Chapter X.   Dissemination

*This draft is based on the text adopted by the UN Statistical Commission for purposes of international recommendations for industrial and distributive trade statistics. Recommendations and encouragements are **in bold**. The Oslo Group members are invited to review the draft and to propose necessary amendments to reflect needs and priorities of energy statistics.*

A. National reporting

10.1   *Data dissemination*. Dissemination of energy statistics is one of the key activities in which national agencies responsible for energy statistics are involved. It is a means not only of providing policymakers, the business community and other users with high-quality statistical information but also of motivating respondents to participate in statistical surveys. If national responsible agencies have the legal power to collect and disseminate energy data, they also have the obligation to protect the confidentiality of respondents.

10.2   *Statistical confidentiality*. Most of the information about individual statistical units, that is either directly collected by responsible agencies or obtained from other sources, is considered confidential. Statistical confidentiality is necessary in order to gain and keep the trust of both those required to respond to statistical surveys and those using the statistical information.

10.3   Principle 6 of the United Nations Fundamental Principles of Official Statistics provides the basis for managing statistical confidentiality. It states:

> Individual data collected by statistical agencies for statistical compilation, whether they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes.

10.4   Legal provisions governing statistical confidentiality at national level are set forth in countries' statistical laws or other supplementary governmental regulations. National definitions of confidentiality and rules for microdata access may differ, but they should be consistent with this fundamental principle. This is especially relevant for countries where the distinction between statistical and non-statistical use of microdata is not fallowed by a long tradition, or is not laid down in any legislation.

10.5   Energy data are usually disseminated by national responsible agencies in the form of various statistical tables. Statistical confidentiality is protected if tabulated data do not allow statistical units to be identified either directly or indirectly, thereby disclosing individual information. Direct identification is possible if data of only one statistical unit are reported in a given data cell, while indirect identification may take place if individual data can be derived from disseminated data (for example, because information about too few units are reflected in a cell, or because of the dominance of one or two units in a cell). To determine whether a statistical

unit is identifiable, account shall be taken of all means that might reasonably be used by a third party to identify it.[1]

10.6   *Statistical disclosure control*. Statistical disclosure control is an application of a set of methods used to reduce the risk of disclosing information on individual units. While application of such methods occurs at the dissemination stage, they are pertinent to all stages of the process of statistical production.

10.7   Statistical disclosure control methods related to the dissemination are usually based on restricting the amount of data or modifying the data release. Disclosure control methods attempt to achieve an optimal balance between the improvement in confidentiality protection and the reduction in data quality. Different types of data pose different types of confidentiality problems and inevitably require different solutions. On the basis of available international guidelines[2] in this area, countries **are encouraged** to develop their own statistical disclosure methods which best suit their specific circumstances. Examples of the most often used methods are presented in paragraph 10.8 below.

10.8   *Methods of protecting confidentiality*. As the first step in statistical disclosure control of tabular data, the sensitive cells need to be identified. The sensitive cells are those that tend to reveal directly or indirectly information about individual statistical units. The most common practices carried out to protect against the disclosure of confidential data include:

(a)   *Aggregation*. A confidential cell in a table is aggregated with another cell and the information is then disseminated for the aggregate and not for the two individual cells. This may result, for example, in grouping of energy data at the class level of ISIC that are confidential with data from another class and presenting and disseminating them at the group level of ISIC;

(b)   *Suppression*. Suppression means removing records from a database or a table that contains confidential data. This is a method that allows statisticians not to publish the values in sensitive cells while publishing the original values in the others (called primary suppression). Suppressing only one cell in a table means, however, that the calculation of totals for the higher levels to which that cell belongs cannot be performed. In this case, some other cells must also be suppressed in order to guarantee the protection of the values under the primary cells, leading to secondary suppression;

(b)   *Other methods*. Controlled rounding and perturbation are more sophisticated techniques for protecting confidentiality of data. Controlled rounding allows statisticians to modify the original value of each cell by rounding it up or down to a near multiple of a base

---

[1] See Eurostat definition of confidentiality, as set out in chap. V (Statistical confidentiality) of Council of the European Union Regulation No. 322/97 of 17 February 1997 on Community statistics; and Council Regulation No. 1588/90 of 11 June 1990 on the transmission of data subject to statistical confidentiality.
[2] See Principles and Guidelines for Managing Statistical Confidentiality and Microdata Access, background document prepared for the Statistical Commission at its thirty-eighth session, held in New York from 27 February to 2 March 2007 (http://unstats.un.org/unsd/statcom/sc2007.htm).

number. Perturbation represents a linear programming variant of the controlled rounding technique.

