rules relates to the hierarchically classified characteristics, namely economic activities and bridging items. It checks that the sum of components equals the corresponding total or subtotal item.

Rule 5. The rule checks that each aggregate of activities is equal to the sum of the reported disaggregated activities. The set of the following checks applies to the detailed template. For the aggregate and medium templates, the checks apply only if the children sectors appear in each template. In the aggregate template for CO₂, for example, the formula “A_U (Total industries) = A + B + C + D + E + F + H + SERVICES + IND-NS” shall be checked and the formula “A = A01 + A02 + A03 + A-NS” is irrelevant.¹:

For Detailed and Medium templates:

$$A_U \text{ (Total industries)} = A + B + C + D + E + F + G + H + I + J + K + L + M + N + O + P + Q + R + S + T + U + \text{IND-NS}$$

For Aggregate template:

$$A_U \text{ (Total industries)} = A + B + C + D + E + F + H + \text{SERVICES} + \text{IND-NS}$$

¹ Users can choose different template by selecting “level of detail” in the “Instructions” worksheet of the global AEA questionnaire.

$A = A01 + A02 + A03 + A\text{-}NS$
$C = C10_C12 + C13_C15 + C16_C18 + C19 + C20 + C21 + C22_C23 + C24_C25 + C26 + C27 + C28 + C29_C30 + C31_C33 + C\text{-}NS$
$C16_C18 = C16 + C17 + C18$
$C22_C23 = C22 + C23$
$C24_C25 = C24 + C25$
$C29_C30 = C29 + C30$
$C31_C33 = C31_C32 + C33$
$E = E36 + E37\text{-}E39$
$G = G45 + G46 + G47 + G\text{-}NS$
$H = H49 + H50 + H51 + H52 + H53 + H\text{-}NS$
$J = J58_J60 + J61 + J62_J63 + J\text{-}NS$
$J58_J60 = J58 + J59_J60$
$K = K64 + K65 + K66 + K\text{-}NS$
$L68A < L$
$M = M69_M70 + M71 + M72 + M73_M75 + M\text{-}NS$
$M69_M71 = M69_M70 + M71$
$M73_M75 = M73 + M74_M75$
$N = N77 + N78 + N79 + N80_N82 + N\text{-}NS$
$Q = Q86 + Q87_Q88 + Q\text{-}NS$
$R = R90_R92 + R93$

$S = S94 + S95 + S96 + S\text{-}NS$
$SERVICES = G + I + J + K + L + M + N + O + P + Q + R + S + T + U$
$HH\text{ total} = HH_TR + HH_HEAT + HH_OTH$
$AEA\text{ total} = A_U + HH$
Total AEA +/- Bridging Items = Total UNFCCC/CLRTAP ²
$NONRESTERR = NRT\text{-}LAND + NRT\text{-}WATER + NRT\text{-}AIR$
$NATRESABROAD = NRA\text{-}FISH + NRA\text{-}LAND + NRA\text{-}WATER + NRA\text{-}AIR$
$H50 > NRA\text{-}WATER$ Note: If H50 is missing but NRA-WATER is not, this is flagged as a [WARNING].
$H51 > NRA\text{-}AIR$ Note: If H51 is missing but NRA-AIR is not, this is flagged as a [WARNING].
Rule for GHG
For each cell, the value reported for GHG must meet the following formula: $GHG = CO_2 + CH_4 * GWP + N_2O * GWP + HFC + PFC + SF_6_NF_3$ Note: GHG should be calculated when the minimum set of the gases (CO ₂ , CH ₄ , and N ₂ O) is reported.
Rule for GHG_BIO
For each cell, the value reported for GHG_BIO must meet the following formula: $GHG_BIO = CO_2 + CO_2_BIO + CH_4 * GWP + N_2O * GWP + HFC + PFC + SF_6_NF_3$

² Consistency check with the total UNFCCC (United Nations Framework Convention on Climate Change) applies only to the years in which country has UNFCCC inventory data. Consistency check with the total CLRTAP applies only to Parties of the CLRTAP (Convention on Long-Range Transboundary Air Pollution) (See the section 3.9 External Consistency).

Note: GHG_BIO should be calculated when the minimum set of the gases (CO ₂ , CO ₂ _BIO, CH ₄ , and N ₂ O) is reported.
Rule for CO ₂ _ROAD
For each cell, the value reported in 'CO ₂ _ROAD' sheet must be smaller or equal to the value reported in the 'CO ₂ ' sheet.
Rule for PM ₁₀ and PM _{2.5}
For each cell, the value reported for PM ₁₀ must bigger or equal to the value reported for PM _{2_5}

This is an [ERROR] rule. Differences which are either larger than 1 in absolute terms or larger than 1% of the parent value in relative terms, or both, trigger an [ERROR]. Minor rounding withing tolerance in individual gases may trigger discrepancies beyond this tolerance in GHG and GHG_BIO due to multiplying with GWP, which can be treated as non-error. Different level of reporting in individual gases may trigger inconsistencies between ISIC Rev.4 letter-level sections and 2-digit level divisions for GHG and GHG_BIO, which can be treated as non-error.

