United Nations | Department of Economic and Social Affairs

Energy Statistics Compilers' Manual

ST/ESA/STAT/SER.F/118

Department of Economic and Social Affairs
Statistics Division

Statistical Papers Series F No. 118

Energy Statistics Compilers' Manual

United Nations
New York, 2022

Department of Economic and Social Affairs

The Department of Economic and Social Affairs of the United Nations is a vital interface between global policies in the economic, social and environmental spheres and national action. The Department works in three main interlinked areas: (i) it compiles, generates and analyses a wide range of economic, social and environmental data and information on which United Nations Member States draw to review common problems and to take stock of policy options; (ii) it facilitates the negotiations of Member States in many intergovernmental bodies on joint courses of action to address ongoing or emerging global challenges; and (iii) it advises interested Governments on the ways and means of translating policy frameworks developed in United Nations conferences and summits into programmes at the country level and, through technical assistance, helps build national capacities.

Note

The designations used and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the United Nations concerning the legal status of any country or of its authorities or the delimitation of its frontiers. The term "country" as used in this publication also refers, as appropriate, to its territories or areas.

The symbols of United Nations documents are composed of capital letters and numbers. Mention of such a symbol indicates a reference to a United Nations document.

ST/ESA/STAT/SER.F/118
United Nations publication
Sales No. E.17.XVII.10
ISBN: 978-92-1-161621-7
eISBN: 978-92-1-060140-5

Preface

*The Energy Statistics Compilers' Manual (*ESCM) complements in a practical manner the *International Recommendations for Energy Statistics (*IRES), which provide a comprehensive methodological framework for the collection, compilation and dissemination of energy statistics in all countries irrespective of the level of development of their statistical system. In particular, IRES provide a set of internationally agreed recommendations covering all aspects of the statistical production process, from the institutional and legal framework, basic concepts, definitions and classifications to data sources, data compilation strategies, energy balances, data quality issues and statistical dissemination.

IRES and ESCM were prepared in response to the request of the United Nations Statistical Commission at its thirty-seventh session (7–10 March 2006) to review the United Nations manuals on energy statistics, develop energy statistics as part of official statistics, harmonize energy definitions and compilation methodologies and develop international standards in energy statistics.

The preparation of IRES and ESCM was carried out by the Statistics Division of the United Nations in close cooperation with the Oslo Group on Energy Statistics and the Intersecretariat Working Group on Energy Statistics (InterEnerStat).

ESCM is aimed primarily at practitioners tasked with building or improving the energy statistics programme of a country or institution in a way that is consistent with the latest international standards and that produces reliable and internationally comparable data.

ESCM aims to assist countries in the collection, analysis and dissemination of energy statistics according to international standards. It provides explanations to facilitate the application of the principles defined in IRES in practical situations, understand specific relationships that facilitate or complicate the adaptation of such principles to national situations and indicates practical ways to implement an energy statistics programme that is consistent with established international recommendations.

ESCM builds heavily on country practices to illustrate various ways to implement the recommendations set out in IRES, and thus suggests a variety of ways to tackle issues that may arise during the implementation of IRES. Additional country examples can be found on the Statistics Division energy statistics website (https://unstats.un.org/unsd/energystats/country-practice).

Acknowledgements

This *Manual* has been prepared by the United Nations Statistics Division in close collaboration with the Oslo Group on Energy Statistics. In particular, as Chair of the Group from 2009 to 2015, Statistics Canada was instrumental in developing a work plan and following through on improving earlier drafts. The Statistics Division took the lead in the final editing phase.

Chapter II was drafted by the Instituto Nacional de Estadística y Geografía (INEGI), Mexico; chapter III was drafted by the Statistics Division; chapters IV and V were initially drafted by Statistics Austria and further developed by the Statistics Division; chapter VI was drafted by Statistics Norway and further developed by the International Energy Agency and the Statistics Division. Chapter VII was drafted by Statistics Canada; and chapter VIII was drafted by Statistics South Africa.

As many chapters were drafted by individuals in different countries and/or organizations, it is inevitable that, even after editing, each chapter will retain its own distinctiveness and style.

Thank you, in particular, to all the countries that provided examples of their practices, both for this publication and for the online collection of country practices.

Contents

Page

Boxes

Page

Figures

Tables

Abbreviations and acronyms

ANZSIC	Australian and New Zealand Standard Industrial Classification
API	American Petroleum Institute
BPM6	Balance of Payments and International Investment Position Manual
Btu	British thermal unit
CHP	combined heat and power
CRIRSCO	Committee for Mineral Reserves International Reporting Standards
CPC	Central Product Classification
DQAF	Data Quality Assessment Framework
ECE	United Nations Economic Commission for Europe
EEA	European Environmental Agency
ESCM	*Energy Statistics Compilers' Manual*
ESS	European Statistical System
Eurostat	Statistical Office of the European Union
GCV	gross calorific value
GDP	Gross domestic product
GHG	greenhouse gas
GTL	gas-to-liquids
GSBPM	Generic Statistical Business Process Model
GWh	gigawatt hour
HS	Harmonized Commodity Description and Coding System
IEA	International Energy Agency
InterEnerStat	Intersecretariat Working Group on Energy Statistics
IPCC	Intergovernmental Panel on Climate Change
IRES	International Recommendations for Energy Statistics
IRIS	International Recommendations for Industrial Statistics
ISIC	International Standard Industrial Classification of All Economic Activities
LPG	liquefied petroleum gas
kWh	Kilowatt-hour
NACE	Statistical Classification of Economic Activities in the European Community
NAICS	North American Industry Classification System
n.e.c.	not elsewhere classified
NQAF	National Quality Assurance Framework

NCV	net calorific value
NGL	natural gas liquids
OECD	Organisation for Economic Co-operation and Development
PRMS	Petroleum Resources Management System
RON	research octane number
SBP	special boiling point
SDMX	Statistical Data and Metadata Exchange
SEA	Swedish Energy Agency
SEEA	System of Environmental-Economic Accounting
SEEA-Energy	System of Environmental-Economic Accounting for Energy
SI	International System of Units
SIEC	Standard International Energy Product Classification
SIMS	Single Integrated Metadata Structure
SNA	System of National Accounts
TES	total energy supply
Tce	ton of coal equivalent
Toe	ton of oil equivalent
UNFC	United Nations Framework Classification for Resourcess
UNFCCC	United Nations Framework Convention on Climate Change
VAT	value added tax

Chapter I: Introduction

A. Background

1.1. Energy plays an essential role in almost all forms of human activity. A successful economy is typically characterized by a reliable and efficient supply of energy that meets the full range of its social and economic needs. Ideally, all households should have access to clean, affordable and reliable energy, while businesses should have access to energy that enables them to produce goods and services in a competitive marketplace. Businesses that supply energy should also be viable and stable. However, growing worldwide demand for energy gives rise to concerns about the sustainability of supply and the impacts on the environment.

1.2. In this modern context, it is essential for countries to monitor and manage their energy resources and various aspects of energy production and use. In order to do this well, it is important to ensure that policy decisions are informed by reliable and appropriate data. The development of systems to produce quality and consistent energy information should be based on internationally agreed standards, classifications and other frameworks as these will enable cross-country comparability and consistency over time.

1.3. There is a broad range of contemporary energy issues that must be addressed. These include ensuring a sustainable supply of energy resources for the future; developing and maintaining the infrastructure for the transport of energy products to market; managing energy price volatility; encouraging investment, innovation and efficiency in the energy sector; ensuring emergency preparedness; and managing the environmental impacts of energy production and use.

1.4. Relevant and timely energy data are necessary to support evidence-based energy policy and decision-making, to monitor and assess programmes, to enable research and analysis, and to inform the public on energy-related matters. As such, energy statistics are vital for a well-functioning economy and a well-maintained environment. In this context, and with no internationally agreed standards for energy statistics collection, the United Nations Statistical Commission considered it necessary to call for the revision and further development of relevant international statistical guidelines at its thirty-sixth (2005) and thirty-seventh (2006) sessions.

B. International Recommendations for Energy Statistics

1.5. The United Nations Statistical Commission, at its forty-second session (New York, 22–25 February 2011), adopted the *International Recommendations for Energy Statistics* (IRES). The present publication, *Energy Statistics Compilers' Manual* (ESCM), has been prepared in accordance with the recommendation made by the Commission at that session in relation to energy statistics, that the *Manual* must contain clear guidelines on data sources, the use of administrative data, and best practices applicable to a wider range of countries.

1.6. The primary purpose of this *Manual* is to assist countries in strengthening official energy statistics by providing guidance on energy concepts and definitions, the legal foundation and institutional arrangements for the collection of energy data, classification systems, data sources and compilation methods, energy balances, practices for ensuring data quality, metadata requirements, and dissemination policies and practices. Such guidance relies on precedents in the field of energy statistics and assists in resolving challenges in the compilation of energy statistics.

1.7. The *Manual* is directed towards all institutions that play a role in the collection, compilation and dissemination of energy statistics; the term "compiler", as used in the *Manual* refers to those institutions. In addition, the *Manual* will enable data users to have a richer understanding of energy statistics and therefore a heightened capacity to both analyse energy information and communicate with data compilers on such subjects as data quality.

1.8. The *Manual* is intended to be a practical and relevant guide for all countries, irrespective of size, stage of economic development, energy resource endowments or institutional arrangements. Country-specific experiences and practices are provided throughout the *Manual* to illustrate how energy data can be collected, compiled and disseminated in a variety of circumstances.

1.9. The IRES was developed largely to address concerns regarding data availability and international comparability for energy statistics. A key element in improving these aspects of energy statistics relates to the clarification and standardization of the scope and boundaries of energy products. The IRES essentially defines the scope of energy products as being products contained within the Standard International Energy Product Classification (SIEC). This classification reflects a historically agreed identification of the products needed to provide a comprehensive and analytically useful picture of the production, transformation and consumption of energy within an economy. The identification of energy products within SIEC in turn facilitates the integration of these product measures with the Central Product Classification (CPC) and the Harmonized Commodity Description and Coding System (HS).

C. Organization of the *Energy Statistics Compilers' Manual*

1.10. This *Manual* covers a number of topics that will provide a step-by-step approach to producing energy statistics. The chapters are as follows:

Chapter I: Introduction

This chapter describes the purpose of the *Manual* and outlines its structure.

Chapter II: Legal framework and institutional arrangements

Each country has a legal framework and institutional arrangements that impact the collection and dissemination of its energy statistics. This chapter examines a range of possible legal and institutional arrangements affecting the design and operation of a national system of official energy statistics. It covers two broad areas:

- Legal framework: the need for a legislative framework as a foundation for energy data collection to provide authority and direction for the collection, handling and processing of information.

- Institutional arrangements: a discussion on the types of working mechanisms among relevant institutions with the potential to assist or otherwise impact official energy statistics. It provides examples from both centralized and decentralized statistical systems. For example, the operation of institutional arrangements between a Ministry of Energy and a National Statistical Office; and the operation of a national statistical committee on energy statistics. The goal of these arrangements is to support the efficient collection and dissemination of cohesive energy statistics.

Chapter III: Classifications

The use of standard classifications is of paramount importance in the collection, compilation and dissemination of statistics. Standard classifications facilitate data collection as they provide a clear definition, with a unique structure, of the objects that are being measured and collected. This chapter elaborates on the classifications used in energy statistics. In particular, it covers the classification of statistical units, providing links with the *International Standard Industrial Classification of All Economic Activities* (ISIC), where appropriate, and the classification of energy products according to SIEC, with its links to the CPC and the HS. The chapter also describes some of the issues relating to the correspondence between these international classifications and provides examples of country practices.

Chapter IV: Generic statistical business process model

This chapter focuses on the different stages of the data production process following the Statistical Business Process Model developed by the United Nations Economic Commission for Europe (UNECE). It also reviews the different stages of the statistical data production process within an energy statistics context in anticipation of the discussion of data sources and collection in the next chapter.

Chapter V: Data sources and data collection

Data sources and the approaches to data collection that support energy statistics vary among countries. This chapter provides a comprehensive assessment of possible data sources and collection tools, including direct collection surveys, administrative data, metering and data modelling, illustrated with country-specific examples. In particular, it provides a number of case studies to demonstrate how different collection approaches can be combined to build a strategy that delivers all required energy information with a minimum of resources. The chapter concludes with some country examples to show how different approaches can be combined to obtain all required information with minimum resources.