10.9   In cases where countries prefer suppression as a method for protecting confidentiality of energy data, **it is recommended that** any data deemed confidential be reported in full detail at the next higher level of classification that adequately protects confidentiality, if data are presented by activities, or the next higher level of aggregation, for any other characteristics. Apart from satisfying the confidentiality protection, this method must result in a minimum loss of information.

10.10   *Confidentiality of complex enterprises.* Careful measures to respect confidentiality should be taken in the case of large enterprises of energy sector. In general, large units are more easily identifiable than small ones, and have a higher probability of being selected in energy surveys; and most likely, their data will dominate the totals in a number of cells, thus allowing identification of such units. The responsible agencies **are encouraged** to discuss with such enterprise condition for dissemination of the data reflecting their activities.

10.11   *Confidentiality rules for energy data*. Rules for protecting confidentiality of energy data should be in line with the provisions of countries' national legislation and practice. As a minimum requirement, the following two factors should be taken into account when defining the confidentiality rules: (a) number of units in a tabulation cell; and (b) dominance of a unit's or units' contribution over the total value of a tabulation cell. A decision with respect to the exact definition of the confidentiality criteria, for example, in terms of the number of units per cell and percentage of dominance, is left to the national responsible agencies. In individual cases, confidentiality rules may be relaxed by requesting the permission of the dominating respondent(s) to authorize the statistical office to disclose the data.

10.12   *Internationalization of confidentiality.* Data collected and disseminated by international organizations depend to a large degree on the quality and completeness of the data supplied by the countries concerned. Therefore, not only does the issue of confidentiality have a national dimension but it is also becoming international in scope, and for the following reasons: (a) high degree of interest in cross-country comparisons; (b) internationalization of users of statistical data (including international organizations); and (c) increase of data dissemination over the Internet. As a result, there is a growing demand for countries' data at a very detailed level and in some cases, even a demand for countries' microdata.

10.13   *Data dissemination timetable.* In producing statistical information, there is usually a trade-off between the timeliness with which the information is prepared and the accuracy and level of detail of the published data. A crucial factor, therefore, in the maintaining of good relations between national responsible agencies as producers of energy statistics and the user community is that of devising an appropriate compilation and release schedule. As this is important for the measurement of timeliness, which is one of the quality dimensions of energy statistics, **it is recommended that** countries develop and announce in advance the precise dates at which energy statistics will be released. The advance release calendar should be posted in the beginning of each year on the country's statistical office website.

10.14   The most important elements that should be taken into account in determining the compilation and release schedule of energy statistics include:

(a)   Timing of the collection of initial data from major energy surveys;

(b)   The extent to which data derived from the major data sources are subject to revisions;

(c)   Timing of preparation of important national energy policy documents that need energy statistics as inputs;

(d)   Modes of data dissemination (press release, electronic or hard copy).

10.15   Timeliness of release of initial monthly, quarterly and annual energy data varies greatly from country to country, mainly reflecting different perspectives on the timeliness-reliability-accuracy trade-off. In keeping with sound statistical practices, countries **are encouraged** to release their initial monthly data […   days after the end of the month in question *[for information – 45 days for industrial statistics]*; their quarterly data, […] months after the end of the quarter *[for information – 3 month for industrial statistics]*; and their annual data, […] months after the end of the year *[for info – 18 months for industrial statistics]*. Monthly and quarterly data should refer to a discrete month or quarter. Most countries use a separate system for compilation of annual energy statistics. In this case, the data for the fourth quarter (or for the twelfth month) need to be published in their own right and should not be derived as the difference between the annual totals and the sum for the first three quarters (or 11 months).

10.16   *Data revisions.* Revisions are an essential part of country practices in respect of the compilation of energy statistics. Their production is a consequence of the trade-off between the timeliness of published data and their reliability, accuracy and comprehensiveness. To resolve these issues, responsible agencies compile provisional data which are later revised when new and more accurate information becomes available. Although, in general, repeated revisions may be perceived as reflecting negatively on the reliability of official energy data, the attempt to avoid them by producing accurate but very untimely data will ultimately fail to satisfy users' needs. It is important to emphasize that the revisions of energy data are produced for the benefit of users, in order to provide them with data that are as timely and accurate as possible. The revisions affect both annual and short-term energy statistics but they are more significant for the short-term data.