3.4 Deletion of previously reported data

Rule 6. Removing values previously transmitted generates a [WARNING].

3.5 Footnotes

The Excel questionnaire contains two types of footnote symbols: symbols for pre-defined footnotes (letters) and symbols for free text footnotes (numbers).

Rule 7. This rule checks the correctness of symbols of pre-defined footnotes. The valid four pre-defined footnotes are the following:

b)	Break in series	Break occurring when there is a change in the standards for defining and observing a variable over time. The flag 'b' is to be attached to the first time period after the break.
----	-----------------	---

c)	Confidential	Confidential data are data which are subject to confidentiality clauses
d)	Secondary confidentiality	Secondary confidential data are data made confidential in order to prevent third parties to indirectly calculate the data points genuinely flagged as confidential.
e)	Estimated data	The 'e' (estimate) flag shall be used only if one or several data points have been calculated using a significantly different methodology and/or sources than the rest of the data points in the questionnaire.
p)	Provisional	The 'p' (provisional) flag shall be used when a data point value is expected to be revised and submitted to OECD/UNSD before the next data collection. In the case of early estimates, the flag 'e' is deemed sufficient and the 'p' flag can be omitted. Notice all 'p' (provisional) flags sent during a given data collection will be systematically removed during the subsequent data collection, unless the 'p' flags are again resubmitted.

The use of non-valid footnote symbols triggers an [ERROR] result.

Rule 8. A free text footnote symbol with no text defined for it in the footnote area of the questionnaire triggers an [ERROR] result.

Rule 9. The following combinations of symbols of pre-defined footnotes and value will trigger a [WARNING]: pre-defined footnote symbols p) or e) in combination with the value (:) (i.e. 'not available').

Rule 10. Only the following combinations of pre-defined footnotes are possible and allowed. Combinations beyond those trigger an [ERROR] result.

Break in time series (b) whilst estimated data (e)
Break in time series (b) whilst provisional (p)

Break in time series (b) whilst estimated data (e) whilst provisional (p)
Confidential (c) whilst break in time series (b)
Secondary confidential (d) whilst break in time series (b)
Estimated data (e) whilst provisional (p)

3.6 Confidentiality

This check verifies certain rules related to confidentiality.

Rule 11. Confidentiality at ISIC section (i.e., 1 digit ISIC level) or at households is not allowed. This is an [WARNING] rule.

Rule 12. Any figure flagged as confidential (pre-defined footnote c) could be derived through calculation. This is a [WARNING] rule.

3.7 Plausibility of reported time series

This check detects implausible changes between consecutive years in time series of the same reporting. Implausible changes are identified against a maximum range of change that is still considered plausible for purposes of the automatic validation.

Rule 13. The existence of implausible changes triggers a [WARNING] result. 'Implausible changes' are defined as follows. The set of the following checks applies to the detailed template. For the aggregate and medium templates, the checks apply only if the corresponding sectors appear in each template. For example, the variation of A01-A03 is irrelevant to the aggregate template for CO₂..

- Totals for the letter-level of ISIC sections (i.e., A, B, C, etc.) which vary more than 30% between consecutive years if they are bigger than 10% of all industries (total A_U) in the second year.
- ISIC industries A01-A03, E36-E39 and H49-H53 which vary more than 80% between consecutive years if they are bigger than 30% of total A, E and H respectively.
- ISIC industries C10-C33 which vary more than 80% between consecutive years if they are bigger than 5% of total C.
- Any other two-digit ISIC industry which varies more than 80% between consecutive years if it is bigger than 5% of all industries (total A_U) in the second year.

- Two-digit Households positions which vary more than 30% between consecutive years if they are bigger than 50% of Total Households in the second year.
- Total Households which vary more than 30% between consecutive years if they are bigger than 10% of Total Industries plus Households in the second year.
- Bridging items which vary more than 50% between consecutive years if they are bigger than 50% of the superior total, i.e. to National residents abroad or National residents in the territory.

Rule 14. Implausible changes must be explained with a free text footnote or, in the case of breaks in the series, with a pre-defined footnote b) and the corresponding free text footnote. The absence of a footnote triggers an [WARNING] result.

3.8 Plausibility of revisions

This check detects implausible revisions between the current and previously reported data (or data previously collected at OECD.Stat based on other form of data publication if the country uses the questionnaire for the first time). Implausible revisions are identified against a maximum range of revisions that is still considered plausible for purposes of the automatic validation.

Rule 15. Implausible revisions trigger a [WARNING]. 'Implausible revisions' are those bigger than 40% if the cell contributes at least to 10 % of Total Industries or Total Households.

3.9 External consistency (cross-domain plausibility)

This series of checks compares AEA to other reported and external data.

Please note that the following cross-domain plausibility checks, except for rule 16, are not expected to be performed by countries before the data is transmitted. Some rules may not be applied to some countries depending on the availability of relevant data at the OECD and the UNSD (i.e. The consistency check with the UNFCCC (United Nations Framework Convention on Climate Change) total applies only to the years for which the country has UNFCCC inventory data. The consistency check with the total CLRTAP applies only to Parties of the CLRTAP (Convention on Long-Range Transboundary Air Pollution)).

Rule 16. The UNFCCC or CLRTAP inventory total reported in the AEA questionnaire (INV-TOTAL) is compared to the UNFCCC or CLRTAP inventory total published by the UNFCCC and the EMEP (European Monitoring and Evaluation Programme) Centre on Emission Inventories and Projections respectively. If the two totals differ, the

country is asked for clarification on the data source used in the questionnaire. This is an [ERROR] rule. Similarly, emissions (or removals) from LULUCF (land use, land-use change and forestry) should match between the AEA questionnaire and the UNFCCC inventory.

Rule 17. There is a comparison between INV-TOTAL (Inventory total) and IND-HH (Total industries and households). For emissions of CO₂, SO_x, NO_x, CO, PM_{2.5} and PM₁₀, if INV-TOTAL equals to the reported IND-HH, it is very likely that the residence principle was not applied to AEA. In this case, the country is invited to provide a footnote, clarifying whether the residence principle was applied or not. This is a [WARNING] rule.

Rule 18. Implausible differences between AEA data and PEFA (Physical energy flow accounts) data trigger a [WARNING] result

For checking the coherence between AEA and PEFA, one needs to relate tonnages of emissions to energy flows expressed in terajoules (TJ). For the resulting emission factors – i.e. tonnes of emissions per TJ of energy use – one may define plausible ranges around IPCC (Intergovernmental Panel on Climate Change) standard emission factors.

Air Emission Accounts (AEA) vs. PEFA Table C	
18.a	If CO₂ (AEA) > 0 then P00 (PEFA Table C) > 0 If CO₂ (AEA) = 0 then P00 (PEFA Table C) = 0 for 'A' to 'U', 'HH', 'HH_HEAT', 'HH_TRA', 'HH_OTH'
18.b	CO ₂ :H51 (AEA) / P15:H51 (Table C) = 72 tCO ₂ /TJ +/-40%
18.c	CO ₂ :H50 (AEA) / (P13:H50 + P17:H50 + P18:H50 + P19:H50) (Table C) = 75 tCO ₂ /TJ +/-40%
18.d	[CO₂ + CO₂_BIO (AEA)] / P00 (Table C) = annual change rate not beyond +/- 30% for 'A' to 'U', 'HH', 'HH_HEAT', 'HH_TRA', 'HH_OTH'

Rule 19. Implausible differences between AEA data and OECD's air transport CO₂ emissions trigger a [WARNING] result.

This check compares AEA data reported for air transport (ISIC H51) and the related bridging items with the respective OECD data on CO₂ emissions from air transport³.

AEA vs. OECD's CO₂ emissions from air transport data and UNFCCC data	
19.a	AEA H51 vs OECD H51 threshold $\pm 5\%$;
19.b	NRA-AIR vs OECD Bridging item - Residents abroad threshold $\pm 10\%$;
19.c	NRT-AIR vs OECD Bridging item - Non-residents on territory threshold $\pm 10\%$;
19.d	AEA H51 vs [UNFCCC 1.A.3.a (Domestic Aviation) – NRT-AIR + NRA-AIR] threshold $\pm 5\%$;

Rule 20. Implausible differences between AEA data and Energy statistics trigger a [WARNING] result. This rule applies only to OECD countries for the moment.

Residential sector detailed data from energy efficiency indicators of the IEA (International Energy Agency) (oil and oil products + gas + coal and coal products for residential space heating + residential space cooling + residential water heating + residential cooking) are compared with CO₂ emissions from heating/cooling activities by households in AEA. The calculated CO₂ emission factor should fall in the range of $\pm 40\%$ around 65 tCO₂/TJ.

Rule 21. Implausible differences between AEA data and Supply and use table (SUT) trigger a [WARNING] result.

The check verifies economic activity for each ISIC division for which AEA reports emissions (value bigger than zero). Emissions for ISIC divisions for which the SUT reported zero economic activity trigger a [WARNING].

³ [https://data-explorer.oecd.org/vis?lc=en&df\[ds\]=DisseminateFinalDMZ&df\[id\]=DSD_AIR_TRANSPORT%40DF_AIR_TRANSPORT&df\[ag\]=OECD.SDD.NAD.SEEA](https://data-explorer.oecd.org/vis?lc=en&df[ds]=DisseminateFinalDMZ&df[id]=DSD_AIR_TRANSPORT%40DF_AIR_TRANSPORT&df[ag]=OECD.SDD.NAD.SEEA)