Chapter VI: Commodity and energy balances

Energy balances are an important basis for making informed policy decisions and are underpinned by a range of sources and methods. Commodity balances, on the other hand, are a useful presentation tool for showing data on different energy products in the same fashion, while maintaining their natural unit, as well as a useful step in the compilation of energy balances. This chapter provides practical guidance in the compilation of both commodity balances and energy balances, and describes the general principles and methods used for doing so. Section A provides a description of the nature and importance of commodity balances and energy balances; section B provides general information pertinent to both commodity balances and energy balances; and section C describes the compilation of commodity balances for one, or all, energy products. The chapter also presents various data sources that can be used for compiling commodity balances, addresses possible data collection challenges, and discusses associated data editing, and estimation and reconciliation methods. In addition, examples are given of a commodity balance for three products for one country, and preliminary balances for all products collated. Section D explains the progression from commodity balances to energy balances, how to use calorific values, and methods for setting the value of primary energy using physical energy content. Validation rules inherent to energy balances are addressed here. This section also contains guidance on how energy statistics compilers may respond to various specific technical issues, such as dealing with differences between the physical energy content method and the partial substitution method, and calculating the renewable energy column in the energy balances. The section ends with a discussion of new energy technologies that may not yet appear in energy balances. Section E provides several country-specific examples on how to compile energy balances.

Chapter VII: Quality assurance frameworks and metadata

Ensuring the production of high-quality energy statistics is a key responsibility and challenge for all statistical offices. This chapter looks at ways in which statistical agencies can ensure data quality by considering quality dimensions when measuring and reporting the quality of statistics, and by implementing quality measures related to each stage of the data production process. In particular, the international context is highlighted with several quality frameworks and examples of both quality reports and quality reviews from various countries and organizations. In keeping with the practical and energy-related spirit of the chapter, section F considers how the International Energy Agency (IEA) makes data quality and validation checks to ensure the quality of energy statistics. The chapter finishes with section H on statistical metadata. This type of documentation provides the means for assessing the fitness for use of energy data and contributes directly to their interpretability, availability and accessibility. To illustrate the application of quality frameworks in energy statistics, examples of country practices related to the quality dimensions in energy statistics are presented, and references are made to data quality and metadata frameworks from various countries and organizations.

Chapter VIII: Data dissemination

Once statistics are produced, it is important that they are disseminated in a consistent, user-orientated manner. This chapter provides practical guidance on how to follow IRES guidelines on dissemination timing, ensuring confidentiality, revision of published data, dissemination format, publishing metadata and compiling preliminary data. The manner in which energy statistics are disseminated varies markedly from country to country; this chapter draws on examples of country practice for these topics. The chapter ends with a brief description of the importance of energy indicators for dissemination purposes. While energy indicators (collating energy data with other statistics to make comparisons across countries and economic sectors and time) are an exceedingly broad topic and other publications are devoted specifically to them, energy indicators are an important part of energy dissemination and it was considered relevant to include them herein.

Chapter II: Legal framework and institutional arrangements

A. Introduction

2.1. The effective management of energy resources has become an increasingly important responsibility of countries around the world. Energy production, transformation, distribution and consumption have all been linked to high-priority and high-profile issues relating to the economy, society and the environment. Therefore, it is crucial to have good-quality, timely and relevant energy data to support the essential planning, monitoring and decision-making for the sustainable development of this sector.

2.2. Countries can take different approaches to establish and maintain a sound national energy statistics system, but they must be developed in such a way as to fit within the local context and circumstances. Nonetheless, regardless of the approach adopted, the national statistics system must be built upon a strong foundation that incorporates a clear legal framework, as well as effective institutional arrangements. This chapter presents the key features of each one and provides a number of examples to illustrate how different countries have implemented the legal framework and institutional arrangements for their energy statistics programmes. In addition, a set of internationally recognized, fundamental principles of official statistics are presented.

B. Legal framework

2.3. According to the International Recommendations for Energy Statistics (IRES), endorsed by the United Nations Statistical Commission in 2011, the existence of a strong legal framework is one of the most important prerequisites for building a sound national statistical system in general, and a national system of energy statistics in particular.

2.4. A legal framework is a set of laws, rules or regulations that establish a data collection entity (such as a national statistics office or agency, a State statistical committee or other similar body), with the mandate and the legal authority to collect, compile and disseminate statistics within a country, along with the responsibility of ensuring the privacy of respondents and the security and confidentiality of the information collected. The framework also provides strategies for the enforcement of the legal authority and responsibilities. Each of these components will be discussed in turn.

1. Establishment of a data collection entity

2.5. Many countries have enacted legislation to create an official statistical entity to oversee the development and operation of a national system of statistics. That entity is given the authority and responsibility to support informed decision-making, research and discussion within governments and the community through the provision of a high-quality, objective and responsive national statistics service.

Box 2.1:
Country practice: Australia

The Australian Bureau of Statistics Act 1975 establishes the Australian Bureau of Statistics (ABS) as an independent statutory authority, defines its functions, establishes the Office of the Australian Statistician and describes the terms under which the Australian Statistician can be appointed to, and removed from, office. The Act also provides for the appointment of the staff of the ABS and creates the Australian Statistics Council.

The Census and Statistics Act 1905 provides the Australian Statistician with the authority to conduct statistical collections, including the Population and Housing Census and, when necessary, to direct a person to provide statistical information. The Act requires the ABS to publish and disseminate compilations and analyses of statistical information and to maintain the confidentiality of information collected thereunder.

Box 2.2:
Country practice: Ghana

In Ghana, the Ghanaian Parliament, by the Energy Commission Act 1997, established the Ghanaian Energy Commission, with the responsibility to create a comprehensive energy database to support national decision-making and to ensure the efficient development and utilization of the energy resources available to the nation. The Commission also has the responsibility for licensing public utilities for the transmission, wholesale supply, distribution and sale of electricity and natural gas. The Act therefore mandates the Commission to collect, compile, analyse and disseminate information for policymakers, researchers and planners.

2.6. The system for the collection of statistics in countries can be set up in different ways. In general, there are two different approaches – a centralized or decentralized approach. Either can be adapted to meet a country's circumstances and needs, and each can work effectively as long as a strong legal foundation and effective institutional arrangements are in place.

2.7. A *centralized statistical approach* is one in which a central or national agency has been given the mandate and authority for all (or most) of a country's official data collection activities. There are advantages for a country to take a centralized approach. Some examples of the advantages are:

- *Economies of scale*: A central agency responsible for a broad range of activities may be able to find savings through the efficient use of infrastructure. For example, there may be a critical mass of data collection work that could allow for a permanent collection infrastructure to be set up and maintained. There also could be opportunities to avoid duplication of effort or to reduce respondent burden by coordinating collection in one place.
- *Centres of expertise*: A centralized organization managing a broad range of surveys may be able to attract and develop a critical mass of expertise in disciplines such as questionnaire design, collection methodology, data analysis, systems development, etc. This might not be feasible, for example, in decentralized operations that may be smaller (e.g. an energy survey unit located in an energy policy department).

- *Centrally managed tools*: An organization that is focused on statistics could allow for the development and maintenance of useful tools, standards and applications that might not be feasible in smaller, more specialized organizations (e.g. business register, metadata repositories, electronic questionnaires).

- *Promoting links between subject areas*: It may be easier to encourage and implement statistical collaboration between related subject areas (e.g. energy and environment) if they are located within the same centralized agency. For example, there could be collaboration on harmonizing concepts and definitions, or conducting data validation and analysis.

- *Coordinated data dissemination*: A centralized statistical agency may be well placed to serve as an official one-stop location for users to get data on a variety of subject areas. At the same time, it may be easier to ensure that all statistics are disseminated or available to all users at the same time.

- *Perceptions of data qu*ality: Where collection activities are separated from policy areas, users may perceive a greater likelihood of data objectivity, integrity, higher professional standards and adherence to established principles of statistical methods.

2.8. On the other hand, a *decentralized statistical approach* can also be an effective strategy for energy data collection, analysis and dissemination. A decentralized system is one in which the statistical responsibilities are co-located with policy or operational areas (e.g. energy departments). Depending upon the country, there can be advantages to this approach. Some examples include:

- *Subject matter expertise*: The energy domain can be a very complex subject matter. Centralized statistical agencies may lack the degree of subject matter knowledge and expertise to be able to determine information needs and the best methods for collecting the necessary information. Decentralized approaches may allow for closer collaboration with data users and energy experts to ensure the right data are collected to meet their needs. This may also facilitate collaboration with respondents from within the energy sector (i.e., being able to discuss the complex subject matter at a reasonable level of expertise).

- *Practical considerations*: Some countries may not have the capacity to establish centralized statistical agencies. It may be more practical or feasible to set up collection activities co-located with other energy operations where resources could be allocated and re-allocated, as required.

- *Greater links to the administrative data sources*: Decentralized statistical systems may allow for closer links to, and understanding of, administrative data sources (e.g. better knowledge of the availability, opportunities and limitations).

Box 2.3:
Country practice: Canada

Statistics Canada serves as the focal point of a centralized statistical system in Canada, operating as a separate department within the national Government. It sets its own work plan and priorities, manages its own budget and gets advice from the National Statistics Advisory Council and a series of subject-matter advisory committees. As a dedicated and centralized statistical organization, it is able to serve as a centre of professional expertise on surveys and methodology, to implement standards, to promote data coherence across subject areas, and to support cross-cutting statistical analysis.

The energy statistics programme operates within Statistics Canada. It gathers data through surveys and from administrative data sources from across the country. As the responsibility for the management of natural resources in Canada is decentralized, there are many stakeholders active in the management of this sector. As such, the energy statistics programme has to collaborate with many departments, regulatory bodies and other organizations across the country on the collection, analysis and dissemination of energy data.

Box 2.4:
Country practice: Sweden

The legal framework for the collection of data in Sweden was established by the Statistics Act of 1756. This has been supplemented by a Statistics Act and Statistics Ordinance, passed in 2001, which cover the collection and dissemination of official statistics.

Sweden has a decentralized statistical system. The Act and Ordinance provide for the designation of statistical authorities responsible for the collection of data for particular subjects. In Sweden, there are 27 statistical authorities responsible for the collection of official statistics, including the Swedish Energy Agency (SEA), which has been designated as the statistical authority for energy data. The Act and Ordinance specify that these statistical authorities have the mandate to decide the nature and extent of the data to be collected, that they should be objective and independent of the Swedish Government, and that all official statistics should be made available to the public and users. They also have the responsibility to determine how these data should be collected (i.e., by subcontracting collection activities to Statistics Sweden, or to a private subcontractor, in accordance with scientific methods and professional standards).

Box 2.5:
Country practice: United Kingdom

The Statistics of Trade Act (1947) and the Statistics and Registration Service Act (2007) provide the legal foundation for the collection of statistics. The United Kingdom has a decentralized system, but with a Statistics Authority that oversees the main overarching statistical legislation. The Office for National Statistics coordinates the national statistics system, but ministerial departments generally have responsibility for the legal framework used to collect and publish statistical information. As such, energy statistics are collected and produced by the Department of Energy and Climate Change (DECC). There are formal institutional arrangements in place to promote collaboration on energy statistics in the United Kingdom. DECC and the Statistics Authority work together to ensure access to statistics and adherence to the code of practice.

2. Legislative authority for data collection, compilation and dissemination

2.9. The mandate and legal authority of the statistical entity should typically include a clear statement of its roles and responsibilities relating to the collection, compilation and dissemination of data. These can include the authority to collect data through the use of a census, sample surveys or by accessing administrative data sources; the authority to compel response from respondents (i.e. mandatory reporting); the requirement to collaborate with other organizations on the development, collection, validation and dissemination of data in order to reduce duplication of effort and minimize respondent burden; and the promotion of the use of common concepts, definitions and standards in order to support data coherence and comparability.

Box 2.6:
Country practice: Canada

The legal framework for data collection in Canada is established by the Statistics Act. This legislation created Statistics Canada as the country's official statistical agency with a mandate to collect, compile, analyse and publish statistical information, and to conduct censuses of population and agriculture. To achieve this, the agency can compel businesses and individuals to respond to surveys, and has been given broad access to all records (i.e., administrative data) held by governments. At the same time, Statistics Canada is directed to collaborate with other government departments (at the federal, provincial and territorial levels) on data collection, development and dissemination, in order to promote an integrated system of national statistics. Throughout the conduct of this work, the agency must avoid duplication of effort, minimize the burden placed on respondents, and guarantee the confidentiality of individually identifiable information.

3. Ensuring privacy, security and confidentiality

2.10. Another important component of the legal framework is the responsibility to ensure the privacy of respondents, the security of information holdings and the confidentiality of the data. These are critical challenges facing all statistical systems.

2.11. Statistical offices must rely on the trust and goodwill of respondents (i.e., individuals, households and businesses) to enable the collection and production of credible, high-quality official statistics. In order to earn that trust and goodwill, statistical offices must ensure:

- *Privacy*: By collecting only the information that is required and using it only for that purpose; by being open and transparent about the information being collected; disseminating all data collected; having protocols in place for ensuring privacy protection and making these known to the public; and having recourse mechanisms (e.g. a Privacy Officer) where respondents can pose questions or lodge complaints for investigation.
- *Security*: By keeping data safe from unauthorized access and use. This is achieved through a variety of mechanisms:
  - Physical building security
  - Computer system security
  - Limiting access to data to only those employees with a need to know
  - Ensuring the safe disposal of records

- *Confidentiality*: By not releasing information that could identify individuals, households or businesses. This can be done by:
  - Removing identifiers from data records
  - Restricting the use of information to statistical purposes only
  - Applying confidentiality patterns to published data
  - Controlling access to micro data files

4. Strategies for enforcement

2.12. The success of a statistical programme relies on the continuous collaboration of participants both from within and outside the statistical entity.