10.17   *Reasons for revisions of data.* In general, there are two types of revisions: (a) revisions arising from "normal" statistical procedures (for instance, availability of new information, change in the methodology, change in data source, change of base year, etc.); and (b) revisions in the form of the correction of errors that may occur in source data or in processing. For normal statistical data revisions (also called current revisions), countries should develop a revision policy. At any moment of time, responsible agencies may decide to carry out a special revision, in addition to the normal statistical data revisions, for the purpose of reassessing the data or investigating in depth some new economic structures. Such revisions are carried out at longer, irregular intervals of time. Often, they may require changes in the time series going as far back

as the beginning of the series to retain methodological consistency. **It is recommended that** these revisions be subject to prior notification from the countries' responsible agencies to users -- notification that covers the reasons and provides information on the impact of the revisions on the data.

10.18   *Revision policy*. To deal with the issues surrounding revisions of energy data, countries **are encouraged** to develop a revision policy that is well designed, carefully managed and well coordinated with other areas of statistics. The development of a revision policy should be aimed not at impeding revisions but rather at providing users with the information necessary for coping with revisions in a more systematic manner. The absence of coordination and of planning of revisions is considered a quality problem by users. Essential features of a well-established revision policy are a predetermined schedule, reasonable stability from year to year, openness, advance notice of reasons and effects, and easy access of users to sufficiently long time series of revised data, as well as adequate documentation of revisions included in the statistical publications and databases.

10.19   *Recommended practices for data revisions*. A sound revision policy is recognized as an important aspect of good governance in statistics, as it will not only help the national users of the data but will also promote international consistency. With a view to assisting countries that have not yet set out such a policy, the following good practices are **recommended:**[3]

(a)   There should be consultations with users to elicit their views on revisions practices;

(b)   A clear, short summary statement of when to expect revisions and why should be readily accessible to users;

(c)   The current revision cycle should be relatively stable from year to year;

(d)   Major conceptual and methodological revisions should usually be introduced every four to six years, balancing need for change and users' concern;

(e)   Revisions should be carried back several years to yield consistent time series;

(f)   Documentation on revisions should be readily available to users;

(g)   Users should be reminded of the size of the likely revisions based on past history;

(h)   When a mistake in reporting or processing is made, the revision should be carried out in a transparent and timely manner.

10.20   *Dissemination formats*. A key to the usefulness of energy statistics is the availability of data and hence its extensive dissemination. Data can be disseminated both electronically and in paper publications. **It is recommended that** countries choose the dissemination format that best

---

[3] For details, see Organization for Economic Cooperation and Development *Data and Metadata Reporting and Presentation Handbook* (Paris, 2007), chap.7, "Guidelines on key reporting practices."

suits their users' needs. For example, press releases of energy statistics have to be disseminated in ways that facilitate re-dissemination by mass media; more comprehensive or detailed statistics need to be disseminated in electronic and/or paper formats. If resources permit, current statistics and longer time series can be organized and accessed (free of charge or for a fee) through the electronic databases maintained by the statistical office. In addition to statistics that are routinely disseminated, responsible agencies can make available to users energy data upon request. For some specific purposes, customized tabulations of data (non-standard activity classification, specific types of units, etc.) can be provided. It is advisable that countries ensure that users are clearly made aware of the availability of additional statistics and the procedures for obtaining them.

10.21 *Dissemination of metadata*. Provision of adequate metadata and quality assessment of energy statistics are as important to users as provision of data. Countries **are encouraged** to follow the recommendations provided in chapter IX on data quality and metadata for energy statistics and to develop and disseminate metadata comprising the following components: (a) data coverage, periodicity and timeliness; (b) access by the public; (c) integrity of disseminated data; (d) data quality; (e) summary methodology; and (f) dissemination formats. **It is recommended** that countries indicate in the metadata all deviations from internationally accepted statistical standards and guidelines. Energy statistics metadata should be made readily accessible through websites and/or publications of responsible agencies. Countries might consider developing different levels of detail of metadata so as to meet the requirements and needs of specialized users.[4]

B. International reporting

This section is to be drafted at a later stage

10.22 Countries **are encouraged** to make energy data available on their websites or to disseminate them internationally as soon as they become available to national users.

10.23 Tables X.1, X.2 and X.3 below provide a list of data items on energy statistics recommended for international dissemination and their level of detail and periodicity.

Table X.1

List of data items on energy statistics for international dissemination with annual periodicity

Item number	Data item	Level of detail	Minimum requirements	Deadline

---

[4] For further details on data and metadata reporting, see Organization for Economic Cooperation and Development, *Data and Metadata Reporting and Presentation Handbook* (Paris, 2007).

Table X.2

List of data items on energy statistics for international dissemination with quarterly periodicity

Item number	Data item	Level of detail	Minimum requirements	Deadline

Table X.3

List of data items on energy statistics for international dissemination with monthly periodicity

Item number	Data item	Level of detail	Minimum requirements	Deadline