# Guidelines for Global DSDs on SEEA



# Introduction, governance and scope

The Ownership Group for SDMX implementation for macro-economic statistics (SDMX-MES OG) mandated a technical sub-group to propose global Data Structure Definitions (DSDs) for the implementation of SDMX in SEEA. The SDMX-MES OG is already responsible for SDMX DSDs in the areas of National Accounts (incl. Government Finance), Balance of Payments, Foreign Direct Investment and Price Indices. More information on the mandate of the SDMX-MES OG and the existing packages can be found on the SDMX website:

SDMX-MES packages on National Accounts and Government Finance as well as on Consumer Price Index (CPI and HICP): https://sdmx.org/?page_id=1498

SDMX-MES packages on Balance of Payments and Foreign Direct Investment: https://sdmx.org/?page_id=1747

The sub-group on SDMX for SEEA is composed of representatives from Eurostat, FAO, OECD, UNECE, UNSD and UN Environment. The overall scope of the work has been defined to cover the following SEEA priority accounts as mandated by the UN Statistical Commission:

1. Air emissions
2. Energy flows
3. Material flows
4. Water flows
5. Land (land use and land cover)

It is noted that some of these accounts will deliver indicators to monitor sustainable development goals (SDGs). In particular, the information in the material flow accounts and water flow accounts can be directly used for SDG indicators. The other priority accounts can also contribute to SDGs depending on what is the main data source for SDG indicators. The methodology and data sources for many SDG indicators are still under discussion (indicators Tier III) and thus there is an opportunity to have other SEEA aligned SDG indicators.

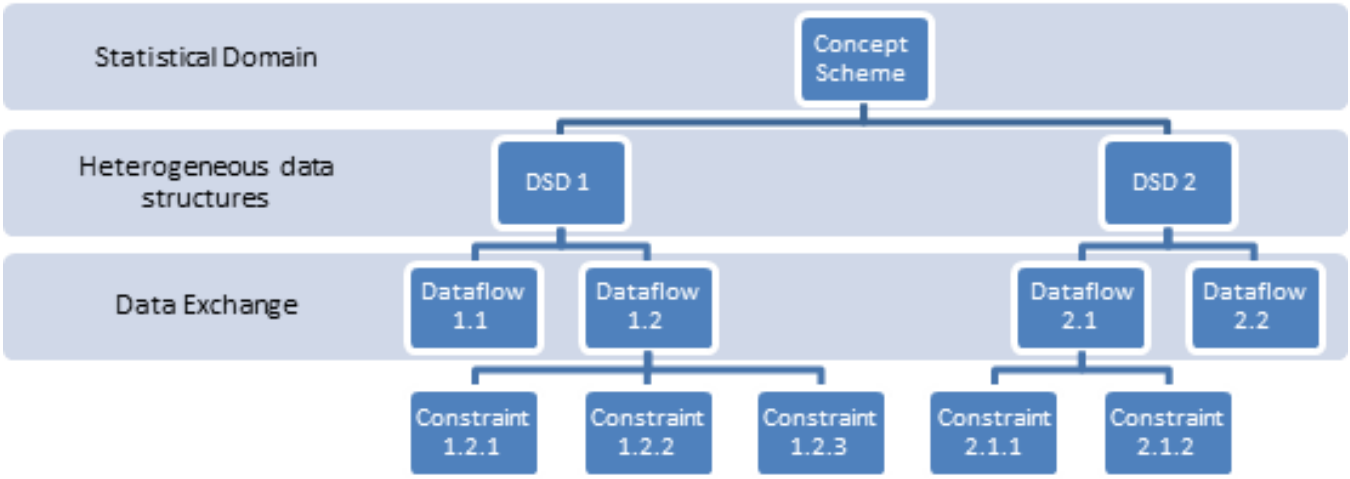
The scope of the work includes some reference metadata which is "close to the data" and is usually transmitted together with data. Other more high level reference metadata (e.g. methodological notes) are currently out of scope. This aspect would be picked up once discussion on Metadata Structure Definitions (MDSs) is elaborated for all domains within the MES area.

The process follows the generic "Checklist for SDMX Design Projects" that is available on the SDMX website: https://sdmx.org/?sdmx_news=checklist-for-sdmx-design-projects. An important milestone in every SDMX project is the pilot review phase. It consists of two phases: a content review and a technical review. The aim of the content review is that the project's *DESIGN* stage material is understandable, well-designed, and that it covers the scope of the project. The main deliverable for review is the "DSD matrix". The aim of the technical review is to test that the project's SDMX artefacts can be implemented at system level as efficiently as possible. It includes to test material such as example data messages and xml-schemas. The SDMX artefacts should be available in a registry, and the implementation & usage guidelines are sent for documentation. Further details on the process can be found in the checklist mentioned above.

# Content review information

## Implementation Model

The model chosen to implement SDMX for SEEA follows the basic model used also in SDMX for National Accounts. The approach is also expressed in the guidelines for "Modelling a Statistical Domain for Data Exchange in SDMX" available from the SDMX Statistical Working Group: https://sdmx.org/?page_id=4345#Modelling. The document suggests a generic implementation model pictured below:




Following this generic model, the outcome of the design project so far are:

- 1 Concept Scheme covering the area of SEEA
- 2 One Data Structure Definition per priority area (i.e. 5 DSDs in total).

Data Flow and Constraint artefacts are not yet created for the pilot review. Further mapping and creation of artefacts will only be done once the overall model is approved. Some examples are provided as part of the pilot package.

An early strategic decision for the data model was whether to extend the existing DSDs for National Accounts to serve also the needs of SEEA or to create separate DSDs for SEEA. It was decided for the latter, i.e. to develop DSDs not integrated with those for National Accounts but sharing as many possible concepts and code lists with them. This means that coding lists from National Accounts are also used for SEEA and extended whenever necessary with additional codes for SEEA. A consequence of this is that some coding lists include codes which are not necessary for SEEA (but are necessary for other macroeconomic statistics) and moreover the codes needed for SEEA may not be obvious to identify in the list. To facilitate this task, SEEA entries are identified in the literal, e.g. added entries in code list STO are identified with the text '(used by SEEA)'.

# SEEA SDMX Data Model

## General aspects (DSD Matrix)

Within a data message, dimensions are used to uniquely identify the observation. They are always mandatory. Attributes can be used to further describe the data. They can be linked to several levels (dataset, series, observation) and can be mandatory or optional. All concepts can be coded (linked to a code list) or be defined through a data type (e.g. String, DateTime,...).

The group has defined a 'DSD Matrix' as an inventory for structuring the relevant data flows in the area of environmental-economic accounts. The Matrix file summarised all concepts and code lists contained in the reporting framework. Those have been selected from various sources along the data flows documented in annex 1. The codification of domain concepts in the DSDs has to be descriptive and comprehensive enough to fulfil any current and future data exchange requirements within the scope of SEEA.

It contains the following sheets:

- **Overview DSDs**, showing the concepts used, their representation and the link to the DSDs;
- **Matrix sheet**, kept as a placeholder for showing the relationship between the data flows and concepts once implemented;
- **Code list sheets**, showing the contents of each of the code lists used and some additional comments like for instance integrity rules. Code lists marked in green are sourced from cross-domain code lists; code lists marked in yellow are macro-economic statistics shared code lists. Note that for some shared code lists, several codes on top are marked in yellow as well. Those are suggested additions which will be merged into the existing code lists upon completion of the pilot.

## SEEA Coding Model Overview

Four of the five priority SEEA accounts are about flows (air emissions, materials, energy and water) and one is about stocks and changes in stocks (land accounts). It was decided to distinguish flows from stocks (including inflows, outflows, opening stocks, closing stocks, etc.), and supply from use with the same approach as in the DSDs for national accounts. A concept named 'accounting entry' is used for this purpose. The same code list as in National Accounts is used. Thus a code 'Credit (resources)' (C) represents a supply of resources (e.g. in a supply table), a code 'Debit (uses)' (D) represents a use of resources (e.g. in a use table), a code 'Assets' (A) represents a stock of resources (e.g. in a land account), etc. It was decided to create two additional codes, for flows between the environment and the economy and between the economy and the environment. These codes do not exist in national accounts because the latter only consider flows within the economy. It is here that SEEA works as an extension of national accounts.

Unsurprisingly, the four flow accounts have many elements in common and differ from the stock account. For flow accounts:

- As regards the columns in a physical supply-use table, see e.g. SEEA CF Table 2.1 it was decided to create one single concept 'interactors' for industries and households (concept INTERACTORS) and to model the accumulation and theflows with the rest of the world as in the national accounts DSDs (using STO codes P.52 for accumulation; P.6 for exports and P.7 for imports). This approach is used in all four flow accounts. As in the SEEA CF, the item 'accumulation' has a different meaning for each account: for air emissions it corresponds to emissions from controlled landfills, for energy and water it corresponds to changes in stocks.

- One concept was created for the 'products' of each of the four flow accounts. This corresponds to the rows in a physical supply-use table, see e.g. SEEA CF Table 2.1. They are: AIRPOL for air emissions, ENERGY_FLOWS for energy flows, MATERIAL for material flows, WATER_FLOWS for water accounts, LAND_COVER and LAND_USE for land accounts. There may be cases in which the same product exists in two accounts, e.g. certain gases in AIRPOL also exist in MATERIAL; certain energy products in ENERGY_FLOWS may also exist in MATERIAL. In those situations one product or flow will have different code in each list. This creates no consistency problem as no SEEA account uses two of those lists simultaneously. The instances of products /flows existing in more than one list are identified in the coding lists. Having the same code for a given product or flow in several lists would break with coding structure of the individual lists.

- One concept was created to report bridging items (BRIDGE_ITEM), i.e. to reconcile between aggregated data for units resident in the economy and aggregated data for national units. This is used for air emissions, material flow accounts, water and energy.

- One concept was created to distinguish footprints (of air emissions, energy, materials, water or land), also known as demand-based measures, from production-based measures (DEMAND_PROD). These footprints represent what is emitted (or the environmental impact) both in the country and abroad because of the use of materials/goods/services within the country. Whereas this concept DEMAND_PROD, in combination with the other concepts in the model, is sufficient to identify univocally footprint data, in certain cases other specific solutions where created, in particular for units (tons in raw material equivalents, or T_RME, see below) or for aggregates (raw material consumption, or RMC, see below). This is for several reasons, among others: tradition (using units and aggregates which are already known and used by environmental accountants), transparency (i.e. allow simple identification of RME or RMC instead of deriving them from combinations of different codes in several concepts) and importance (as some aggregates such as RMC are SDG indicators and other important indicators). See details about using this code below, in section 'detailed coding instructions', e.g. for air emission accounts fifth dataset and material flow accounts fourth dataset. For energy, water and land this is a future-proof feature to allow for the reporting of footprints as currently there are no transmissions of energy footprints nor water footprints.

- Raw material equivalents used for imports and exports in material flow accounts are modelled as a new unit (entry T_RME in code list CL_UNIT). Trade flows expressed in RME are larger than the actual weight of the products traded: almost all products go through different stages of manufacturing, starting from the extraction of raw materials, then the transformation into raw products, followed by further processing and assembly into semi-manufactured products and finally into finished products. The mass weight of the extracted raw materials, e.g. gross ore, is generally much higher than the weight of the traded products that contain the processed material only. The initial purpose of raw material equivalents is to measure imports and exports of products using the weight of the raw materials which are needed to produce them. In this sense, raw material equivalents simply correspond to a new unit of measure. However raw material equivalents are also used to produce the aggregate raw material consumption (RMC), which is a footprint recorded as such along a specific dimension of the DSD. It is for this reason that raw material equivalents and footprints are distinguished in the DSD.

- One concept for product (PRODUCT) was necessary for breakdowns of footprints by product according to the classifications CPC/CPA, e.g. energy footprints by CPC. This is used for air emissions, energy flows and material flows.

- One concept was necessary for the geographical code of counterpart areas in imports and exports (AREA). This was necessary in particular in Europe, where questionnaires distinguish between intra EU and extra EU imports/exports. The same code list was used as in the DSDs for national accounts.

- It was decided to model the flows in land accounts closer to the stocks in land accounts than to the flows in the other accounts (see next).

For the land accounts (land cover and land use), which encompass flows and stocks:

- There is one code list for land cover and one for land use (CL_LAND_COVER and CL_LAND_USE).

- A second concept was necessary for double-entry tables of changes in land cover from one type to another (land cover change matrix, see SEEA CF Table 5.14). This is LAND_COVER_TO. It uses the same code list CL_LAND_COVER and concept LAND_COVER

- One concept was necessary to distinguish opening stocks of resources, additions to and reductions from stocks, and closing stocks. The same approach as in the DSDs for national accounts was followed. This is modelled in the concept NA_STO. In particular, opening stocks are coded with code LS in CL_NA_STO, closing stocks are coded with code LE, changes in stocks are coded with code LX. Additions to stocks are changes with positive value and reductions from stocks are changes with negative value.