2.13. Within the statistical entity, a variety of activities need to occur on an ongoing basis to ensure that the provisions of the legal framework are being applied. These include:

- Providing staff with ongoing training and information to raise awareness and to emphasize the importance of ensuring privacy, security and confidentiality.
- Ensuring senior management leadership and oversight through regular reviews, messages and reminders to staff to emphasize the importance of the initiative.
- Conducting regular reviews of infrastructure, processes, and protocols for ensuring security.
- Developing and maintaining expertise in methodology and strategies for ensuring confidentiality.
- Conducting periodic audits to evaluate the degree to which these objectives are being met and to identify issues to be addressed.
- Studying and sharing best practices.

2.14. These efforts and activities should be backed up by formal processes as part of the legal framework. For example, all staff should be required to take a formal oath to keep data secure and confidential. This oath should extend beyond the period of employment with the statistical agency – typically for life. Any breach could result in fines or criminal prosecution.

2.15. Similarly, as part of the legal framework, there should be provisions dealing with partners outside the statistical organization. These provisions should include:

- *Laws to make reporting mandatory*. Failure to comply, or any attempt to disrupt data collection activities would be subject to criminal prosecution, fines and possible imprisonment.
- *Requirements to inform data suppliers and users*. Information must be disseminated to educate partners about the importance of reporting, the use of the collected information, and the protocols in place to ensure privacy, security and confidentiality, etc. This information should be made readily available (e.g. on the website, on questionnaires and in data releases).
- *Selective prosecution to raise awareness*.

Box 2.7:
Country practice: New Zealand

Statistics New Zealand is required by law to protect the information it collects. These requirements are outlined in the Statistics Act, 1975 and the Privacy Act, 1993.

Section 37 of the Statistics Act governs the use of individuals' information and specifies that the information collected can only be used for producing statistics, must be kept secure to prevent unauthorized access, and must not be released where it could lead to disclosure of individuals' details. All staff are bound by the Act until death and can face individual penalties for infringement.

The Privacy Act, 1993 covers the management of collected information. It established an expert Privacy Officer and a Privacy Advisory Group to ensure that Statistics New Zealand complies with all provisions of the Act. This includes monitoring, evaluating and reporting on compliance, and the handling of complaints.

C. Institutional arrangements

2.16. Institutional arrangements refer to those processes or mechanisms put in place to support collaboration between organizations in order to manage or improve the functioning of a national statistics programme. Institutional arrangements can be established formally or informally. Formal arrangements may be specified in legislation, as in the case where a statistical agency is granted access to all government information holdings for statistical applications. That allows the agency to work with other departments to establish an ongoing partnership (e.g. gaining access to data, determining the format and frequency). An example of an informal arrangement would be the creation of an advisory committee to share ideas and expertise.

1. Designation of a single agency responsible for the dissemination of official energy statistics or agencies responsible for specific data sets

2.17. Where many organizations are involved in the management of the energy sector, there may be questions about where to source the official country statistics on energy. This could lead to confusion or the proliferation of energy-related data sets.

2.18. Efforts should be made to coordinate the dissemination of data by partners within the country. It should be made clear to data users where they can access official data on each related topic. The dissemination process and tools should be known, accessible and understood.

Box 2.8:
Country practice: Mexico

In Mexico, the Law of the National System of Statistical and Geographical Information established a centralized national statistical agency. This agency has technical and management autonomy, as well as legal status. While the Law does not specifically provide a legal mandate for elaborating and disseminating energy data and indicators, there are internal regulations of the Ministry of Energy that refer to the national energy balance.

Institutional arrangements are also in place to promote collaboration among stakeholders on the collection and dissemination of energy statistics. The Ministry of Energy participates in two national subsystems of information – geography and environment, and economic.

Box 2.8 (continued)

A variety of bodies and commissions at the central and state levels in Mexico are active in the energy sector, and institutional arrangements are in place to engage these users and stakeholders in the discussion of energy statistics. This is done through consultations, working groups on specific activities and a Specialized Technical Committee on Information for the Energy Sector.

Box 2.9:
Country practice: United States

The United States Energy Information Administration (EIA) is the statistical and analytical agency within the United States Department of Energy. EIA collects, analyses and disseminates independent and impartial energy information to promote sound policymaking, efficient markets and public understanding of energy and its interaction with the economy and the environment. It is the nation's premier source of energy information and, by law, its data, analyses and forecasts are independent of approval by any other officer or employee of the federal government.

The Department of Energy Organization Act of 1977 established EIA as the primary federal government authority on energy statistics and analysis. It conducts a comprehensive data collection programme that covers the full spectrum of energy sources, end uses and energy flows. EIA also prepares informative energy analyses, monthly short-term forecasts of energy market trends, and long-term United States and international energy outlooks.

2. Clear definition of roles and responsibilities of relevant agencies

2.19. In most countries, there are a variety of organizations involved in the collection, compilation, management and dissemination of energy statistics. There may be valuable opportunities to work together on a variety of fronts. The nature and extent of those collaborative arrangements will differ from country to country. The challenge is to identify those opportunities and to put in place formal or informal processes to promote and support those efforts.

2.20. Some of the areas where institutional collaboration can result in benefits are:

- *Data collection and sharing*: Rather than different organizations each collecting their own survey data, data-sharing agreements could be negotiated to have one organization do the collection, then share the data with the other. This could save resources while reducing the burden on respondents.
- *Use of administrative data sources*: Similarly, if an organization (e.g. a regulatory body) is already collecting energy data for their own purposes, the statistical agency could tap into that source as a valuable alternative to survey collection. This could reduce costs and the burden on respondents. However, there are challenges with the use of administrative data for statistical purposes. These are discussed in chapter V on data sources and data collection.

3. Establishment of working groups

2.21. Another feature of institutional arrangements is the use of committees or working groups to bring together the various organizations contributing to the collection, compilation, management and dissemination of statistics. These can be an effective means for the coordination of efforts and effective management on several fronts.

- *Priority setting*: Working groups can be set up to bring stakeholders (e.g. data collectors, data users) together on a regular or periodic basis to discuss new and emerging data needs, to establish priorities and to identify new opportunities for collaboration, etc. This could help to inform decisions, for example, about the allocation of resources to priority initiatives.
- *Harmonization of concepts*: Organizations could collaborate to standardize concepts and definitions. These types of efforts are critical for improving data quality, coherence and utility. They could also be valuable in reducing the likelihood of conflicting or inconsistent data sets being released by different agencies.
- *Data validation and analysis*: Each organization involved in an aspect of the energy sector (e.g. regulatory, statistical, management) develops its own particular subject matter knowledge and industry expertise. There may be opportunities for collaboration on the sharing of this knowledge and expertise for the purposes of validating or explaining the data, or for conducting other analyses.

Box 2.10:
Country practice: Sweden

In terms of institutional arrangements, Sweden has a number of processes in place. Statistics Sweden coordinates the work of 25 statistical authorities through a Council for Official Statistics, which meets twice a year to discuss issues and make decisions. This Council also provides guidelines on conducting surveys, publishing official statistics, standardizing definitions and ensuring quality.

In this decentralized system, the SEA falls under the authority of the Ministry of Infrastructure. To provide advice and guidance on official energy statistics, there is a User Council consisting of statistical authorities, industry associations and energy researchers, which meets four times a year. That group discusses data needs, gaps and relevant issues.

D. Fundamental Principles of Official Statistics

2.22. Both the legal framework and institutional arrangements should conform to the Fundamental Principles of Official Statistics, which were first adopted by the United Nations Statistical Commission in 1994, and reaffirmed in 2013. The Principles were subsequently endorsed by the Economic and Social Council on 24 July 2013 and by the General Assembly on 29 January 2014.

Box 2.11:
Fundamental Principles of Official Statistics

Principle 1. Official statistics provide an indispensable element in the information system of a democratic society, serving the Government, the economy and the public with data about the economic, demographic, social and environmental situation. To this end, official statistics that meet the test of practical utility are to be compiled and made available on an impartial basis by official statistical agencies to honor citizens' entitlement to public information.

Principle 2. To retain trust in official statistics, the statistical agencies need to decide according to strictly professional considerations, including scientific principles and professional ethics, on the methods and procedures for the collection, processing, storage and presentation of statistical data.

Principle 3. To facilitate a correct interpretation of the data, the statistical agencies are to present information according to scientific standards on the sources, methods and procedures of the statistics.

Principle 4. The statistical agencies are entitled to comment on erroneous interpretation and misuse of statistics.

Principle 5. Data for statistical purposes may be drawn from all types of sources, be they statistical surveys or administrative records. Statistical agencies are to choose the source with regard to quality, timeliness, costs and the burden on respondents.

Principle 6. 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.

Principle 7. The laws, regulations and measures under which the statistical systems operate are to be made public.

Principle 8. Coordination among statistical agencies within countries is essential to achieve consistency and efficiency in the statistical system.

Principle 9. The use by statistical agencies in each country of international concepts, classifications and methods promotes the consistency and efficiency of statistical systems at all official levels.

Principle 10. Bilateral and multilateral cooperation in statistics contributes to the improvement of systems of official statistics in all countries.

E. Conclusion

2.23. Countries may adopt the model for the collection and dissemination of energy statistics that best suits their needs and circumstances. Different approaches can be successful and appropriate, depending on the country's circumstances and needs. However, any approach should incorporate a strong legal framework and effective institutional arrangements and should adhere to the Fundamental Principles of Official Statistics.

Chapter III: Classifications

A. Introduction

3.1. The use of standard classifications is of paramount importance in the collection, compilation and dissemination of statistics. Standard classifications facilitate data collection, as they provide a clear definition, with a unique structure, of the objects being measured and collected. They facilitate the compilation of data, as the classifications define relationships between concepts and objects. Finally, they allow for better integration of data collected across different statistical domains, such as energy, environment and economic statistics.

3.2. During the preparation of the *International Recommendations for Energy Statistics* (IRES), considerable work was undertaken to identify and harmonize the classifications used for energy statistics. The two most prominent classifications presented in IRES are the classification for statistical units and the classification of energy products. Other classifications used in energy statistics, for example, the classification of energy resources, are also presented in IRES and addressed in this chapter.

3.3. Section B of this chapter discusses the classification of statistical units in a way relevant for energy statistics. Section C presents the classification of energy products and the issues currently encountered in the correspondence with other international product classifications. Section D discusses the classification of energy resources.

3.4. It should be noted that, with the exception of the classifications of energy products and energy resources, the other "classifications" discussed here are lists of relevant items for these specific purposes and have not been devised as fully recognized classifications that would need to meet additional criteria, as defined by the Expert Group on International Statistical Classifications.

3.5. This chapter also provides a number of country examples (see boxes 3.1–3.7) illustrating practical implementation of classification concerns in some countries. These and other examples are also available on the Statistics Division Energy Statistics website, under "Country Practice Examples" (https://unstats.un.org/unsd/energystats/country-practice/).

B. Classifications of statistical units in energy statistics

3.6. A statistical unit is an entity about which information is sought and for which statistics are ultimately compiled. The definitions of statistical units referenced in IRES are consistent with those used for industrial statistics (see *International Recommendations for Industrial Statistics* (IRIS), 2008) and the units should be classified according to the *International Standard Industrial Classification of All Economic Activities* (ISIC), Rev. 4.

3.7. IRES recommends that the statistical units should be the establishment and the household. The establishment is recommended because it is the most detailed unit for which the range of data required for energy statistics is normally available. To be analytically useful, data need to be grouped according to such characteristics as kind of activity, geographical area and size, and which is facilitated by the use of the establishment as the statistical unit.

Box 3.1:
Energy, Water and Environment Survey in Australia

The Energy, Water and Environment Survey (EWES) collects information on the usage of energy and water by a broad range of businesses in Australia. It was originally run as a one-off survey in 2008–2009, and again in 2011–2012 (albeit with a different sample and some content changes). This survey also collects information on environmental management activities. Data from the EWES constitute a vital input into the Energy Account, Australia (ABS cat. no. 4614.0), which provides statistics to monitor changes over time in the supply and use of energy within Australia from an economic and environmental perspective.

The scope of the 2008–2009 EWES was all Australian-based activities of business entities with a non-cancelled Australian Business Number and an active Income Tax Withholding or Goods and Service Tax role, including government-owned or controlled public trading enterprises, except for micro non-employers and those businesses classified according to the following Australian and New Zealand Standard Industrial Classification (ANZSIC) or Standard Institutional Sector Classification of Australia (SISCA) areas:

Agriculture (ANZSIC06 Subdivision 01)

Water Supply, Sewerage and Drainage Services (ANZSIC06 Subdivision 28)

Finance (ANZSIC06 Subdivision 62)

Insurance (ANZSIC06 Subdivision 63)

Public Administration (ANZSIC06 Subdivision 75)

Defence (ANZSIC06 Subdivision 76)

Private Households Employing Staff (ANZSIC06 Subdivision 96)

General Government (SISCA 3000)

Finance sector (SISCA 2110, 2121, 2129, 2131, 2132, 2133, 2141, 2142, 2191, 2199, 2301, 2309)

The EWES captures information on consumption (volume) and expenditures of electricity and natural gas from electricity/natural gas suppliers. Consumption and expenditure of other fuel types, including petrol, diesel, LPG, aviation turbine fuel, fuel oil, coal, coke and coal by-products, renewable fuels and crude oil are also collected. Electricity, natural gas and other fuels supplied to businesses on a non-monetary basis are included. Electricity, natural gas and other fuels on-sold to other businesses are specifically excluded. Also captured is electricity generation for sale or own use.

The Australian Bureau of Statistics Business Register is a list of "Type of Activity units" (TAU) defined as "one or more business entities, sub-entities or branches of a business entity within an enterprise group that can report production and employment data for similar economic activities. When a minimum set of data items is available, a TAU is created, which covers all the operations within an industry subdivision (and the TAU is classified to the relevant subdivision of the ANZSIC). Where a business cannot supply adequate data for each industry, a TAU is formed, which contains activity in more than one industry subdivision."

A population, based on the scope of EWES, is identified, and a sample is then taken from that population.

3.8. Statistical units are classified according to the type of economic activity they carry out in accordance with ISIC (or equivalent national classifications). Each statistical unit has therefore a corresponding ISIC code, which characterizes the principal activity of the unit involved.

3.9. The correspondence between the groups of economic activities that are relevant for energy statistics and the appropriate ISIC divisions will be discussed in the next sections. This is particularly important for collecting and presenting internationally comparable statistics.

3.10. It should be noted that these correspondences can be easily translated into a correspondence with the *Statistical Classification of Economic Activities in the European Community* (NACE), which is the classification in use in the European Statistical System. NACE is derived from ISIC but is more detailed than ISIC. Both have exactly the same items at the highest levels, but NACE is more detailed at lower levels. The same is the case for other related classifications, such as the *North American Industry Classification System* (NAICS) and the *Australian and New Zealand Standard Industrial Classification* (ANZSIC) and many national classifications used around the world.[1] Box 3.1 shows an example of the use of industry classification categories in an energy survey.

3.11. In energy statistics, a distinction, useful for the development of data collection strategies, is made between different groups of statistical units – energy industries, other energy producers and energy consumers. These groups are presented below.

1. Energy industries

3.12. Energy industries are defined as those economic units whose principal activity is primary energy production, energy transformation or energy distribution. The principal activity is determined in accordance with the standard top-down method described in ISIC.

3.13. The activities defining energy industries are very diverse and their detailed technical descriptions are quite complex. However, for the purposes of energy statistics, the activities of economic units belonging to energy industries can be conveniently identified by the establishments (plants) in which they occur. For example, typical representatives of primary production are coal mines and oil and gas extraction sites.

3.14. IRES recommends that, in order to improve cross-country comparability of statistics on energy production by energy industries, countries should identify, as far as feasible and applicable, the energy industries listed in the left column of table 3.1, which is equivalent to table 5.1 in IRES. The table also provides information on the ISIC, Rev. 4 division/group/class in which the different energy industries are included.

[1] For a more detailed discussion about the links between ISIC, Rev. 4 and the other classifications (such as NACE, NAICS and ANZSIC), please refer to the International Standard Industrial Classification of All Economic Activities (ISIC), Rev. 4 (United Nations, 2008).

Table 3.1:
Energy industries

Energy industry	ISIC Rev. 4
Electricity, CHP and heat plants[a]	Division 35 – Electricity, gas, steam and air conditioning supply
Pumped storage plants	
Coal mines	Division 05 – Mining of coal and lignite
Coke ovens	Group 191 – Manufacture of coke oven products
Coal liquefaction plants	Group 192 – Manufacture of refined petroleum products
Patent fuel plants	Group 192 – Manufacture of refined petroleum products
Brown coal briquette plants	Group 192 – Manufacture of refined petroleum products
Gas works[b] (and other conversion to gases)	Group 352 – Manufacture of gas; distribution of gaseous fuels through mains
Gas separation plants	Division 06 – Extraction of crude petroleum and natural gas
Gas-to-liquids (GTL) plants	Group 192 – Manufacture of refined petroleum products
Liquefied natural gas (LNG) plants/ re-gasification plants	Group 091 – Support activities for petroleum and natural gas extraction Class 5221 – Service activities incidental to land transportation
Blast furnaces	Group 241 – Manufacture of basic iron and steel
Oil and gas extraction	Division 06 – Extraction of crude petroleum and natural gas Group 091 – Support activities for petroleum and natural gas extraction
Oil refineries	Division 19 – Manufacture of coke and refined petroleum products
Charcoal plants[c]	Class 2011 – Manufacture of basic chemicals
Biogas production plants[d]	Group 352 – Manufacture of gas; distribution of gaseous fuels through mains
Nuclear fuel extraction and fuel processing	Class 0721 – Mining of uranium and thorium ores Class 2011 – Manufacture of basic chemicals
Other energy industry not elsewhere specified[e]	Class 0892 – Extraction of peat ….

[a] Also including the distribution of electricity and heat to consumers.

[b] Also including the distribution of these gases.

[c] The provided ISIC link refers to the production of charcoal through the distillation of wood. If charcoal is produced in the forest using traditional methods, the activity would be classified in ISIC 0220 – "Logging".

[d] Plants having the production of biogases as their main activity would be classified in ISIC class 3520. However, biogases may also be produced as by-products of other activities, such as those classified in ISIC 3700 – "Sewerage" and 3821 – "Treatment and disposal of non-hazardous waste".

[e] The given ISIC link provides an example, namely the extraction of peat, but is not exhaustive.

3.15. As can be seen from the table, the list of the energy industries is usually more detailed than the classes of ISIC, Rev 4. For example, coal liquefaction plants can be found in the same ISIC group 192 "Manufacture of refined petroleum products" as oil refineries, patent fuel plants and the brown coal briquettes plants. However, these plants are quite different because of the different inputs, transformation processes and energy products produced. Therefore, in energy statistics a distinction is made between coal liquefaction plants, patent fuel plants, brown coal briquettes plants and oil refineries to account for the different inputs and outputs of these plants.

3.16. Another example of more targeted detail for energy statistics are charcoal plants, which are included under ISIC class 2011 "Manufacture of basic chemicals" when the production of charcoal is carried out through the distillation of wood. ISIC class 2011 is broader in scope, as it includes the manufacture of many chemicals using basic processes, such as thermal cracking and distillation. The outputs of these processes are usually separate chemical elements or separate chemically-defined compounds covering more than just charcoal. ISIC class 2011 includes, for example, the manufacture of distilled water, the manufacture of synthetic aromatic products, etc. However, it should be noted that if charcoal is produced in the forest using traditional methods, the activity would be classified in ISIC class 0220 "Logging".

3.17. It is similar for blast furnaces used in the manufacture of iron and steel. The ISIC group 241 is broader in scope as it also covers a number of other activities within the manufacture of iron and steel, such as the production of ferro-alloys, granular iron and iron powder, steel in ingots or other primary forms, the production of cold-rolled or cold-drawn steel products, etc. Only the operation of blast furnaces is considered to be part of the energy industry (because of its production of blast furnace gas).

3.18. The ISIC correspondence is therefore useful for the identification of the energy industries, but its level of detail is not sufficient for energy statistics, and either a more detailed industry classification or an additional indicator identifying the energy industry units in these ISIC classes is needed.

2. Other energy producers

3.19. Other energy producers are defined in IRES as those economic units (including households) which choose, or are forced by circumstances, to produce energy for their own consumption and/or to supply energy to other units, but for which energy production is not the principal activity. These units are engaged in the production, transformation and transmission/distribution of energy as a secondary or ancillary activity. This means that the "energy" output generated by these activities and measured in terms of value added does not exceed that of the principal activity of the unit.[2] In the case of ancillary activities, the activities are carried out to support the principal and secondary activities of the unit.

3.20. For example, geographically remote industries may have no access to electricity unless they produce it themselves; iron and steel works requiring coke and the heat from it for their own production purposes will often produce their own coke and electricity. Sugar mills nearly always burn the bagasse they produce for generating steam, and process heat and electricity. An enterprise whose primary activity is the production of animal products (i.e., raising and breeding of pigs, sheep, etc.) could use its animal waste as fuel in a biogas system to generate electricity for its own use or to sell to a local market.

3.21. The most common examples of other energy producers refer to the production of heat and electricity. Many industrial establishments and commercial organizations may have electricity or heat generating equipment that they can turn on in the event of a failure in the public supply system (in which case they might even sell electricity to other consumers or to the public supply system). Households that use solar panels for generating electricity for their own use (and sometimes even for sale to third parties) are another example of other energy producers.

3.22. In order to have a full understanding of the energy supply and demand in a country, it is important that these energy production activities are also included in the official energy statistics. This information, however, may not be readily available as energy is not the primary business interest of these industries. Therefore, the collection of data on energy production by other energy producers might be a challenge.

3.23. IRES recommends that in countries where such producers account for a significant part of total energy production, every effort should be made to obtain from them the detailed data and incorporate them in their official energy statistics, including in

[2] Since the principal activity is determined using the top-down method and not simply by choosing the activity with the largest value-added contribution, this is not always the case, but should mostly be true in the scenarios considered for energy statistics.

the energy balance. Countries, where production of energy by other energy producers is small might limit their data collection from such industries to appropriate aggregates only or prepare estimates as necessary.

3.24. Care should be taken to ensure full coverage of the energy production activities of units which produce energy as a secondary activity, since these economic units are classified under a different division/group/class than those in table 3.1. In such cases, additional efforts should be made to collect data on energy production from economic units that are not classified as energy industries (including, for example, most manufacturing units, agriculture, fishing and forestry units, households, etc.). It should be noted that the share of secondary and ancillary activities in a unit's overall activities depends on the choice of unit. If the statistical unit is the establishment, the unit will be engaged in a smaller number of activities (ideally just one) than when the associated enterprise is chosen as the statistical unit.

3. Electricity and heat

3.25. To identify the producers of electricity and heat according to the type of producers and type of plants, the following lists of categories are used (but they are not formally designated as classifications).

(a) Two types of producers are identified:

- **Main activity producers**

  These are units which produce electricity or heat as their principal activity. Formerly known as public utilities, these units may be privately- or publicly-owned companies.

- **Autoproducers**

  - Autoproducers (electricity)

    These are units which produce electricity but for whom the production of electricity is not their principal activity.

  - Autoproducers (heat)

    These are units which produce heat for sale but for whom the production of heat is not their principal activity.

(b) Three types of generating plants are also identified for the producers of electricity and heat (both main and autoproducers):

- **Electricity plants**

  These refer to plants producing only electricity. The electricity may be obtained directly from natural sources, such as hydro, geothermal, wind, tidal, marine, solar energy or from fuel cells, or from the heat obtained from the combustion of fuels or nuclear reactions.

- **Combined heat and power (CHP) plants**

  These refer to plants which produce both heat and electricity from at least one generating unit in the plant. They are sometimes referred to as "co-generation" plants.

- **Heat plants**

  These refer to plants (including heat pumps and electric boilers) designed to produce heat only for delivery to third parties.

3.26. These lists apply to all producers of electricity and heat, i.e. to producers in energy industries, as well as other energy producers. Similar to comments in the previous sections, the allocation of producers to ISIC categories is relatively straightforward for main activity producers but may be linked to many ISIC classes for autoproducers.

4. Energy consumers

3.27. Energy consumers comprise economic units (establishments/enterprises and households) that act as final users of energy. In other words, they use energy products for energy purposes (heat raising, transportation and electrical services) and/or for non-energy purposes (e.g. manufacture of plastic products).

3.28. IRES recommends that countries identify, as far as feasible and applicable, the groups of energy consumers as listed in table 3.2 below, which is identical to IRES table 5.3. To facilitate the collection of energy statistics and their integration with other economic statistics, table 3.2 also provides a correspondence between the identified groups of energy consumers and the relevant categories of ISIC, Rev. 4. The scope of each consumer group is defined by the scope of the economic units belonging to the corresponding ISIC categories in tsable 3.2, except for "households", which includes all households in their capacity as final consumers and not only those engaged in economic activities (as covered by ISIC).

3.29. It should be noted that, while the level of detail presented in the left column of Table 3.2 is often the basis of international data collection, data at the detailed ISIC level (division/group/class) are important for all industries and their collection is encouraged.

Table 3.2:
Energy consumers

Energy consumers	Correspondence to ISIC Rev. 4
Manufacturing, construction and non-fuel mining industries	
Iron and steel	ISIC Group 241 and Class 2431 Note that the consumption of energy products in coke ovens and blast furnaces is excluded, as these plants are considered part of the energy industries.
Chemical and petrochemical	ISIC Divisions 20 and 21 Note that the consumption of energy products by plants manufacturing charcoal or carrying out the enrichment/production of nuclear fuels (both classified in ISIC 2011) is excluded, as these plants are considered part of the energy industries.
Non-ferrous metals	ISIC Group 242 and Class 2432
Non-metallic minerals	ISIC Division 23
Transport equipment	ISIC Divisions 29 and 30
Machinery	ISIC Divisions 25, 26, 27 and 28
Mining and quarrying	ISIC Divisions 07 and 08 and Group 099, excluding the mining of uranium and thorium ores (Class 0721) and the extraction of peat (Class 0892).
Food and tobacco	ISIC Divisions 10, 11 and 12
Paper, pulp and print	ISIC Divisions 17 and 18
Wood and wood products (other than pulp and paper)	ISIC Division 16
Textile and leather	ISIC Divisions 13, 14 and 15
Construction	ISIC Divisions 41, 42 and 43
Industries not elsewhere specified	ISIC Divisions 22, 31 and 32
Household	ISIC Divisions 97 and 98
Commerce and public services	ISIC divisions: 33, 36-39, 45-96 and 99, excluding ISIC 8422
Agriculture, Forestry	ISIC Divisions 01 and 02
Fishing	ISIC Division 03
Not elsewhere specified (including Defence activities)	ISIC Class 8422

3.30. The boxes on the following pages show examples of surveys conducted to gather information on energy users, demonstrating the classifications aspects of these surveys. Box 3.2 describes the survey of energy consumption in the industrial sector in Finland, based on the international standard classifications recommended in this *Manual*. Box 3.3 shows the classifications detail (in terms of the North American Industry Classifications System) used in the industrial consumption of energy survey in Canada. Box 3.4 describes the choice of statistical units in the energy consumption survey by industry in Denmark. Box 3.5 describes the concepts applied in the end use survey in Ghana, including the choice of unit and the industry breakdown. Both concepts deviate from the recommendations set out in IRES and do not follow a formal classification in Ghana but show a national adaptation of the recommended concepts.

Box 3.2:
Energy use in manufacturing survey in Finland

The aim of the inquiry on energy use in manufacturing is to describe the energy consumption of manufacturing. This inquiry has been conducted yearly since 2008. It collects information about the consumption of electricity, heat and a range of fuels, using the national standard industrial classification TOL 2008, which is based on NACE, Rev. 2, which in turn is based on ISIC, Rev. 4.

The population of the inquiry consists of all places of business that belong to Divisions B (Mining and quarrying, or C (Manufacturing) of ISIC, Rev. 4:

All places of business are available from the enterprise register. The sample is generally done by the two-digit ISIC level, but some of the groups are combined, while others are selected at the more detailed level. For example, Divisions C13 (Manufacture of textiles) and C14 (Manufacture of wearing apparel) are combined, while in Division 24, separate classes (C24.5.1 (Casting of iron) and C24.5.2 (Casting of steel)) at the more detailed level are chosen.

The sample may be split into three strata:

(i) Large-size enterprises, i.e. all enterprises fulfilling any of the following criteria:
- The total number of employees exceeds 1000.
- It is part of main grid industry.
- It belongs to an emissions trading scheme.
- Its functioning requires an environmental licence.
- It is a user of fuel, rarely used in Finland.

These enterprises are surveyed every year.

(ii) Medium-size enterprises, i.e. enterprises that fewer at least 10 employees but are not included in stratum (i).

For these enterprises, a sample is taken yearly.

(iii) Small-size enterprises, i.e. enterprises that have less than 10 employees and are not included in other strata.

For these enterprises, a sample is taken at intervals of five years. This inquiry is executed between the months of March and May. A part of the data is available from other administrative data sources and is excluded from this inquiry, which reduces the burden on the informant.

The data release is scheduled 10 to 11 months after the end of the statistical year. Collected data are released on levels of fuels (including electricity and heat), standard industrial classifications (ISIC, Rev. 4/NACE, Rev. 2) and regional breakdown (NUTS 2).

For more information, see www.stat.fi/til/tene/index_en.html.

Box 3.3: Industrial Consumption of Energy Survey in Canada

Since 1995, the Manufacturing and Energy Division at Statistics Canada has conducted the Industrial Consumption of Energy (ICE) survey to obtain information on the annual demand for energy of the industrial sector in Canada. The ICE survey includes about 4,600 manufacturing establishments and about 200 non-manufacturing units covering mining, oil and gas extraction and utilities located in Canada.

The statistics cover the consumption of energy by manufacturing companies classified under the North American Industry Classification System (NAICS).

Manufacturing (31–33):

- Food manufacturing (311)
- Beverage and tobacco product manufacturing (312)
- Textile mills (313)
- Textile product mills (314)
- Clothing manufacturing (315)
- Leather and allied product manufacturing (316)
- Wood product manufacturing (321)
- Paper manufacturing (322)
- Printing and related support activities (323)
- Petroleum and coal product manufacturing (324)
- Chemical manufacturing (325)
- Plastics and rubber products manufacturing (326)
- Non-metallic mineral product manufacturing (327)
- Primary metal manufacturing (331)
- Fabricated metal product manufacturing (332)
- Machinery manufacturing (333)
- Computer and electronic product manufacturing (334)
- Electrical equipment, appliance and component manufacturing (335)
- Transportation equipment manufacturing (336)
- Furniture and related product manufacturing (337)
- Miscellaneous manufacturing (339)

The ICE survey covers the following fuels: electricity, natural gas, propane, diesel, light fuel oil, kerosene and other middle distillates, heavy fuel oil, wood and wood waste, spent pulping liquor, refuse, steam, coal (bituminous, sub-bituminous, lignite and anthracite), coal coke, coal tar, light coal oil, coke oven gas, petroleum coke, refinery fuel gas, coke on catalyst, bitumen emulsion, ethane, butane, naphtha, by-product gas, flared gas, steam sales and others.

Information is also collected on the different usages of energy commodities: as fuel, to produce steam for sale, to produce electricity and for non-energy use, as well as reasons for changes in energy consumption.

For more information, see www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=5047.

Box 3.4:
Choice of the statistical unit in the survey on energy consumption by industry in Denmark

The purpose of the survey – energy consumption by industry – is to analyse the volume and composition of energy consumption in the manufacturing industries. Data on the manufacturing industries' energy consumption are available from 1980, and have been compiled every second or third year ever since.

The collected information is part of the overall energy statistics (energy balances) and is used in statistical reports in several environmental areas, in environmental-economic accounts (energy accounts) and national accounts. This information is also important as part of Denmark's energy and contingency planning.

The survey of industrial energy consumption is a cut-off census and covers enterprises with more than 20 employees – representing about 90 per cent of industrial energy consumption – and all establishments within the manufacturing industries related to those enterprises. The number of establishments is around 4,000, which are classified according to the Danish Industrial Classification, which is based on the European Union classification (NACE).

When conducting the census, a questionnaire is sent to every establishment. This is key for several reasons, including: (a) the availability of the information, since information on the physical use of energy is usually only available at the establishment level where the meters are located and the measurements recorded; and (b) the geographical coverage, since having information at the establishment level is crucial in order to gather regional statistics on consumption.

Furthermore, the questionnaire asks for information on energy consumption in both physical and monetary units. This allows for additional uses of the data, e.g. as input in the compilation of national accounts, as well as checks during the compilation of the data.

For more information, www.dst.dk/en/Statistik/dokumentation/documentationofstatistics/energy-consumption.

Box 3.5:
Statistical Unit Used for End-use Survey in Ghana

The Energy Commission of Ghana, as part of its mandate as stated in the *Energy Commission Act (Act 541)*, conducted a nationwide survey to collect data on the supply and consumption of energy in the residential, industrial, and commercial and services sectors of the economy. The data collected from the survey were used to calculate end-use fuel shares, the level of energy use and specific energy use by end-use activities. The survey is planned to be conducted every 3–5 years.

For the residential sector, the statistical unit used in the survey is a household. A household is defined as a person or a group of persons who live together in the same dwelling, share the same house-keeping arrangements and are catered for as one unit.

In the case of the commercial and services sector, the statistical unit is the premises, which is defined as a building or collection of buildings and/or specific floor area in a location undertaking commercial or service business. The premises may consist of several buildings that are all occupied by the same business entity and each building may have several floors and floor areas. The premises may also be a portion of a building occupied by a business entity.

Box 3.5 (continued)

The statistical unit for industry is the establishment.

The sectoral classification was done to ensure a homogeneous energy use pattern. The classification for the various sectors and subsectors is presented below:

(a) Households:	Urban	• Accra/Tema
		• Other Urban
	Rural	• Forest
		• Coastal
		• Savannah
(b) Commercial and Services		• Offices
		• Hotels
		• Health facilities
		• Stores and supermarkets
		• Schools
		• Restaurants and chopbars[3]
(c) Industry:	Formal	• Food
		• Beverages and tobacco
		• Textile and textile products
		• Leather and allied products
		• Printing and related support
		• Plastics and rubber support
		• Non-metallic mineral products
		• Primary metal
		• Fabricated metal products
		• Furniture and related support
		• Chemical and pharmaceutical products
		• Machinery
		• Computer and electronic products
		• Electrical equipment, appliances and components
		• Sawmill and planing of wood
		• Miscellaneous
	Informal:	• Alcohol brewing
		• Bakeries
		• Clay firing
		• Fish smoking
		• Gari processing
		• Oil production
		• Sawmill and planing of wood

[3] Traditional restaurants which use mainly wood and charcoal as fuel.

Box 3.5 (continued)

Data collected during the survey include data on electricity, diesel, gasoline, residual fuel oil, liquefied petroleum gas (LPG), kerosene, premix (a blend of gasoline and diesel), wood and charcoal.

The Commission also publishes the annual energy statistics, which contain data on the supply and consumption of energy in the country. Statistics published in the annual energy statistics are mainly made up of data collected from the various institutions within and outside the country's energy sector.

C. Classification of energy products

3.31. The classification of energy products presented in IRES is the Standard International Energy Product Classification (SIEC), which was developed as part of the preparation of IRES. This classification is an important milestone in energy statistics as it represents the first internationally agreed classification of energy products based on a set of internationally unified product definitions.

3.32. SIEC was the result of a harmonization process of definitions used by international, regional and supranational organizations involved in the collection and compilation of energy statistics. A classification structure, with a coding system, was then developed based on the harmonized definitions. SIEC went through a wide consultation process with countries, as well as with experts on statistical standard classifications.

3.33. The benefits of a standard international classification of energy products are numerous: it focuses solely on products that are relevant to understanding the energy situation (thus excluding products not related to energy); it provides definitions for all energy products and reduces the ambiguity among different energy products; it facilitates data collection and allows for a better integration of data collected in different statistical domains (thus reducing the response burden due to multiple data requests); and it improves the comparability of data. However, SIEC still faces a number of challenges, mainly because of two interlinked reasons: on the one hand, SIEC is based on mixed classification criteria, which depend partly on the purpose of use and partly on the characteristics of the product; and on the other hand, on the correspondence between SIEC and the Central Product Classification (CPC) and the Harmonized Commodity Description and Coding System (Harmonized System or HS), which is not always straightforward. The latter is a particularly important point as a better correspondence would facilitate the integration of statistics collected across statistical domains (e.g. economic and trade statistics) and may reduce the response burden on the data providers.

3.34. The full SIEC classification is provided in annex 3A, with the correspondence to recent product classifications, namely CPC, Ver. 2, CPC, Ver. 2.1, HS 2007, HS 2012 and HS 2017. The correspondences with CPC, Ver. 2 and HS 2007 were established at the time of the creation of IRES. The other international classifications have since been revised (CPC, Version 2.1 and HS 2017 are currently in effect). Updated correspondence tables with the new and future classifications will also be made available on the Statistics Division website.

3.35. It should be noted that in energy statistics the distinction between primary and secondary energy products, as well as renewable and non-renewable energy, is often made and reflected both in the compilation of energy balances and in the calculation of indicators for policymaking. Although these are not explicit classification criteria in SIEC, in many cases a complete SIEC category can clearly be assigned to one or the other set. IRES provides the list of products considered primary or secondary and renewable or non-renewable.

Why was CPC not enough?

3.36. The *Central Product Classification* (CPC) is a comprehensive and multipurpose classification of all products. Its main purpose is to provide a framework for the international comparison of statistics on products. Among other roles, it is intended to function as a basis for recompiling basic statistics from their original classifications into a standard classification for analytical use. Since its inception in 1990, the goods part of CPC has been largely built upon and influenced by the HS used in international trade, but deviations have been made in areas where the level of detail in the HS was not sufficient to measure domestic production, such as for agricultural products.

3.37. Because of its multipurpose nature, CPC was not developed with a particular focus on specific subject areas like energy statistics. Therefore, it does not fully reflect and respond to the needs of energy statistics; energy products are often not explicitly identified within a CPC class and may sometimes cover different CPC classes, and there is no higher aggregate for energy products.

3.38. On the other hand, SIEC was developed to reflect more appropriately the products relevant for a better understanding of the supply, availability and consumption of energy in a country. The definitions that form the basis of SIEC were only recently internationally agreed upon, and with significant international effort. Until this point, the concepts and definitions in energy statistics had been developed independently and varied across organizations and countries. During the discussions to agree on the energy product definitions, no explicit consideration of CPC or HS concepts was included, adding to the differences between SIEC and these classifications.

3.39. The criteria and concepts used by CPC and SIEC are therefore quite different, leading to differing product definitions and hierarchies. In some cases, the differences might reasonably be overcome, whereas in a few cases the underlying criteria are so different that it will be very difficult to reconcile the two systems.

What are the main issues in the correspondence between SIEC, CPC and HS?

3.40. As indicated above, the categories do not always match across the three classifications due to the specific uses/needs of energy statistics which are more adequately addressed in SIEC as compared to the uses/needs provided for under CPC and HS.

3.41. One notable issue in the correspondence between the classifications is related to the definitions of the categories since there are a number of products in SIEC which are defined in terms of their use rather than their physical characteristics (which is the underlying classification criterion in CPC and HS). This is the case of "refinery feedstocks", for example, which are defined as consisting of "oils or gases from crude oil refining or the processing of hydrocarbons in the petrochemical industry which are destined for further processing in the refinery, excluding blending. Typical feedstocks include naphthas, middle distillates, pyrolysis gasoline and heavy oils from vacuum distillation and petrochemical plants." Feedstocks may cover a wider range of products, including naphthas (HS 2710.11) and pyrolysis gasoline (HS 2707.50), among others. Since the definition of feedstocks is primarily based on intended use rather than the types of products it covers, giving an explicit CPC/HS link could be misleading.

3.42. Another example is "additives and oxygenates", which are defined as "compounds added to or blended with oil products to modify their properties (octane, cetane, cold properties, etc.)". The examples provided in the definition are: (a) oxygenates, such as alcohols (methanol, ethanol) and ethers – MTBE (methyl tertiary butyl ether), ETBE (ethyl tertiary butyl ether), TAME (tertiary amyl methyl ether); (b) esters (e.g. rapeseed or dimethylester, etc.); and (c) chemical compounds (such as tetramethyl lead (TML), tetraethyl lead (TEL) and detergents). Some additives/oxygenates may be derived from biomass, while others may be of fossil hydrocarbon origin.

3.43. Similar to feedstocks, the "additives and oxygenates" category covers a number of different products that are added to or blended with oil products. The correspondence between SIEC, CPC and HS for these categories is expected to remain problematic. The examples listed above in the definition are, for example, included in HS categories 2207, 2905.11, 2909.19, etc.

3.44. In other cases, the definition of the product is not consistent across classifications. This is the case of biodiesel, for example. While SIEC distinguishes the pure "bio" component in the biofuel section, the newly introduced category "biodiesel" (HS 2012 code 3826) covers also the blended mixture containing less than 70 per cent by weight of petroleum oils or oils obtained from bituminous minerals. Therefore, when using custom data for the trade of biodiesel, adjustments are necessary to properly use the data in energy statistics.

3.45. Another issue is related to the different levels of detail of the different classifications. This leads for example to a SIEC category linked to categories (sometimes broader in scope) belonging to different CPC divisions. For example, SIEC 511 – Fuelwood, wood residues and by-products – can be matched with a set of three CPC subclasses taken from three different divisions of CPC: CPC 03130–Fuel wood, in logs, in billets, in twigs, in faggots or in similar forms; CPC 31230–Wood in chips or particles; and CPC 39280–Sawdust and wood waste and scrap. It is not possible to further divide this aggregation, since SIEC 5119 is linked to all three of the related CPC subclasses.

3.46. HS also does not provide much detail for oil products, resulting in the fact that many of the products in SIEC Division 46 are linked to two HS 2007 subheadings: 2710.11 and 2710.19. This particular issue is evident when import/export data for oil products, such as motor gasoline or kerosene, are sought to compile an energy balance. The relevant categories at the most detailed level of HS 2007 distinguish only "Light oils and preparations" (2710.11) and "Other" (2710.19) – two categories to which 13 different SIEC products could be linked. HS 2012 essentially maintains this level of detail but adds a separate category for oil products containing biodiesel.[4] While the level of detail in HS at the international level is thus not suitable for an easy conversion to SIEC, and therefore the use of customs data, many countries have adapted HS to a more detailed national version. Such national versions often provide detail that allows the direct use of customs data for energy statistics. An example is given in box 3.6, showing (part of) the expanded version of chapter 27 of the Philippine Standard Commodity Classification (PSCC 2015).

3.47. At the international level, further discussions with the World Customs Organization and its Harmonized System Committee will be needed to establish more detail for these types of products in HS. CPC, on the other hand, already provides detail that more closely reflects the needs of energy statistics, identifying another area where it deviates from HS by providing additional detail.

3.48. There are also products in SIEC that are not reflected at all in HS or CPC, since they do not meet product criteria in those classifications. An example would be Black liquor (SIEC 514). Since this product is primarily produced and used by the same unit (typically a paper manufacturer), there are no market transactions and no international trade, which are the main focus of CPC and HS applications. Another example may be recovered gases, for which no significant international trade exists.

[4] To indicate the change in scope of the original HS subheading, 2710.11 was renumbered 2710.12 in HS 2012. It should be noted that the definition of the new subheading 2710.20 does not reflect the data collection requirements for biodiesel products in SIEC and therefore does not improve the link between HS and SIEC for actual data conversions.

Box 3.6:
Philippine Standard Commodity Classification (PSCC 2015)

Heading 2710 of the PSCC 2015 is subdivided into the following categories:

2710.12	Light oils and preparations:
2710.12.11	of RON* 97 and above, leaded
2710.12.11-01	Bioethanol blended
2710.12.11-09	Other
2710.12.12	of RON 97 and above, unleaded
2710.12.12-01	Bioethanol blended
2710.12.12-09	Other
2710.12.13	of RON 90 and above, but below RON 97, leaded
2710.12.13-01	Bioethanol blended
2710.12.13-09	Other
2710.12.14	of RON 90 and above, but below RON 97, unleaded
2710.12.14-01	Bioethanol blended
2710.12.14-09	Other
2710.12.15	Other, leaded
2710.12.15-01	Bioethanol blended
2710.12.15-09	Other
2710.12.16	Other, unleaded
2710.12.16-01	Bioethanol blended
2710.12.16-09	Other
2710.12.20-00	Aviation spirit, not of a kind used as jet fuel
2710.12.30-00	Tetrapropylene
2710.12.40-00	White spirit
2710.12.50-00	Low aromatic solvents containing by weight less than 1% aromatic content
2710.12.60-00	Other solvent spirits
2710.12.70-00	Naphtha, reformates and other preparations of a kind used for blending into motor spirits
2710.12.80-00	Other alpha olefins
2710.12.90-00	Other
2710.19	Other:
2710.19.20-00	Topped crudes
2710.19.30-00	Carbon black feedstock
2710.19.41-00	Lubricating oil feedstock
2710.19.42-00	Lubricating oils for aircraft engines
2710.19.43-00	Other lubricating oils
2710.19.44-00	Lubricating greases
2710.19.50-00	Hydraulic brake fluid
2710.19.60-00	Transformer and circuit breakers oils
2710.19.71-00	Automotive diesel fuel
2710.19.72-00	Other diesel fuels
2710.19.79	Fuel oils
2710.19.79-01	Bunker oil

Box 3.6 (continued)

2710.19.79-02	Gas oils
2710.19.79-09	Other
2710.19.81-00	Aviation turbine fuel (jet fuel) having a flash point of 23º xC or more
2710.19.82-00	Aviation turbine fuel (jet fuel) having a flash point of less than 23º C
2710.19.83	Other kerosene
2710.19.83-01	Lamp kerosene
2710.19.83-02	Other kerosene, including vaporising oil
2710.19.89	Other medium oils and preparations
2710.19.89-01	Normal paraffin
2710.19.89-09	Other
2710.19.90	Other
2710.19.90-01	Mineral oil
2710.19.90-02	Low aromatic solvents of heavy petroleum oils
2710.19.90-03	Other solvents, heavy
2710.19.90-09	Other
2710.20.00	Petroleum oils and oils obtained from bituminous minerals (other than crude) and preparations not elsewhere specified or included, containing by weight 70 per cent or more of petroleum oils or of oils obtained from bituminous minerals
2710.20.00-01	High speed diesel fuel
2710.20.00-02	Other diesel fuel
2710.20.00-03	Other light oil
2710.20.00-09	Other

Source: http://nap.psa.gov.ph/activestats/pscc/.

* RON: research octane number

3.49. Since national adaptations of CPC and HS sometimes include details that are more applicable for energy statistics purposes, such as the creation of separate categories for motor gasoline and for other products in the national external trade classification, the linkages between SIEC and national classifications may be more straightforward in those countries. Box 3.7 below shows an example of SIEC links to nationally used product classifications in Azerbaijan.

Box 3.7:
Energy product classification in Azerbaijan

The National Classification of Energy Products was developed in Azerbaijan on the basis of IRES, agreed to with national authorities implementing energy policy in the country and approved by relevant decrees of the State Statistical Committee and State Committee for Standardization, Metrology and Patent of the Republic of Azerbaijan.

The classification ensures comparability of data at the international level and, at the same time, provides links with the various classifications used at the country level. The National Classification of Energy Products contains correspondence links not only to CPC and HS, but it also provides links to the Statistical Classification of Products by Activity (CPA) and the List of Products of the European Community (PRODCOM). Since CPA and PRODCOM classifications are used in collecting statistics on the production and consumption of

Box 3.7 (continued)

products in Azerbaijan, their harmonization with the National Classification of Energy Products is critically important and more suitable for users.

An example of the National Classification of Energy Products of Azerbaijan is provided below for the first few products.

Energy Product Classification

Section/ Division/ Group	Class	Products	Correspondences			
			CPA 2008	2008 PRODCOM	CPC, Ver. 2	HS 2007
0		Coal				
01		Hard coal				
011	0110	Anthracite	05.10.10*	05.10.10.10.1	11010*	2701.11
012		Bituminous coal				
	0121	Coking coal	05.10.10*	05.10.10.10.2	11010*	2701.19
	0129	Other bituminous coal	05.10.10*	05.10.10.20.0	11010*	2701.12
02		Brown coal, Lignite				
021	0210	Sub-bituminous coal	05.20.10	05.20.10.00.1	11030*	2702.10*

The classification also facilitates the compilation of the energy balances as it provides a clear aggregation scheme for the energy product categories of the balances. In addition, by having a clear hierarchical structure with detailed definitions, the National Classification of Energy Products has allowed for an improved coordination of energy statistics with other branches of economic statistics.

What is being done to improve the links between SIEC, CPC and HS?

3.50. Ideally, the three classifications, SIEC, CPC and HS, should be fully compatible, that is, the definitions of the categories, at least at the lowest level of detail, should be the same. This would foster data integration and reduce the response burden. While SIEC is the first internationally agreed classification of energy products, which is a major step forward in energy statistics, it also clearly identifies areas for improvement/refinement to better align it with CPC and HS. Some of those examples have been given above. To improve the linkages between these classifications, changes are required on all sides. There are a number of areas where CPC and HS could better accommodate the detail requirements of energy statistics and some areas where SIEC definitions could be brought more in line with other standards. It is expected that improvements will take place over time.

3.51. Some changes have already been made in CPC, Ver. 2.1 to improve the linkages and reduce differences in terminology. However, not all problems could be solved, such as those resulting sometimes from the ambiguity in SIEC definitions, and additional work is being carried out to identify further changes that could be undertaken in the next update of CPC.

3.52. In addition, SIEC, like other statistical standard classifications, is expected to be reviewed and updated on a regular basis and further improvement will be considered in the next update of the classification.

D. Classification of mineral and energy resources

3.53. Data items on mineral and energy resources are important for the assessment of their availability in the environment, and of their depletion. This information is often used in the compilation of asset accounts in the *System of National Accounts* (SNA), as well as in the *System of Environmental-Economic Accounting* (SEEA). While information on the resources is not needed to build an energy balance, it is an important set of data that is needed to inform on issues related to energy sustainability and energy security. Consequently, classifications used for the measurement of energy resources should be consistent with concepts used in standard components of energy statistics.

3.54. There are two dimensions of classifications to consider for the description and measurement of energy resources – the type of energy product that can be extracted from those resources and the characteristics of the resources.

Types of energy resources

3.55. The type of energy resources is determined by the energy product extracted from the resource. While theoretically a detailed classification like SIEC could be used to distinguish the different types of energy resourcesl, in practice a list of aggregated categories is used, as shown in table 3.3. This list, though not formally designated as a classification, sets out the types of resources for which work on classifying and measuring resources has been carried out or is being planned in the context of the United Nations Framework Classification for Resources (UNFC).

Table 3.3:
Types of energy resources

Type of energy resources	Related energy products (SIEC)
Mineral and energy resources	
Oil resources	20 Oil shale/oil sands 41 Conventional crude oil 42 Natural gas liquids
Natural gas resources	30 Natural gas 42 Natural gas liquids
Coal and peat resources	01 Hard coal 02 Brown coal 11 Peat
Uranium and other nuclear fuels	9101 Uranium ores
Natural timber resources	511 Fuelwood, wood residues and by-products
Renewable sources	
Solar	70 Electricity 80 Heat
Hydro	70 Electricity
Wind	70 Electricity
Wave and tidal	70 Electricity
Geothermal	70 Electricity 80 Heat
Other electricity and heat	70 Electricity 80 Heat

Characteristics of energy resources

3.56. UNFC 2009 is a scheme for classifying and evaluating mineral and energy resources. Mineral and energy resources are distinguished according to their characteristics in terms of socioeconomic viability (dimension "E", 3 levels), project feasibility (dimension "F", 4 levels) and geological knowledge (dimension "G", 4 levels), thus allocating each resource/project to a three-digit code that represents these three dimensions. Some of the levels in dimensions E and F are further subdivided. (See https://unece.org/DAM/energy/se/pdfs/UNFC/unfc2009/UNFC2009_ES39_e.pdf for details.)

3.57. UNFC provides detailed definitions and specifications for the application of the classification. These specifications have been built on, and are also consistent with, other existing resources classification systems used for specific industries, such as the Committee for Mineral Reserves International Reporting Standards (CRIRSCO) for mineral resources and the Petroleum Resources Management System (PRMS) for oil resources.

3.58. The application of UNFC and the development of related guidance are currently being developed for renewable energy sources.

3.59. While UNFC thus provides a detailed classification of mineral and energy reserves, for statistical purposes, such as SEEA, a more aggregated version is used. This essentially combines categories of UNFC into three main classes for known deposits, defined as follows:

Table 3.4:
Categorization of mineral and energy resources

SEEA Class	Corresponding UNFC-2009 project categories		
	E Economic and social viability	F Field project status and feasibility	G Geological knowledge
Class A: Commercially recoverable resources	E1. Extraction and sale have been confirmed to be economically viable.	F1. Feasibility of extraction by a defined development project or mining operation has been confirmed.	Quantities associated with a known deposit that can be estimated with a high (G1), moderate (G2) or low (G3) level of confidence.
Class B: Potentially commercially recoverable resources	E2. Extraction and sale are expected to become economically viable in the foreseeable future.	F2.1 Project activities are ongoing to justify development in the foreseeable future, or F2.2 Project activities are on hold and/or where justification as a commercial development may be subject to significant delay.	
Class C: Non-commercial and other known deposits	E3. Extraction and sale are not expected to become economically viable in the foreseeable future or evaluation is at too early a stage to determine economic viability.	F2.2 Project activities are on hold and/or where justification as a commercial development may be subject to significant delay, or F2.3 There are no current plans to develop or to acquire additional data due to limited potential, or F4. No development project or mining operation has been identified.	

3.60. Note that the above categories do not include potential deposits, since they are outside of the scope of SEEA or IRES.

3.61. While UNFC already defines an aggregated level grouping – "Commercial projects", "Potentially commercial projects", "Non-commercial projects" and others – SEEA deviates from these aggregations and defines the similarly named, but different aggregations shown above.

Annex 3A: Standard International Energy Product Classification (SIEC)

SIEC headings			Correspondences				
			Central Product Classification		Harmonized System		
Section/ Division/ Group	Class		CPC, Ver. 2	CPC, Ver. 2.1	HS 2007	HS 2012	HS 2017
0		**Coal**					
01		**Hard coal**					
011	0110	Anthracite	11010*	11010*	2701.11	2701.11	2701.11
012		Bituminous coal					
	0121	Coking coal	11010*	11010*	2701.19	2701.19	2701.19
	0129	Other bituminous coal	11010*	11010*	2701.12	2701.12	2701.12
02		**Brown coal**					
021	0210	Sub-bituminous coal	11030*	11030*	2702.10*	2702.10*	2702.10*
022	0220	Lignite	11030*	11030*	2702.10*	2702.10*	2702.10*
03		**Coal products**					
031		Coal coke					
	0311	Coke oven coke	33100*	33100*	2704*	2704*	2704*
	0312	Gas coke	33100*	33100*	2704*	2704*	2704*
	0313	Coke breeze	33100*	33100*	2704*	2704*	2704*
	0314	Semi cokes	33100*	33100*	2704*	2704*	2704*
032	0320	Patent fuel	11020	11020	2701.20	2701.20	2701.20
033	0330	Brown coal briquettes (BKB)	11040	11040	2702.20	2702.20	2702.20
034	0340	Coal tar	33200*	33200*	2706	2706	2706
035	0350	Coke oven gas	17200*	17200*	2705*	2705*	2705*
036	0360	Gas works gas (and other manufactured gases for distribution)	17200*	17200*	2705*	2705*	2705*
037		Recovered gases					
	0371	Blast furnace gas	17200*	17200*	2705*	2705*	2705*
	0372	Basic oxygen steel furnace gas	17200*	17200*	2705*	2705*	2705*
	0379	Other recovered gases	17200*	17200*	2705*	2705*	2705*
039	0390	Other coal products	33500*, 34540*	33500*, 34540*	2707, 2708.10*, 2708.20*, 2712.90*	2707, 2708.10*, 2708.20*, 2712.90*	2707, 2708.10*, 2708.20*, 2712.90*
1		**Peat and peat products**					
11		**Peat**					
111	1110	Sod peat	11050*	11050*	2703*	2703*	2703*
112	1120	Milled peat	11050*	11050*	2703*	2703*	2703*
12		**Peat products**					
121	1210	Peat briquettes	11050*	11050*	2703*	2703*	2703*
129	1290	Other peat products	11050*, 33100*, 33200*, 33500*	11050*, 33100*, 33200*, 33500*	2703*, 2704*, 2706*, 2712.90*	2703*, 2704*, 2706*, 2712.90*	2703*, 2704*, 2706*, 2712.90*
2		**Oil shale / oil sands**					
20		**Oil shale / oil sands**					
200	2000	Oil shale / oil sands	12030	12030	2714.10	2714.10	2714.10
3		**Natural gas**					
30		**Natural gas**					

SIEC headings			Correspondences				
			Central Product Classification		Harmonized System		
Section/ Division/ Group	Class		CPC, Ver. 2	CPC, Ver. 2.1	HS 2007	HS 2012	HS 2017
300	3000	Natural gas	12020	12020	2711.11, 2711.21	2711.11, 2711.21	2711.11, 2711.21
4		**Oil**					
41		**Conventional crude oil**					
410	4100	Conventional crude oil	12010*	12010*	2709*	2709*	2709*
42		**Natural gas liquids (NGL)**					
420	4200	Natural gas liquids (NGL)	33420*	33421*, 33429*	2711.14, 2711.19*, 2711.29*	2711.14, 2711.19*, 2711.29*	2711.14, 2711.19*, 2711.29*
43		**Refinery feedstocks**					
430	4300	Refinery feedstocks	[a]	[a]	[a]	[a]	[a]
44		**Additives and oxygenates**					
440	4400	Additives and oxygenates	34131*, 34139*, 34170*, others	34131*, 34139*, 34170*, others	2207.20*, 2905.11, 2909.19*, others	2207.20*, 2905.11, 2909.19*, others	2207.20*, 2905.11, 2909.19*, others
45		**Other hydrocarbons**					
450	4500	Other hydrocarbons	12010*, 34210*	12010*, 34210*	2709*, 2804.10	2709*, 2804.10	2709*, 2804.10
46		**Oil products**					
461	4610	Refinery gas	33420*	33429*	2711.29*	2711.29*	2711.29*
462	4620	Ethane	33420*	33429*	2711.19*, 2711.29*	2711.19*, 2711.29*	2711.19*, 2711.29*
463	4630	Liquefied petroleum gases (LPG)	33410	33410	2711.12, 2711.13	2711.12, 2711.13	2711.12, 2711.13
464	4640	Naphtha	33330*	33330	2710.11*	2710.12*	2710.12*
465		Gasolines					
	4651	Aviation gasoline	33310*	33312	2710.11*	2710.12*	2710.12*
	4652	Motor gasoline	33310*	33311	2710.11*	2710.12*	2710.12*
	4653	Gasoline-type jet fuel	33320	33320	2710.11*	2710.12*	2710.12*
466		Kerosenes					
	4661	Kerosene-type jet fuel	33342	33342	2710.19*	2710.19*	2710.19*
	4669	Other kerosene	33341	33349	2710.19*	2710.19*	2710.19*
467		Gas oil / diesel oil and Hhavy gas oil					
	4671	Gas oil / diesel oil	33360*	33360*	2710.19*	2710.19*, 2710.20*	2710.19*, 2710.20*
	4672	Heavy gas oil	33360*	33360*	2710.19*	2710.19*, 2710.20*	2710.19*, 2710.20*
468	4680	Fuel oil	33370	33370	2710.19*	2710.19*	2710.19*
469		Other oil products					
	4691	White spirit and special boiling point industrial spirits	33330*	33350	2710.11*	2710.12*	2710.12*
	4692	Lubricants	33380*	33380	2710.19*	2710.19*	2710.19*
	4693	Paraffin waxes	33500*	33500*	2712.20*	2712.20*	2712.20*
	4694	Petroleum coke	33500*, 34540*	33500*, 34540*	2708.20*, 2713.11, 2713.12	2708.20*, 2713.11, 2713.12	2708.20*, 2713.11, 2713.12
	4695	Bitumen	33500*	33500*	2713.20	2713.20	2713.20

SIEC headings			Correspondences				
			Central Product Classification		Harmonized System		
Section/ Division/ Group	Class		CPC, Ver. 2	CPC, Ver. 2.1	HS 2007	HS 2012	HS 2017
	4699	Other oil products n.e.c.	33330*, 33350*, 33380*, 33420*, 33500*, 34540*	33390*, 33421*, 33429*, 33500*, 34540*	2707*, 2708.10*, 2710.11*, 2710.19*, 2711.14*, 2712.10*, 2712.20*, 2712.90*, 2713.90	2707*, 2708.10*, 2710.12*, 2710.19*, 2710.20*, 2711.14*, 2712.10*, 2712.20*, 2712.90*, 2713.90	2707*, 2708.10*, 2710.12*, 2710.19*, 2710.20*, 2711.14*, 2712.10*, 2712.20*, 2712.90*, 2713.90
5		**Biofuels**					
51		**Solid biofuels**					
511		Fuelwood, wood residues and by-products					
	5111	Wood pellets	39280*	39281	4401.30*	4401.31	4401.31
	5119	Other fuelwood, wood residues and by-products	03130, 31230, 39280*	03131, 03132, 31230, 39282, 39283	4401.10, 4401.21, 4401.22, 4401.30*	4401.10, 4401.21, 4401.22, 4401.39	4401.11, 4401.12, 4401.21, 4401.22, 4401.39, 4401.40
512	5120	Bagasse	39140*	39141	2303.20*	2303.20*	2303.20*
513	5130	Animal waste	34654*	34654*	3101*	3101*	3101*
514	5140	Black liquor	39230*	39230*	3804.00*	3804.00*	3804.00*
515	5150	Other vegetal material and residues	01913, 21710, 34654*, 39120*, 39150*	01913, 21910, 34654*, 39120*, 39150*	0901.90*, 1213, 1802*, 2302*, 2304, 2305, 2306, 3101	0901.90*, 1213, 1802*, 2302*, 2304, 2305, 2306, 3101	0901.90*, 1213, 1802*, 2302*, 2304, 2305, 2306, 3101
516	5160	Charcoal	34510	34510	4402	4402	4402
52		**Liquid biofuels**					
521	5210	Biogasoline	34131*, 34139*, 34170*	34131*, 34139*, 34170*	2207.20*, 2905.11*, 2905.13*, 2905.14*, 2909.19*	2207.20*, 2905.11*, 2905.13*, 2905.14*, 2909.19*	2207.20*, 2905.11*, 2905.13*, 2905.14*, 2909.19*
522	5220	Biodiesels	35490*	35491	3824.90*	3826.00*	3826.00*
523	5230	Bio jet kerosene					
529	5290	Other liquid biofuels					
53		**Biogases**					
531		Biogases from anaerobic fermentation					
	5311	Landfill gas	33420*	33429*	2711.29*	2711.29*	2711.29*
	5312	Sewage sludge gas	33420*	33429*	2711.29*	2711.29*	2711.29*
	5319	Other biogases from anaerobic fermentation	33420*	33429*	2711.29*	2711.29*	2711.29*
532	5320	Biogases from thermal processes					
6		**Waste**					
61		**Industrial waste**					

SIEC headings			Correspondences				
			Central Product Classification		Harmonized System		
Section/ Division/ Group	Class		CPC, Ver. 2	CPC, Ver. 2.1	HS 2007	HS 2012	HS 2017
610	6100	Industrial waste	3921, 39220, 39240, 39250, 39260, 39270, 39290	3921, 39220, 39240, 39250, 39260, 39270, 39290	2525.30, 2601, 3915, 4004, 4012.20, 4115.20, 4707, 5003, 5103.20, 5103.30, 5104, 5202, 5505, 6309, 6310	2525.30, 2601, 3915, 4004, 4012.20, 4115.20, 4707, 5003, 5103.20, 5103.30, 5104, 5202, 5505, 6309, 6310	2525.30, 2601, 3915, 4004, 4012.20, 4115.20, 4707, 5003, 5103.20, 5103.30, 5104, 5202, 5505, 6309, 6310
62		**Municipal waste**					
620	6200	Municipal waste	39910	39910	3825.10	3825.10	3825.10
7		**Electricity**					
70		**Electricity**					
700	7000	Electricity	17100	17100	2716	2716	2716
8		**Heat**					
80		**Heat**					
800	8000	Heat	17300	17300	2201.90*	2201.90*	2201.90*
9		**Nuclear fuels and other fuels n.e.c.**					
91		**Uranium and plutonium**					
910		Uranium and plutonium					
	9101	Uranium ores	13000*	13000*	2612.10	2612.10	2612.10
	9109	Other uranium and plutonium	33610, 33620, 33630*, 33710, 33720	33610, 33620, 33630*, 33710, 33720	2844.10, 2844.20, 2844.30*, 2844.50, 8401.30	2844.10, 2844.20, 2844.30*, 2844.50, 8401.30	2844.10, 2844.20, 2844.30*, 2844.50, 8401.30
92		**Other nuclear fuels**					
920	9200	Other nuclear fuels	33630*, 33690*	33630*, 33690*	2844.30*, 2844.40*	2844.30*, 2844.40*	2844.30*, 2844.40*
99		**Other fuels n.e.c.**					
990	9900	Other fuels n.e.c.					

Note: "Coal products" refer to the products derived from hard coal and brown coal; "Peat products" refer to products derived from peat; "Oil products" refer to products derived from the processing of conventional crude oil, natural gas liquids, other hydrocarbons, refinery feedstock, etc.

Descriptions and definitions of CPC and HS codes can be accessed on the websites of their custodians: United Nations Statistics Division and World Customs Organization (WCO), respectively.

[a] Since the definition of feedstocks is primarily based on intended use, giving an explicit CPC/HS link could be misleading. Feedstocks may cover a wider range of products, including naphthas (HS 2710.11) and pyrolysis gasoline (HS 2707.50), among others.

Chapter IV:
Generic Statistical Business Process Model

A. Introduction

4.1. The act of data collection, whether carried out through a survey or not, should not be considered in isolation but as part of a statistical process having various phases, from the identification of users' needs and the objectives of the generic statistical data collection, through the design and processing phases to the final stages of data dissemination and evaluation. When planning a survey, for example, it is important to have a clear understanding of the data needs, the data that are already available from other surveys or data sources, and the type of information that is intended to be disseminated and the means for doing so.

4.2. In anticipation of the description of different data collection methods for energy statistics in chapter V of this publication, it is important to review the different stages of the statistical data production process and how they relate to each other. This chapter presents an overview of the Generic Statistical Business Process Model (GSBPM), which was developed by the United Nations Economic Commission for Europe (ECE) to provide a standard framework for the business processes needed to produce official statistics. "The GSBPM can also be used for integrating data and metadata standards, as a template for process documentation, for harmonization of statistical computing infrastructures, and to provide a framework for process data quality assessment and improvement."[5]

4.3. Because of its flexible nature, the GSBPM can be used to describe the collection and compilation of energy statistics. While this chapter focuses to a large extent on the description of the GSBPM, a specific example in energy statistics is provided, together with examples of issues to consider in the various stages of the GSBPM.

5 Economic Commission for Europe (ECE), *Generic Statistical Business Process Model (GSBPM)* (Version 5.0, December 2013), available at www1.unece.org/stat/platform/display/GSBPM/GSBPM+v5.0.

B. Generic Statistical Business Process Model

4.4. The GSBPM consists of eight phases in the statistical business process, and within each phase a number of subprocesses are further identified. Figure 4.1 provides an overview of the eight phases and sub-processes of the GSBPM.

4.5. It should be noted that the GSBPM is not a rigid framework and should be applied with flexibility. The different phases of the GSBPM and their subprocesses do not have to be followed in the order described in the model and may be implemented in a different order and sometimes in iterative ways. The phases of the GSBPM are described below.

Figure 4.1:
Generic Statistical Business Process Model

Quality Management / Metadata Management							
Specify needs	**Design**	**Build**	**Collect**	**Process**	**Analyse**	**Disseminate**	**Evaluate**
1.1 Identify needs	2.1 Design outputs	3.1 Build collection instrument	4.1 Create frame and select sample	5.1 Integrate data	6.1 Prepare draft outputs	7.1 Update output systems	8.1 Gather evaluation inputs
1.2 Consult and confirm needs	2.2 Design variable descriptions	3.2 Build or enhance process components	4.2 Set up collection	5.2 Classify and code	6.2 Validate outputs	7.2 Produce dissemination products	8.2 Conduct evaluation
1.3 Establish output objectives	2.3 Design collection	3.3 Build or enhance dissemination components	4.3 Run collection	5.3 Review and validate	6.3 Interpret and explain outputs	7.3 Manage release of dissemination products	8.3 Agree on an action plan
1.4 Identify concepts	2.4 Design frame and sample	3.4 Configure workflows	4.4 Finalize collection	5.4 Edit and impute	6.4 Apply disclosure control	7.4 Promote dissemination products	
1.5 Check data availability	2.5 Design processing and analysis	3.5 Test production system		5.5 Derive new variables and units	6.5 Finalize outputs	7.5 Manage user support	
1.6 Prepare business case	2.6 Design production systems and workflow	3.6 Test statistical business process		5.6 Calculate weights			
		3.7 Finalize production system		5.7 Calculate aggregates			
				5.8 Finalize data files			

1. Specify needs phase

4.6. The specify needs phase of the GSBPM is used to ascertain the need for the statistical information; establish output objectives based on identified data needs; clarify concepts from the point of view of data users; check data availability; and prepare a business case in order to obtain approval to carry out a new or modified statistical business model.

4.7. Identifying users' needs is particularly important as it defines the purpose for which the information is being collected and the uses it could have in decision-making or for research. It is good practice to consult potential data users at this stage and to elicit their guidance in developing the objectives and scope of the survey programme. A clear statement of objectives, well-defined concepts and the level of data quality expected will allow the potential user to determine whether the information to be collected will serve their purposes. Next, users should be consulted to determine if the objectives identified align with their needs and whether there are any other concepts or content to be considered. This will aid in developing a robust survey programme and help ensure the relevance of the statistical data being produced. Users can be organizations, groups, agencies and individuals who are expected to use the information for policy, research or other purposes.

4.8. Before designing a new survey process, available data sources should be reviewed to determine whether information already exists to meet the users' needs and the conditions under which that information could be available. An assessment of possible alternatives would normally include research into potential administrative data sources and their methodologies to determine whether they would be suitable for use for statistical purposes. When existing sources have been assessed, a strategy for filling any remaining gaps in data requirement is prepared.

4.9. Once users have been consulted and their needs established, it is important to identify the conceptual, methodological and operational issues that should be addressed in developing an energy statistics programme. The preparation of a business case, which also includes an assessment of the costs and benefits of undertaking the data collection project, the quality expected by users and the expected delivery dates of the survey data are necessary to obtain approval for the overall data production process.

Important issues to consider in this phase

- When developing survey objectives and concepts, be sure to involve important users and other stakeholders.
- When developing a survey programme, try to make it as cost-effective as possible. It may be necessary to make a trade-off between cost and quality.
- Where explicit quality targets are established, ensure that they are included in the planning process.
- Elicit feedback from the group being surveyed (e.g. respondents in the energy sector, compilers of greenhouse gas inventories) to test concepts and questions, to manage the response burden and to increase respondent participation once the programme goes into the collection stage.

2. Design phase

4.10. The design phase of the GSBPM describes the development and design activities, and any associated practical research work needed to define the statistical outputs, concepts, methodologies, collection methods and operational processes. This phase also specifies all relevant metadata ready for use later in the statistical business process, as well as quality assurance procedures. For statistical outputs produced on a regular basis, this phase usually occurs for the first iteration, and whenever improvement actions are identified in the evaluate phase of a previous iteration. During that phase, an assessment of the consistency with international and national standards is explicitly made in order to reduce the length and cost of the design process, and to enhance the comparability and usability of outputs.

4.11. During this phase, a number of subprocesses are identified. The content and quality of the final outputs should be identified at this stage of the data collection process as subsequent work would depend on the choice of the statistics to be disseminated and their format. Disclosure control methods should also be considered and made explicit at this stage, as well as the correspondence to existing standards.

4.12. Once the outputs have been properly specified and described, the relevant statistical variables to be collected can be identified and the data collection method and instrument determined. The actual activities in this subprocess vary according to the type of collection instruments required, which can include computer assisted interviewing, paper questionnaires, electronic questionnaires, telephone surveys, administrative data interfaces and data integration techniques.

4.13. Designing a questionnaire should take into account the statistical requirements of data users and be based on the outputs defined in the first stage. When designing a questionnaire there are a number of considerations that are important for an effective data collection, such as the use of words and concepts that are easily understood by both respondents and questionnaire designers. It is usually the practice to consult with respondents during this stage about content and wording to ensure their understanding of what is to be reported. Metadata requirements should also be specified (e.g. through the creation of a data dictionary). Ideally, data quality guidelines should also be established for questionnaire design.

4.14. If the chosen data collection method is based on sampling, the design of the frame and sample is decided during this phase of the GSBPM. The target population, that is, the set of elements for which the estimated data are to represent, and the sampling frame should be clearly identified. Due to the various constraints that exist in executing a statistical programme, it is very likely that only a subset of this population will be used in compiling the statistics. This is the survey population or sample, the set of units that the constraints of the survey programme force us to narrow down to.

4.15. In the context of quality assurance, one desirable goal is to minimize the overcoverage or undercoverage between a target population and its sample. The frame should be kept as up-to-date as possible, as errors due to misclassification, duplication, erroneous conclusions or omissions will have a direct result on the quality of the estimates. Characteristics of the frame unit (e.g. contact information, address, size, identification) should also be of high quality as these aspects of the frame are essential when used for stratification, collection, follow-up with respondents, data processing, imputation, estimation, record linkage, and quality assessment and analysis.

4.16. Finally, the workflow from collection to dissemination should be identified to ensure a smooth data collection process. Designing a workflow for a survey process should take into account the quality dimensions of accuracy, timeliness, credibility and cost-effectiveness.

4.17. Statistical processing methodology defines how the collected data will be treated after collection has ended. Designing a robust system is essential to ensure the data quality of final estimates. The processing system should have clearly defined editing specifications for correcting errors in the collected data and to help populate variables that are only partially reported. This ensures data quality by providing consistent treatment of partially completed surveys and builds consistency checks into the initial phase of data capture and analysis.

4.18. The imputation method for survey non-response should also be clearly defined. It is necessary to determine the form of imputation that will be used for each scenario and to examine the implications of using a particular imputation method. Some imputation methods do not preserve the relationships between variables and may distort the underlying distributions. To ensure the highest possible quality, variances due to imputation (non-response error) should be taken into account when producing estimates.

4.19. Production systems are the means by which microdata are collected and processed into the final estimates. At this stage, a process flow that collects and processes the data through each stage of the survey process must be designed. Specifications should be designed that outline the needs of each stage of the survey process and allow the output file of one stage to be loaded into the next.

Important issues to consider in this phase

Frame design:

- Test possible frames at the design stage of a survey for their suitability and quality.
- When more than one frame is available, consider using a combination of them if this is likely to improve quality.
- Where possible, use the same frame or combination of frames when conducting multiple surveys with the same target population (e.g. for monthly and annual surveys).
- Be sure to actively monitor the frame quality on an ongoing basis by assessing its coverage of the desired target population and the accuracy of the description of the units' characteristics (e.g. proper industrial coding, births/deaths).

Questionnaire design:

- When designing the questionnaire, use a flow and wording that allow the respondent to reply to questions as accurately as possible. Use clear and concise language.
- In the introduction to the questionnaire, include a survey title and subject and clearly identify the purpose of the survey. Indicate the authority under which the survey is being taken.
- To reduce respondent errors made, provide instructions that are clear, concise and noticeable.

Imputation method design:

- Develop and test imputation methods before implementation.
- When designing a survey programme to produce energy statistics, take into account the relevant international standards in concepts and measurement procedures used for the collection and compilation of energy statistics as this will ensure their comparability.

3. