



DEPARTMENT OF ECONOMIC AND SOCIAL AFFAIRS
STATISTICS DIVISION
UNITED NATIONS

ESA/STATISTICS/AC.228
EGM-FDES/1/9

**Expert Group Meeting on the Revision of the Framework
for the Development of Environment Statistics (FDES)
New York, 8-10 November 2010**

Towards a Sound Data Basis for Environment Statistics: Requirements for Multiple Use of Data and Institutional Cooperation

Michael Nagy
Environment Agency Austria

Eva Milota
Statistics Austria

Introduction

The FDES sets out the scope of environment statistics by relating the components of the environment to information categories that are based on the recognition that environmental problems are the result of human activities and natural events reflecting a sequence of action, impact, and reaction. The main objective of environment statistics is to provide statistics, indicators and meta-information on environmental pressures and the state of the environment to support the development, implementation, monitoring and evaluation of environmental policies on various levels and to inform the public. Environmental data are compiled, stored and disseminated in different contexts for different purposes at different institutions like .e.g. control and enforcement of environmental law by national and local government departments, monitoring of the status of the environment by government departments and research institutes, surveys carried out by NSIs for statistical purposes etc.

Environment statistics should provide a sound data basis (i.e. the bottom of the information pyramid) which can be used for multiple purposes (research and analysis, indicator development, integration with social and economic statistics etc.). Ideally, the data produced in the different contexts by the different institutions are complementary to each other and the standard attributes of high quality statistics (relevance, accuracy, timeliness, accessibility, interpretability and coherence) are fully achieved. In reality, for a certain number of reasons, this is not the case. There is on the one hand side duplication of work and on the other side lack of cooperation between institutions as well as the problem that statistical concepts and quality standards often are not applied in administration and research institutions.

Multidisciplinary of environment statistics and dispersed sources

The mandate to produce environment statistics is with National Statistical Institutes (NSIs) and in many countries this is organizationally reflected in the existence of a Department or Unit on Environment Statistics. There are still countries without a specific Department/Unit dealing with environment statistics and the compilation of environment statistics lies within other Departments/Units (e.g. agriculture statistics, land use statistics, national accounts etc.). Speaking institutionally it can be said that environment statistics is a relative young discipline compared to other statistics carried out by NSIs, such as business statistics, population statistics, national accounts etc.).

Environment statistics are statistics that describe the state and trends of the environment, covering the media of the natural environment (air/climate, water, land/soil), the biota within the media, and human settlements (see [OECD Glossary of Statistical Terms](#)). The 1984 FDES recognizes environment statistics as multi-disciplinary with dispersed sources and a variety of methods applied in the compilation (see FDES 1984). It furthermore recognizes that a number of producers (mainly government departments, central statistical services, research institutes etc.) of data are also the major users of environment statistics, next to users from businesses and industry, mass media or the general public.

This variety of data sources, methods and the required multidisciplinary may lead to:

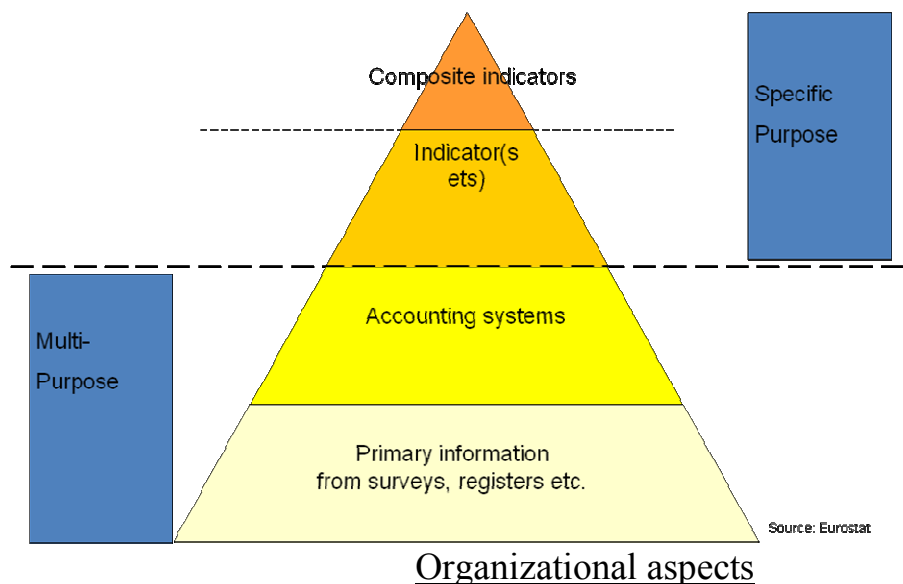
- a) Overlaps in terms of responsibility (within NSI or with other institutions), which have to be overcome from an organizational perspective (inter-institutional steering of environment statistics, clarification of responsibilities, priorities, budget, legal background etc.).
- b) Data quality problems (all dimensions) as external data providers typically generate the data within another context and for a different purpose than for environment statistics. Even data collected within an NSI might not necessarily fit the needs of environment statistics when the context is different (e.g. data from agriculture statistics, business statistics or national accounts used for environment statistics).

Both the organizational aspect and the data quality aspect have to be taken into consideration when developing environment statistics from scratch or when improving existing systems.

Figure 1 shows the classical “information pyramid” where a sound data basis (primary information from different sources” fits all purposes. This picture reflects an ideal situation. There are 2 major problems related to this:

- Primary information is often produced for a single specific purpose and thus does not qualify other purposes
- The data considered as primary information is not always coherent and complementary.

Figure 1: Classical „Information pyramid“



When considering organizational aspects of environment statistics potential data providers can roughly be grouped as follows:

- National Statistical Institutes and their related bodies (e.g. regional offices)

- Public Administration (e.g. Ministries) and related bodies (e.g. Environment Agencies, National Geographical Institutes, local authorities)
- Research Institutes (public and private)
- Private or semi-private interest groups (e.g. Water Association, Chamber of Commerce, Agriculture Lobby Groups)
- Non-Governmental Organizations (NGOs)
- Private companies

The goal of official statistics (and thus of official environment statistics) is to produce relevant, objective and accurate statistics to answer the public and users demands for better access to information, assisting good policy and decision-making. It is within the mandate of the National Statistical Institute to produce the official environment statistics, even if there is no separate department of environment statistics. NSIs are usually not the only official and national institutes to produce official statistics. Central banks and some ministries or other central authorities may have statistical functions as well (but not all of them may be relevant for environment statistics). Together, all producers of official statistics form the statistical system of a country. In some countries, especially those that have a federal structure, producers of official statistics exist also at regional or even municipal levels. The NSI is in charge of the coordination between statistical producers and of ensuring the coherence and compliance of the statistical system to the principles. The driving force behind the generation of primary statistical data by NSIs (e.g. via survey techniques) is the statistical law (with a statistical programme) or a similar legal act.

The drivers for primary data production by other institutions than NSIs are heterogeneous:

- Public administration has to enforce laws, produce reports, develop new legal acts, support policy making and to inform and serve the public.
- Research institutes carry out research and development, teaching and can be involved in policy support
- Private, semi-private interest groups and NGOs want to support their clients and influence policy making into a specific direction
- Private companies want to make business with data or have a specific interest that their data is used for official statistics

When setting up a statistical system for environment statistics all potential data providers should be taken into consideration. The decision whether data from an external data provider (other than NSI) can be used for official environment statistics has to be taken on the basis of the following criteria:

- a) Is data (nationally) representative?

- b) Is the data collection/production part of a structured data collection system which is part of a series? Is the data collection repeated regularly?
- c) Are sound statistical principles applied or are the applied principles consistent/comparable with statistical principles?

Generally it can be said that it will be NSIs and the public administration which are able to fulfill the above criteria. However, research institutes and interest groups (such as water associations) could provide important input to develop methods for data gap filling and grossing-up of data from sample surveys (e.g. by application of resource use or emission factors).

There will be only exceptional cases where data produced by private companies might qualify for official environment statistics.

Once data sources for official environment statistics are identified, other organizational aspects (which are often linked to data quality issues) have to be taken into consideration:

- **Duplication of data flows:** For example in many countries manufacturing industries have to report annually their water uses to the water administration (Ministry of Environment, Ministry of Water, local authority or Environment Agency) for environmental control reasons. The reporting units are defined by the water use permit and the activities are classified according to the main water using production process. On the other hand there is an annual statistical survey on water use and wastewater discharged covering the same industries plus some others which is carried out by the NSI. This survey uses ISIC classification and the reporting units are individual establishments. There is a clear duplication of data flows, a double reporting burden to some enterprises and their establishments. As there are different classifications and reporting units in the two data collections and the statistical law forbids to provide individual (non-aggregated) data to other governmental bodies the problem cannot be solved by just omitting one of the data collections.
- **Contradictory results:** Even within the public administration it can happen that data collected on the same topic but by different institutions do not match or lead to contradictory results. For example data on sewage sludge generation, reuse and disposal is collected by different governmental institutions for different reasons:
 - a) Governmental body responsible for urban wastewater treatment (typically Ministry of Environment, Department on Water Management): Goal is to ensure that the generated sewage sludge is treated, reused and disposed according to the law. Data on volumes or weight of sewage sludge generation, treatment, reuse and disposal is produced.
 - b) Governmental body responsible for Agriculture (Ministry of Agriculture): Goals is to ensure that only sewage sludge of a certain quality is used in

agriculture. Data on quality, volumes or weight of sewage sludge used in agriculture is produced

- c) Governmental body responsible for waste generation and waste disposal (Ministry of Environment, Department on Waste): Goals is to ensure that only waste (including sewage sludge) of a certain quality after a certain treatment is disposed or used to cover landfills. Data on quality, volumes or weight of sewage sludge disposed on landfills is produced.

It would be naïve to put the figures of the 3 governmental bodies and to consider them coherent or complementary and to use them directly for environment statistics. It could for example happen that the total number of sewage sludge disposed plus sewage sludge used in agriculture results in a higher number than sewage sludge generated. On a first glance this seems impossible. In practice there are usually big differences in the numbers which may be explainable, but it needs a good technical understanding to explain the differences. Reasons could be problems with the unit of measurement (wet/dry weight), temporal storage, change of weight due to decomposition processes, adding of sludge from other industries before disposal etc.). Therefore conventions and standards are to be defined to overcome such an issue and to “release” the three mentioned governmental bodies from their apparent competition for the most accurate data.

- **Institutional data sharing** (on the governmental level): In most countries governmental bodies are obliged to share data. However, in practice this often does not work because of a number of reasons:
 - There is no legal background for a particular exchange of data: e.g. parts of environment statistics are not legally based (e.g. voluntary reporting of some data sets to international organization such as Eurostat, OECD, UNSD). As there is usually no national legal background for this kind of reporting, potential data providers (e.g. Environment Agencies) see no reason to provide data for such a purpose to the NSI.
 - Additional efforts (and budget constraints) to provide data in the needed format: Usually data needs to be manipulated before handing it over to other institutions (in this case: NSI). This requires additional manpower on the side of the data producer for which no additional budget is available.
 - Unclear legal situation in individual cases and strategic issues: Who is the owner of the data? Will an existing (voluntary) data collection be endangered if data is handed over to a third party (NSI)?
 - It works in one direction: from public administration to the NSI, but the rule of statistical confidentiality does not allow it in the other direction.

Data quality aspects

When talking about data quality aspects this has to be discussed on the basis of the main quality parameters of official statistics (relevance, accuracy, accessibility, coherence, interpretability, timeliness).

Even if it can be taken for granted that data produced by the public administration are of certain relevance (for policy making in a certain sector – e.g. water use or air emissions) there are often problems related to representativeness or coverage of the data. For example the European Pollutant Release and Transfer Register (ePRTR) is one of the most important sources of information on industrial emissions to air, water and soil. EU Member States are obliged to maintain such registers and to provide emission data for individual industrial reporting units (which are not necessarily identical to “establishments”) and to update the data annually. The register covers big industries above given threshold values (depending on production parameters) and intends to cover 80-90% of the national emissions of a given pollutant. The ePRTR is not representative (but covering dominating emissions) and its reporting units are not statistical reporting units. However, since the activities are defined according to ISIC (NACE) it is possible to use it for environment statistics and to complete information gaps (e.g. on smaller enterprises/establishments) with the help of sample surveys or calculations.

Timeliness and accessibility of data can be a problem in federal countries. If data is collected by a local authority, then this could involve a number of steps of quality assurance, data manipulation etc. until the data is centrally available. In some cases there might not be any kind of data transmission to one single central body; therefore the NSI has to contact a number of local bodies to obtain the data. All these steps need time or it may be impossible to receive data from the local level.

Primary data produced for other purposes than for environment statistics may not fulfill accuracy requirements or are not suitable for methodological reasons. This could be the case when e.g. samples on air or water emissions are taken only for the purpose to measure the exceeding of a legally binding threshold value. Such samples are often taken at periods of high production but are not representative for a year.

National data collection strategy as a solution

Development of Environment Statistics is to be considered as one important, but not the only use of environmental data. Thus, NSIs and their mandate to provide official statistics are key stakeholders, next to public administration or even research institutes. Ideally, the different data producers produce complementary data, apply classifications which can be transformed to each other, have common QA/QC procedures and share their data without restrictions (taking into consideration existing law).

Guiding rules to develop such a strategy can for example be derived from the European Strategy on a Shared Environmental Information System (SEIS, European Commission 2008):

- Information should be managed as close as possible to its source;
- Information should be collected once, and shared with others for many purposes;
- Information should be readily available to public authorities and enable them to easily fulfill their legal reporting obligations;

- Information should be readily accessible to end-users, primarily public authorities at all levels from local to European, to enable them to assess in a timely fashion the state of the environment and the effectiveness of their policies, and to design new policy;
- Information should also be accessible to enable end-users, both public authorities and citizens, to make comparisons at the appropriate geographical scale (e.g. countries, cities, catchments areas) and to participate meaningfully in the development and implementation of environmental policy;
- Information should be fully available to the general public, after due consideration of the appropriate level of aggregation and subject to appropriate confidentiality constraints, and at national level in the relevant national language(s); and
- Information sharing and processing should be supported through common, free open source software tools.

In order to develop such an inter-governmental strategy a several steps approach is needed, this includes the identification of the main data producers of environmental data, high-level agreements by the relevant Ministers (e.g. via Memoranda of Understandings) and other high-level representatives of the main stakeholders as well as the set-up of a steering Committee.

There should be a common framework which is consistent with FDES (e.g. DPSIR) and the existing data collections should be checked against this framework, and modified if needed.

It is also important that all data producers use temporal and spatial aggregations which fit all purposes of the data. In environment statistics the temporal resolution is usually the calendar year.

Reporting units used by the different stakeholders should be coherent so that data can be combined at an agreed aggregation level.

Statistical classifications (such as ISIC) are not necessarily to be applied by administrative data producers or by research institutes. But there should be common coding systems in place which allow the linkage of data from outside NSIs e.g. directly with data of the business register. Then all attributes which are available at the NSI can be added to a given reporting unit.

Measured values have to be representative or the representative data have to be indicated (metadata).

Maintaining metadata in the information system is crucial for a proper use of the data by all potential users.

Additionally, e-Government strategies could support the development of technical solutions for data sharing, development of standards and rules going beyond the needs of environment statistics and thus allowing more complex analysis of the data.

Examples

Water Emission Statistics

The European Water Framework Directive requires the EU Member States to develop registers on significant emissions to water. On the other hand the European Pollutant Release and Transfer Register also requires data on pollutant emission to water (air and soil).

Within the EDM-Strategy (Electronic Data Management) one single portal for data input by operators of industrial establishments has been created. They report only once their emissions to water¹, but the data is used for many purposes such as:

- Reporting under the EU Water Framework Directive
- Reporting under the EU Urban Wastewater Treatment Directive
- Working out the national River Basin Management Plan
- Information of the public via the Austrian Water Information System (WISA)
- Emission data for environment statistics to Statistics Austria

A number of agreements had to be made and legal acts had to be modified to make the system work. There is a high-level steering committee (involving stakeholders from local authorities and national authorities) deciding upon strategic issues and financing and technical working groups dealing with technical and legal aspects.

QA/QC procedures involve experts from local authorities, the Austrian Environment Agency and other national authorities. QA/QC processes are part of the electronic system and quality assured data is clearly indicated.

Statistical classifications and statistical units are considered in this electronic reporting system and thus data can be used for statistical purpose. As data is not representative (i.e. covering industries above threshold values as defined by environmental law) data gap filling and grossing up for statistical purposes is done by the Austrian Environment Agency applying emission factors provided by experts of the Technical University of Vienna.

Integrated NAMEA

The Austrian “Integrated NAMEA” (NAMEA – National Accounting Matrix including Environmental Accounts) follows the fundamental idea of a NAMEA - that is to adjust the classifications of economic and environment-related data to enable a direct comparison of parameters from both domains. Using a standardized classification of industries (NACE classification) and households, economic indicators such as the

¹ Next to water emissions they have to report on waste (generation, transport, treatment, disposal) and air emissions. The system will be further developed for radioactive material and it is also possible to include water abstraction and water use.

production value or labour force can be linked to environmental material flows and/or environmental expenditure of a given branch. In its simplest form a NAMEA refers to a specific environmental aspect, e.g. air emissions. In contrast an integrated NAMEA is a complex model comprising data of several environmental domains, which are compared with economic parameters.

The economic accounts of the Austrian integrated NAMEA comprise the production value, gross value added and labour force in full-time equivalences. The environmental data consist of the modules material input, energy consumption, air emissions, environmental protection expenditure (for protection of ambient air and climate, and for waste management), environmental taxes and wastes.

Data sources are:

Statistics Austria

- National Accounts for
 - production value
 - gross value added
 - labour force in full-time equivalences
 - environmental taxes
- Environmental Accounts for
 - material inputs (material flow accounts)
 - environmental protection expenditure for protection of ambient air and climate, and for waste management (environmental protection expenditure accounts)
- Energy Accounts for
 - energy consumption

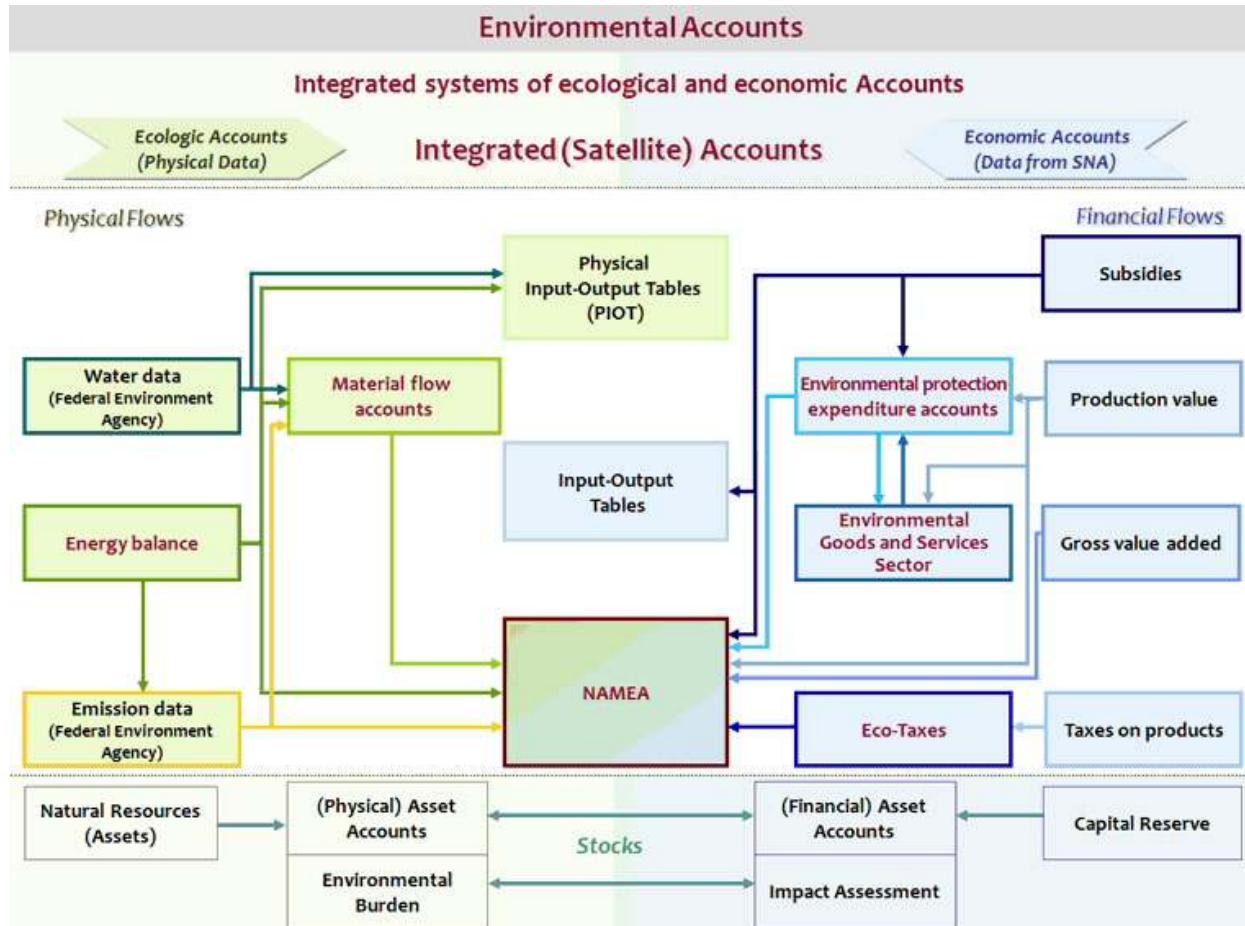
Federal Environment Agency

- National Greenhouse Gas Inventories and
- National Air Emission Inventory for
 - air pollutants such as (amongst others) NO_x, SO₂, NMVOC, NH₃ and CO, Particulate Matter (PM), Persistent Organic Pollutants (POPs), CO₂
- Austrian waste management data
 - hazardous wastes
 - non hazardous wastes

Data for Environmental Accounts are based on different data sources such as:

- structural business statistics (from Statistics Austria)
- short term statistics (from Statistics Austria)
- statistics for agriculture, forestry and fishing (from Statistics Austria)
- overall agricultural and forestry accounts (from Statistics Austria)
- the Federal Ministry of Agriculture, Forestry, Environment and Water Management
- the Federal Ministry of Economy, Family and Youth

Figure 2: Structure and data sources for the Austrian Integrated NAMEA (Source: Statistics Austria)



Although the data are recorded and evaluated for different purposes and therefore also reflect differences, they can at the lowest common denominator - the structure of NACE - be made somewhat matchable

Basis for a possible exchange of data between the participating institutions are the Federal Statistics Act and the Environmental Information Act.

References

EUROPEAN COMMISSION (2008): TOWARDS A SHARED ENVIRONMENTAL INFORMATION SYSTEM. [HTTP://EUR-LEX.EUROPA.EU/LEXURISERV/LEXURISERV.DO?URI=COM:2008:0046:FIN:EN:PDF](http://eur-lex.europa.eu/lexuriserv/lexuriserv.do?uri=COM:2008:0046:FIN:EN:PDF)
OECD: GLOSSARY OF STATISTICAL TERMS. [HTTP://STATS.OECD.ORG/GLOSSARY/](http://stats.oecd.org/glossary/)
UNITED NATIONS (1984): A FRAMEWORK FOR THE DEVELOPMENT OF ENVIRONMENT STATISTICS (FDES).