## Concepts and Coding

Note that, as in the other DSDs within macro-economic statistics, we work with a single measure (OBS_VALUE). No multiple measures are used. Thus only dimensions and attributes are described below. For the coded dimensions and attributes, code lists have been taken as far as available in that order from:

- existing cross-domain code lists (e.g. CL_FREQ)
- existing code lists from other macro-economic statistics DSDs (e.g. CL_UNIT, CL_ACCOUNT_ENTRY). In three cases existing code lists where extended: first in CL_UNIT more physical units (i.e. non-monetary units) were added, such as 'Tons of CO2 equivalents'. Secondly, in CL_ACCOUNT_ENTRY, a few codes were added e.g. flows from the environment to the economy and from the economy to the environment. Thirdly, in CL_NA_STO a few codes were added for main aggregates such as 'Domestic Material Consumption'

- derived from existing classifications, e.g. CL_ENERGY_FLOWS derived from classifications of energy products used in energy statistics
- new code lists created by the group for the needs of SEEA, e.g. CL_AIRPOL with list of greenhouse gases and air pollutants, and other similar lists for the other SEEA priority accounts.

For the last three cases, the guidelines for the Creation and Management of SDMX Code Lists have been used (https://sdmx.org/?page_id=4345#CodeListGuideline). Some existing code lists have been extended. In order not to impact the existing DSDs, they have been created in the pilot registry with ESTAT as maintenance agency and a different version number for the purpose of the pilot. Once approved, the original Code Lists would be updated and used across all MES DSDs. These are marked in red in the DSD matrix for "draft maintenance agency / version for pilot use".

Consequently, the following concepts are used in the DSD for each environmental account:

**Air emission accounts**

DEMAND_PROD: distinguishes between production-based accounts and demand-based accounts (i.e. footprints).

ACCOUNTING_ENTRY: Generic national accounts list to distinguish assets and flows. Further distinguishes flows inside the economy, from the economy to the environment (and viceversa), incoming flows (credits) and outgoing flows (debits), etc. All the entries in air emission accounts are flows from the economy to the environment.

COUNTERPART_AREA: needed for international transactions. This is to attribute emissions to other economies in the case of footprints.

INTERACTORS: represents units in the economy or the environment. The former correspond to production units, broken down by ISIC/NACE, or households, broken down by activity type. In the case of air emission accounts, all economic units act as emitters and the environment absorbs the emissions.

BRIDGE_ITEM: bridging items to reconcile the economy defined in terms of resident units (as in the national accounts and air emission accounts) and units physically located on the territory (as in the UNFCCC inventories).

AIRPOL: list of greenhouse gases and air pollutants.

STO: list of transactions and aggregates from SNA with additional transactions and aggregates from SEEA. For air emission accounts this concept is only needed to attribute air footprints to types of final demand (final consumption expenditure, gross capital formation, exports, etc.). See detailed coding instructions below for air emissions dataset 5.

SECTOR: institutional sectors from SNA, needed in SEEA e.g. to distinguish between final consumption by government sector and final consumption by household sector

**Energy accounts**

DEMAND_PROD: distinguishes between production-based accounts and demand-based accounts (i.e. footprints). Currently, energy footprints are not compiled in Europe.

ACCOUNTING_ENTRY: distinguishes assets and flows. Further distinguishes flows inside economy, from economy to environment (and viceversa), incoming flows (credits) and outgoing flows (debits), etc. Only flows, of different types, used in energy accounts.

COUNTERPART_AREA: needed for international transactions (imports/exports)

INTERACTORS: represents units in the economy or the environment. The former correspond to production units, broken down by ISIC/NACE, or households, broken down by activity type.

BRIDGE_ITEM: bridging items to reconcile the economy defined in terms of resident units (as in the national accounts and energy accounts) and units physically located on the territory (as in energy balances).

ENERGY_FLOWS: list of energy-related natural inputs, energy products and energy-related residuals. This list is based on the *Standard International Energy Product Classification* (*SIEC*), supplemented and aggregated as needed. CL_ENERGY_FLOWS encompasses three types of flows (natural inputs, energy products, residuals), of which only the second type is covered by SIEC, as SIEC is a classification of energy products. Correspondingly, ENERGY_FLOWS has additional codes for natural inputs and residuals. Secondly, SIEC codes are aggregated because the groupings of natural inputs and residuals are more aggregated than the SIEC-4-digit level, which is not needed for energy accounts. The aggregation categories are the result of a discussion in Europe and are used for the European data collection. The compromise reached in Europe became a European legal act (Commission Delegated Regulation (EU) 2016/172).

STO: list of transactions and aggregates from SNA with additional transactions and aggregates from SEEA. For energy accounts this concept is only needed for the flows in the physical supply –use table corresponding to imports (P.7), exports (P.6), accumulation (P.52), transformation use of energy (TUTU) and final use of energy (TUEU). More about the latter two codes in the detailed coding instructions below.

SECTOR: institutional sectors from SNA. Not really needed for energy accounts but temporarily left in case the data model is adjusted.

**Material flow accounts**

DEMAND_PROD: distinguishes between production-based accounts and demand-based accounts (i.e. footprints).

ACCOUNTING_ENTRY: distinguishes assets and flows. Further distinguishes flows inside the economy, from the economy to the environment (and viceversa), incoming flows (credits) and outgoing flows (debits), etc. Only flows, of different types, used in material flow accounts.

COUNTERPART_AREA: needed for international transactions.

BRIDGE_ITEM: bridging items to reconcile the economy defined in terms of resident units (as in the national accounts and material flow accounts) and units physically located on the territory.

MATERIAL: list of materials, with a breakdown by stage of manufacturing for imports/exports.

STO: list of transactions and aggregates from SNA with additional transactions and aggregates from SEEA. MFA aggregates added for transparency and due to their importance (some of them are SDG indicators). Added entries in STO for the purpose of SEEA are identified with the text '(environmental accounts)' in the literal.

SECTOR: institutional sectors from SNA. Not really needed for energy accounts but temporarily left in case the data model is adjusted

**Water accounts (only flows)**

DEMAND_PROD: distinguishes between production-based accounts and demand-based accounts (i.e. footprints). Currently, water footprints are not compiled in Europe.

ACCOUNTING_ENTRY: distinguishes assets and flows. Further distinguishes flows inside the economy, from the economy to the environment (and viceversa), incoming flows (credits) and outgoing flows (debits), etc. Only flows, of different types, used in water flow accounts.

COUNTERPART_AREA: needed for international transactions (imports/exports)

INTERACTORS: represents units in the economy or the environment. The former are further distinguished between production units, broken down by ISIC /NACE, and households, broken down by activity types

WATER_FLOWS: list of water flows

**Land accounts**

DEMAND_PROD: distinguishes between production-based accounts and demand-based accounts (i.e. footprints). Currently, land footprints are not compiled in Europe.

ACCOUNTING_ENTRY: distinguishes assets and flows. Further distinguishes flows inside the economy, from the economy to the environment (and viceversa), incoming flows (credits) and outgoing flows (debits), etc. Both flows and stocks are covered in land accounts.

LAND_COVER and LAND_COVER_TO: lists of land cover types (origin/destination)

LAND_USE and LAND_USE_TO: lists of land use types (origin/destination)

**Generic concepts used in all DSDs:**

UNIT_MEASURE: unit of measure

UNIT_MULT: multiplier of unit of measure

FREQ: data frequency (for SEEA, usually annual)

OBS_STATUS: to distinguish normal estimates, provisional estimates, forecasts, etc.

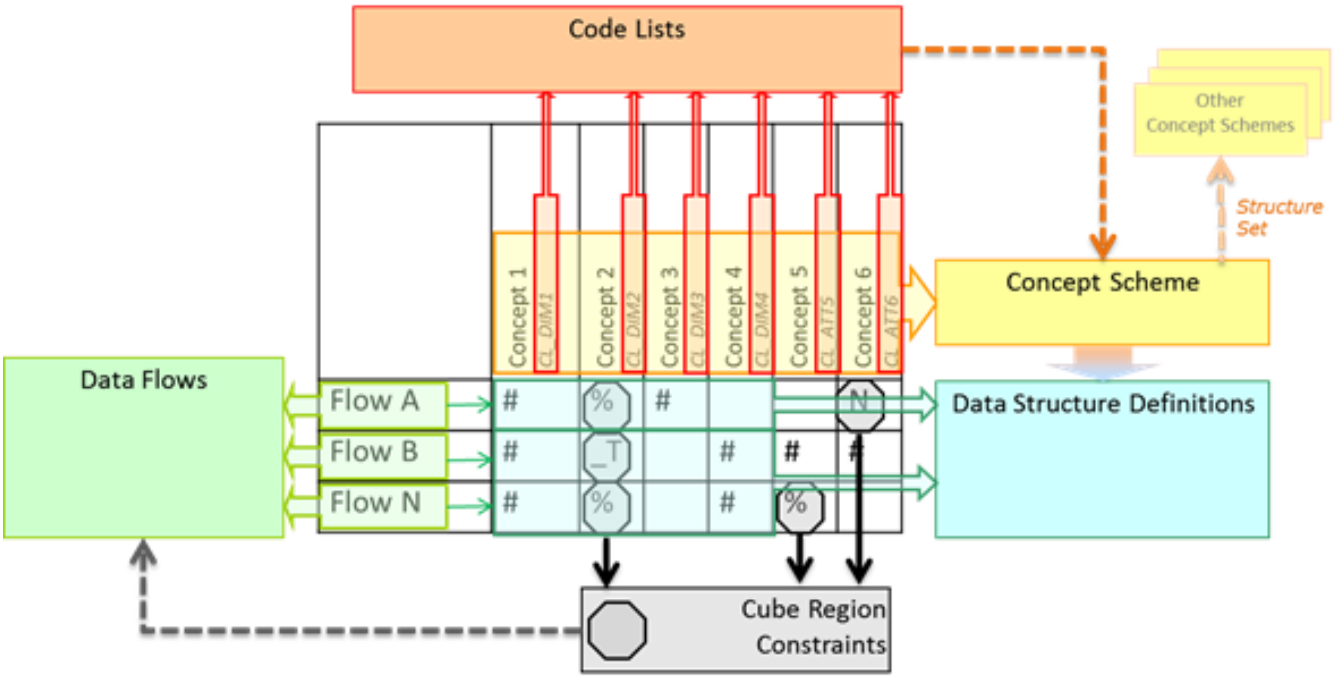
CONF_STATUS: to distinguish e.g. confidential data from non-confidential data

DECIMALS: number of decimal digits

# Pilot review information

## Link between the DSD matrix and the SDMX artefacts

Following the generic implementation model, the DSD matrix is linked to the SDMX artefact as follows:




For the pilot review, the matrix documents the Concepts Scheme, the DSDs and the Code Lists. Data Flows have been used to analyse the Concepts and Code Lists but will only be fully documented after the pilot review (see also annex 1 for data flows included for final implementation).

## SDMX artefacts

The pilot artefacts can be accessed through the Fusion Cloud registry (https://registry.sdmxcloud.org/). Once approved, the final artefacts will be published in the Global SDMX Registry (https://registry.sdmx.org/). Note that the interfaces look different, but both registries allow user interface as well as web service access. The DSD identifiers and URLs to the pilot DSDs are the following:

| **DSD ID** | **Pilot DSD version** | **REST URL** ***(SDMX 2.1, all references)*** |
|---|---|---|
| SEEAAIR | 0.2 | https://registry.sdmxcloud.org/ws/public/sdmxapi/rest/datastructure/ESTAT/SEEAAIR/0.2/?format=sdmx-2.1&detail=full&references=all&prettyPrint=true |
| SEEANRG | 0.2 | https://registry.sdmxcloud.org/ws/public/sdmxapi/rest/datastructure/ESTAT/SEEANRG/0.2/?format=sdmx-2.1&detail=full&references=all&prettyPrint=true |
| SEEAMFA | 0.2 | https://registry.sdmxcloud.org/ws/public/sdmxapi/rest/datastructure/ESTAT/SEEAMFA/0.2/?format=sdmx-2.1&detail=full&references=all&prettyPrint=true |
| SEEAWTR | 0.2 | https://registry.sdmxcloud.org/ws/public/sdmxapi/rest/datastructure/ESTAT/SEEAWTR/0.2/?format=sdmx-2.1&detail=full&references=all&prettyPrint=true |
| SEEALND | 0.2 | https://registry.sdmxcloud.org/ws/public/sdmxapi/rest/datastructure/ESTAT/SEEALND/0.2/?format=sdmx-2.1&detail=full&references=all&prettyPrint=true |

Note that the maintenance agency for the DSDs is currently Eurostat for purposes of the pilot only. The final maintenance agency for the DSDs might change depending on Ownership Group agreement after the pilot and before the DSDs are finally published in the global DSD registry. Other parameters of the REST URL can be modified as needed by the implementer.

# Detailed coding instructions

This section provides more detailed instructions for each of the SEEA priority accounts and the datasets for each account.

### Air emission accounts

The data model for air emission accounts is based on existing data collections and dissemination standards in place in Eurostat, OECD and FAO.

The SDMX data model considers five datasets related to air emissions. A first dataset is for data collections of air emission accounts. This is represented e.g. in SEEA CF Table 3.7. This table corresponds to the Eurostat questionnaire (although the Eurostat questionnaire has rows, columns and layers arranged differently this has no consequences for the SDMX encoding).

Moreover, the tables in the Eurostat questionnaire have a second part, at the bottom of each table, collecting bridging items. This is a reconciliation of the accounts, based on the residence principle from national accounts, with the territorial principle used in the greenhouse gas inventories and air pollution inventories. As said, in the Eurostat questionnaire, the bridging items are part of the same table, but for the purpose of the SDMX encoding it is helpful to imagine them as two different (sub)tables: one as SEEA CF Table 3.7 on the top, and one with the bridging items at the bottom.

The (sub)table at the top of the Eurostat questionnaire, corresponding to SEEA CF 3.7 and with no bridging items, should be encoded as follows:

- All measures are production-based, i.e. the accounts are broken-down by producers of the emissions. They are coded DEMAND_PROD=P.
- The types of substances in the rows of SEEA CF Table 3.7 correspond to the concept AIRPOL; the code list of air pollutants (CL_AIRPOL) encompasses codes in use by Eurostat, OECD and FAO.
- The columns of SEEA CF Table 3.7 follow the concept INTERACTORS. The item 'accumulation' in this list corresponds to emissions from controlled landfills.
- All entries in the table are flows from economic units into the environment and they are coded as ACCOUNTING_ENTRY=EC. This is true for all the air emission accounts.
- Other concepts with fixed values are: COUNTERPART AREA must be coded W2 meaning all the emissions are attributed to the domestic economy; PRODUCT=T meaning no breakdown by product; SECTOR=S1 meaning no breakdown by institutional sector; BRIDGE_ITEM=RES meaning emissions by residents; STO=_Z meaning not applicable i.e. there are no monetary national accounts transactions.
- The code list for units CL_UNIT is based on national accounts, but additional codes were created for tonnes of gases i.e. tonnes of CO2-equivalent, SO2-equivalent, NO2-equivalent and NMVOC equivalent.
- Other concepts that must be used as needed are the standard SDMX attributes (and the associated code lists) UNIT_MULT, DECIMALS, FREQ, OBS_STATUS, CONF_STATUS, with no particular instructions.
- ORGANISATION must be coded according to the receiving organisation.

The (sub)table at the bottom of the Eurostat questionnaire, corresponding to the bridging items, should be encoded as follows:

- The types of substances in the rows of SEEA CF Table 3.7 correspond to the concept AIRPOL.
- The different items of the concept BRIDGE_ITEM must be reported with the codes in CL_BRIDGE_ITEM.
- The bridging items are not reported separately for each economic activity nor the households. INTERACTORS takes the value ATU_HH (All NACE and households activities).
- Other concepts with fixed values are: DEMAND_PROD=P; ACCOUNTING_ENTRY=EC; COUNTERPART AREA=W2; PRODUCT=T; SECTOR=S1; STO=_Z.
- Other concepts that must be used as needed are: UNIT_MEASURE, UNIT_MULT, DECIMALS, FREQ, OBS_STATUS, CONF_STATUS, with no particular instructions.
- ORGANISATION must be coded according to the receiving organisation.

A second dataset corresponds to the air emission accounts for dissemination by Eurostat. This is encoded in the same way as SEEA CF Table 3.7 or bridging items explained above, with the only difference that there are additional unit codes in CL_UNIT for dissemination indicators such as tonnes of CO2-equivalents per euro at current prices (and also for SO2, NO2, NMVOC), tonnes of CO2-equivalents per euro in chain-linked volumes (and also for SO2, NO2, NMVOC) and tonnes of CO2-equivalents per capita.

A third dataset corresponds to air emission intensities for dissemination by Eurostat. This is encoded in the same way as SEEA CF Table 3.7 with the following differences:

- INTERACTORS may be ISIC/NACE industries.
- BRIDGE_ITEM=RES.
- STO can be P1 or B1G according to the reference used to calculate the intensities (output or value added).

A fourth dataset corresponds to emissions with breakdown by product (in sense of CPC/CPA), rather than by air pollutant. These are published by FAO. These are not demand-based estimates (footprints) but production-based estimates. The encoding would be as follows:

- DEMAND_PROD=P
- AIRPOL=T
- The concept PRODUCT follows the list CL_PRODUCT, which in turn has CPC/CPA codes.
- The concept BRIDGE_ITEM = TER.
- The concept UNIT takes the code KG of CO2E by KG of Product.
- All entries in the table are flows from economic units into the environment and they are coded as ACCOUNTING_ENTRY=EC. Other concepts with fixed values are: COUNTERPART AREA=W2;  SECTOR=S1; STO=_Z; The concept INTERACTORS must be valued Z.
- Other concepts that must be used as needed are: UNIT_MULT, DECIMALS, FREQ, OBS_STATUS, CONF_STATUS, with no particular instructions.
- ORGANISATION must be coded according to the receiving organisation.

A fifth dataset corresponds to emissions induced by final use (footprints). Presently there are no data transmissions of carbon footprints or pollutant footprints, but Eurostat disseminates air emission footprints. Air emission footprints should be encoded as follows:

- Footprints are primarily identified with the code DEMAND_PROD=D
- The concept AIRPOL takes values from CL_AIRPOL
- The concept INTERACTORS takes the values ATU (i.e. all ISIC/NACE producers) and HH (i.e. households) of CL_INTERACTORS
- STO could be coded using potentially any final demand code. Eurostat only uses P3, P5, P51G, P52, P53 and P6.
- Potentially there may be breakdowns by institutional sector of final use categories, e.g. final consumption by sector or gross fixed capital formation by institutional sector. This is coded together with STO. However not all possible combinations make sense, e.g. STO=P6 (exports) do not make sense broken down by sector (only the 'sector' rest of the world induces exports and their associated footprints). Eurostat only has sectoral breakdowns for STO =P3 (i.e. final consumption) and SECTOR can be coded S13 or S14 or S15. In any other case sector should be coded SECTOR=S1 meaning no breakdown by institutional sector
- The concept COUNTERPART AREA takes values from CL_AREA. If footprints are reported without geographical origin, COUNTERPART AREA must be coded W0 meaning that the footprint includes emissions embodied in imports and hence originating from elsewhere in the world. If values are reported for the domestic economy and rest-of-the-world, codes W2 (domestic/reference area) and W1 (rest of the world) should be used. In case the modelling framework is based on source data with multiple interlinked economies (e.g. so-called multi-regional input-output tables), COUNTERPART AREA may also use other codes from CL_AREA to identify the countries/regions where the emissions occur.

- Other concepts with fixed values are: All entries in the table are flows from economic units into the environment and they are coded as ACCOUNTING_ENTRY=EC.
- The concept PRODUCT takes the different values of CL_PRODUCT.
- BRIDGE_ITEM=RES meaning emissions by residents.
- The code list for units CL_UNIT is based on national accounts, but additional codes were created for tonnes of gases i.e. tonnes of CO2-equivalent, SO2-equivalent, NO2 equivalent and NMVOC equivalent.
- Other concepts that must be used as needed are: UNIT_MULT, DECIMALS, FREQ, OBS_STATUS, CONF_STATUS, with no particular instructions.
- ORGANISATION must be coded according to the receiving organisation.

## Energy accounts

The data model for physical energy flow accounts is based on existing data collections in place in Eurostat and upcoming work by UNSD. A simplified version of the tables in the existing data collections is represented in SEEA CF Table 3.5 (physical supply-use table for energy flows). This data model further breaks down the use table into two sub-tables for transformation use of energy and final use (see ACCOUNTING_ENTRY below).

- In such dataset, all measures are production-based, i.e. they are not footprints. They are coded DEMAND_PROD=P.
- The energy flows in the rows of SEEA CF Table 3.5 follow the concept ENERGY_FLOWS. The code list of energy flows (CL_ENERGY_FLOWS) encompasses codes in use by Eurostat and UNSD, both based on the international classification SIEC as regards energy products (to which natural inputs and residuals are added in the code list of energy flows).
- The columns ‘Production (including household production)’ in SEEA CF Table 3.5 follow the concept INTERACTORS. The column ‘accumulation’ follow the concept STO code P52; the column ‘flows from the rest of the world: imports’ follow the concept STO code P7; the column ‘flows to the rest of the world: exports’ follow the concept STO code P6.
- All entries in the supply table (Eurostat questionnaire, Table A) which are flows between economic units (i.e. not with the environment) are coded ACCOUNTING_ENTRY=C. The flows from the environment into the economy in the supply table are coded as ACCOUNTING_ENTRY=EN
- All entries in the use table (Eurostat questionnaire, Table B and C and sub-tables B.1, B.2) which are flows between economic units (i.e. not with the environment) are coded ACCOUNTING_ENTRY=D, with further sub-codes DT, DE and DER for transformation use and end use, as relevant. The flows to the environment in the use table are coded as ACCOUNTING_ENTRY=EC.
- Final use of energy products in the table must either be 'transformation use of energy' or 'final use of energy'. These are coded respectively TUTU or TUEU. Transformation use of energy consists of transforming energy products to create other energy products, e.g. crude oil into gasoline, whereas final energy use consists of using energy products to product energy. Note that before the creation of the codes TUTU and TUEU, the STO codes P2 (intermediate consumption) and P3 (final consumption) were proposed as proxies borrowed from national accounts. However they were imperfect proxies because 'transformation use of energy' and 'final use of energy' are not identical to P2 and P3, for instance burning gasoline to produce energy (to move cars) may be done by businesses for P2 or by families for P3. For this reason P2 and P3 are not to be used anymore for energy accounts, and instead TUTU or TUEU must be used for final use of energy products, depending on whether it is transformation use (TUTU) or not (TUEU). Note that total use TU is TUTU+TUEU.
- The concept COUNTERPART AREA is only relevant for imports in the supply table and for exports in the use table.
- The concept PRODUCT must be coded T (meaning no product breakdown; note the breakdowns in the physical supply-use table follow the concept ENERGY_FLOWS) and SECTOR must be coded S1 (meaning no institutional sector breakdown).
- Other concepts that must be used as needed are UNIT=TJ, UNIT_MULT, DECIMALS, FREQ , OBS_STATUS, CONF_STATUS, with no particular instructions.
- ORGANISATION must be coded according to the receiving organisation.
- The Eurostat questionnaire Table E on Bridging items is coded with the corresponding entries of concept BRIDGE_ITEMS and with DEMAND_PROD=P, ACCOUNTING_ENTRY=D, INTERACTORS=T, ENERGY_FLOWS=T, PRODUCT=T, SECTOR=S1, UNIT=TJ.

The data model could also accommodate energy footprints. This is future-proof, as there are presently no data transmissions of energy footprints. Footprints will be primarily identified with the code DEMAND_PROD=D.

## Material flow accounts

The data model for material flow accounts is based on existing data collections in place in Eurostat and UN Environment. The reporting of physical flow accounts for materials consists of several datasets. For convenience this is explained separately here although they may all be bundled as part of the same data transmission.

A first dataset is about domestic extractions broken down by material flows (input side of the material flows). This corresponds to the Eurostat questionnaire, table A. All the flows are extractions from the environment and there are no flows with other economies (imports/exports). In this case we are not considering measures in raw material equivalents (footprints).

- In such dataset, all measures are production-based, i.e. they are not footprints. They are coded DEMAND_PROD=P.

- The code list of material flows (CL_MATERIALS) encompasses codes in use by Eurostat and UN Environment. This code list also includes some memo items.

- All the flows are from the environment into the economy and are coded as ACCOUNTING_ENTRY=EN.
- STO must be coded MDE (domestic extraction) or BDMI (direct material inputs), as relevant.
- Units are measured in UNIT=TN (i.e. tonnes), UNIT_MULT=3, i.e. thousand tonnes.
- COUNTERPART_AREA must be coded _Z (meaning not applicable, as all flows are between the environment and the national economy).
- SECTOR must be coded S1 (meaning no institutional sector breakdown).
- BRIDGE_ITEM must be coded RES (meaning all the values are accounted following the residence principle).
- Other concepts that must be used as needed are DECIMALS, FREQ , OBS_STATUS, CONF_STATUS, with no particular instructions.
- ORGANISATION must be coded according to the receiving organisation.

A second dataset is about imports/exports of material flows. This corresponds to the Eurostat questionnaire, tables B and D. Whenever they are measured in 'normal terms', i.e. not as footprints, they should be encoded as follows:

- All measures are production-based. They are coded DEMAND_PROD=P
- The code list of material flows is CL_MATERIALS. In particular there are codes for imports/exports according to the stage of manufacturing of the traded product: raw, semi-finished and finished products.

- All the flows are between economic units (i.e. not with the environment) in different countries. Imports are coded ACCOUNTING_ENTRY=C, exports are coded ACCOUNTING_ENTRY=D. Physical trade balance is not reported but derived automatically.
- STO must be coded P6 (exports), P7 (imports) or PPTB (physical trade balance) as corresponds.
- Units are measured in UNIT=TN (i.e. tonnes), UNIT_MULT=3, i.e. thousand tonnes.
- COUNTERPART AREA must be coded according to the country of origin of imports or destination of exports, or W1 (rest of the world) if the country detail is not available.
- SECTOR=S1 (meaning no institutional sector breakdown).
- BRIDGE_ITEM must be in general coded RES (meaning all the values are accounted following the residence principle). However, material MF4 and its sub-items (Fossil energy materials/carriers) might be reported following the territory principle. Two different encoding scenarios have to be distinguished, both for imports and exports:

  For imports (following residence principle): The adjustment for the material code MF42 must not be reported (See table below, column Opt.1).

  For imports (following territory principle): The BRIDGE_ITEM code for MF4 and its sub-items must be TER and the adjustment for residence principle (MF42) must be reported with code RES_ABR, i.e. residents abroad (See table below, column Opt.2).

  For exports (following residence principle): The adjustment (MF42 and its breakdown) must not be reported (See table below, column Opt.3).

  For exports (following territory principle): the BRIDGE_ITEM code for MF4 and its sub-items must be TER and the adjustment for residence principle (MF42) must be reported with code TER_NRES i.e. residents abroad (See table below, column Opt.4).

| SDMX | MATERIAL | **Imports (Table B)** | | **Exports (Table D)** | |
|---|---|---|---|---|---|
| MF4 | **MF.4 Fossil energy materials/carriers, raw and processed** | Opt. 1 | Opt. 2 | Opt. 3 | Opt. 4 |
| MF41 | MF.4.1 Coal and other solid energy materials/carriers | RES | TER | RES | TER |
| MF411 | MF.4.1.1 Lignite (brown coal) | RES | TER | RES | TER |
| MF412 | MF.4.1.2 Hard coal | RES | TER | RES | TER |
| MF413 | MF.4.1.3 Oil shale and tar sands | RES | TER | RES | TER |
| MF414 | MF.4.1.4 Peat | RES | TER | RES | TER |
| MF42 | MF.4.2 Liquid and gaseous energy materials/carriers | RES | TER | RES | TER |
| MF421 | MF.4.2.1 Crude oil, condensate and natural gas liquids (NGL) | RES | TER | RES | TER |
| MF422 | MF.4.2.2 Natural gas | RES | TER | RES | TER |
| MF42 | MF.4.2.3 Fuels bunkered (Imports: by resident units abroad; Exports: by non-resident units domestically) | not reported | RES_ABR | not reported | TER_NRES |
| MF42 | MF.4.2.3.1 Fuel for land transport | not reported | RES_ABR_LTR | not reported | TER_NRES_LTR |
| MF42 | MF.4.2.3.2 Fuel for water transport | not reported | RES_ABR_WTR | not reported | TER_NRES_WTR |
| MF42 | MF.4.2.3.3 Fuel for air transport | not reported | RES_ABR_ATR | not reported | TER_NRES_ATR |
| MF43 | MF.4.3 Products mainly from fossil energy products | RES | TER | RES | TER |

- Other concepts that must be used as needed are: DECIMALS, FREQ , OBS_STATUS, CONF_STATUS, with no particular instructions.
- ORGANISATION must be coded according to the receiving organisation.

A third dataset is about output flows from the economy to the environment, balancing items and other indicators (e.g. domestic processed output, net additions to stock). This corresponds respectively to tables F, G and H in Eurostat's questionnaire. This is encoded as follows:

- All measures are production-based, i.e. they are not footprints. They are coded DEMAND_PROD=P.
- The code list of material flows is CL_MATERIALS. This code list includes not only codes for material inputs but also for the categories of domestic processed outputs and balancing items.
- Most flows are from the economy to the environment and are coded ACCOUNTING_ENTRY=EC. Balancing items in the input side are also coded ACCOUNTING_ENTRY=EC. Flows from the environment to the economy (such as oxygen for combustion, water for drinking) are coded ACCOUNTING_ENTRY=EN. Balancing items in the output side are also coded ACCOUNTING_ENTRY=EN. Net additions to stock are coded ACCOUNTING_ENTRY=D.
- STO can be coded MDPO (domestic processed output), KNAS (net additions to stock) or BBI (balancing items), as needed.
- Units are measured in UNIT=TN (i.e. tonnes), UNIT_MULT=3, i.e. thousand tonnes.
- COUNTERPART AREA=_Z
- SECTOR=S1 (meaning no institutional sector breakdown).
- Other concepts that must be used as needed are DECIMALS, FREQ , OBS_STATUS, CONF_STATUS, with no particular instructions.
- ORGANISATION must be coded according to the receiving organisation.

A fourth dataset is about material flow accounts in raw material equivalents (MFA-RME). The data model accommodates the needs of Eurostat and UN Environment. The relevant flows are recorded in Table I of Eurostat's questionnaire. For imports and exports, RME means that the products crossing the border are not measured as the weight of these products but as the weight of the natural resources needed to produce them. These raw material equivalents represent the material footprint of these products. Building on imports and exports measured in RME, other indicators can be constructed, such as raw material consumption and raw material input (codes for them have been created in CL_STO). They should be encoded as follows:

- All measures are consumption-/demand-based. They are coded DEMAND_PROD=D.

- The code list of material flows is CL_MATERIALS.

- ACCOUNTING_ENTRY can take several codes. Imports are coded ACCOUNTING_ENTRY=C and exports are coded ACCOUNTING_ENTRY=D. RMC and RMI are coded ACCOUNTING_ENTRY=D (i.e. demand perspective).
- STO can be coded P6 (exports), P7 (imports), BRMC (RMC) or BRMI (RMI), as needed. If available, subcategories of RMC can be coded using STO codes P3, P5, P51G, P52 and P53.
- Units are measured in UNIT=TN_RME (i.e. tonnes of raw material equivalents), UNIT_MULT=3, i.e. thousand tonnes of raw material equivalents.
- COUNTERPART_AREA must be coded according to the country of origin of the materials if this information is available. If no geographical breakdown is provided, COUNTERPART_AREA=W0.
- Potentially there may be sectoral breakdowns of final use categories, e.g. final consumption or gross fixed capital formation by institutional sector. This is coded together with STO. Eurostat only has sectoral breakdowns for STO=P3 (i.e. final consumption). In this case, SECTOR can be coded S13 or S14 or S15. In any other case, sector should be coded SECTOR=S1, meaning no breakdown by institutional sector.
- Other concepts that must be used as needed are: DECIMALS, FREQ , OBS_STATUS, CONF_STATUS, with no particular instructions.
- ORGANISATION must be coded according to the receiving organisation.

## Water accounts (only flows)

The data model for physical water flow accounts is based on methodological work by UNSD. A simplified version of the tables in the existing data collections is represented in SEEA CF Table 3.6 (physical supply-use table for water flows).

- In such dataset, all measures are production-based, i.e. they are not footprints. They are coded DEMAND_PROD=P.
- The water flows in the rows of SEEA CF Table 3.6 follow the concept WATER_FLOWS. The corresponding code list is CL_WATER_FLOWS.
- The columns ‘Production (including household production)’ in SEEA CF Table 3.6 follow the concept INTERACTORS. The column ‘accumulation’ follow the concept STO code P52; the column ‘flows from the rest of the world: imports’ follow the concept STO code P7; the column ‘flows to the rest of the world: exports’ follow the concept STO code P6.
- All entries in the supply table which are flows between economic units (i.e. not with the environment) are coded ACCOUNTING_ENTRY=C. The flows from the environment to the economy in the supply table are coded as ACCOUNTING_ENTRY=EN.
- All entries in the use table which are flows between economic units (i.e. not with the environment) are coded ACCOUNTING_ENTRY=D. The flows to the environment in the use table are coded as ACCOUNTING_ENTRY=EC.
- The concept COUNTERPART AREA is only relevant for imports in the supply table and for exports in the use table.
- The concept UNIT must be coded M3 (cubic meters of water); BRIDGE_ITEMS must be coded as TER, meaning production and consumption activities on the territory; PRODUCT must be coded T (meaning no product breakdown); SECTOR must be coded S1 (meaning no institutional sector breakdown); STO=_Z (meaning not applicable i.e. there are no monetary national accounts transactions).
- Other concepts that must be used as needed are: UNIT_MULT, DECIMALS, FREQ , OBS_STATUS, CONF_STATUS, with no particular instructions.
- ORGANISATION must be coded according to the receiving organisation.

The data model could also accommodate water footprints. This is future-proof, as there are presently no data transmissions of water footprints. Footprints will be primarily identified with the code DEMAND_PROD=D.

## Land accounts

The data model for land accounts (physical account for land use, land use change matrix, physical account for land cover, and land cover change matrix) is based on existing data sets and upcoming improvements by FAO. The code lists of land cover and land use encompass codes in use by this organisation. Data structure definitions are created for four datasets. For convenience they are explained separately although they may all be bundled as part of the same data transmission.

A first dataset is a physical account for land cover, see e.g. SEEA CF Table 5.13. Rows correspond to opening stock, additions to stock, reductions in stock, and closing stock. The structure corresponds to a balance sheet in National Accounts. Columns correspond to land cover types. This is reported as follows:

- All measures are production-based, i.e. they are not footprints. They are coded DEMAND_PROD=P.
- The row entries are reported with the following combinations of concepts ACCOUNT_ENTRY, STO and sign of the value:

| | ACCOUNT_ENTRY | STO | Value |
|---|---|---|---|
| opening stock | N | LS (opening balance) | Must be positive |
| additions to stock | AI | LX (changes in balance) | Must be positive |
| reductions in stock | AD | LX (changes in balance) | Must be negative |
| closing stock | A | LE (closing balance) | Must be positive |
| changes in stock | N | LX (changes in balance) | Can be positive, negative, or zero |

- The types of land cover follow the concept LAND_COVER.
- LAND_COVER_TO must be coded T.
- The concept UNIT must be coded HA (hectare) or other unit of area. Land cover can only be measured in surface units.
- Other concepts that must be used are: UNIT_MULT, DECIMALS, FREQ , OBS_STATUS, CONF_STATUS, with no particular instructions.
- ORGANISATION must be coded according to the receiving organisation.

A second dataset is a land cover change matrix, see e.g. SEEA CF Table 5.14. This table has a typology of land cover in rows and columns, and changes happening during the year are reported inside the matrix, as floows:

- All measures are production-based, i.e. they are not footprints. They are coded DEMAND_PROD=P.

- The types of land cover in rows follow the concept LAND_COVER. The types of land cover in columns follow the concept LAND_COVER_TO. Both concepts use the same code list CL_LAND_COVER.
- ACCOUNT_ENTRY must be coded to A (net acquisition of assets); STO must be coded as LX (changes in balance).
- The rwo entries are reported with the following combination of concepts ACCOUNT_ENTRY, STO and sign of the value:

| | ACCOUNT_ENTRY | STO | Value |
|---|---|---|---|
| opening stock | N | LS (opening balance) | Must be positive |
| additions to stock | AI | LX (changes in balance) | Must be positive |
| reductions in stock | AD | LX (changes in balance) | Must be negative |
| closing stock | A | LE (closing balance) | Must be positive |
| changes in stock | N | LX (changes in balance) | Can be positive, negative, or zero |

---

- The types of land cover follow the concept LAND_COVER.
- LAND_COVER_TO must be coded T.
- The concept UNIT must be coded HA (hectare) or other surface unit.
- Other concepts that must be used as needed are: UNIT_MULT, DECIMALS, FREQ , OBS_STATUS, CONF_STATUS, with no particular instructions.
- ORGANISATION must be coded according to the receiving organisation.

A third dataset is about physical account for land use. It is identical to the first dataset above but using the concept LAND_USE and the code list CL_LAND_USE instead of LAND_COVER and CL_LAND_COVER, respectively. The only possible unit for land use is UNIT=HA. This dataset is found in the SEEA Agriculture, Forestry and Fisheries.

A fourth dataset is about land use change matrix. This dataset is not yet used. It is identical to the second dataset above but using the concepts LAND_USE and LAND_USE_TO as well as the code list CL_LAND_USE. Moreover, measurement units for land use can be hectares (UNIT=HA), another surface unit, or a monetary unit for monetary estimates.

# Annex

## Annex 1 - Data Flow documentation

All SEEA related data flows (marked in green) shown in the diagram below have been used to analyse current data exchanges between countries and international organisations. Non-SEEA flows (marked in blue) have been added for documentary purposes.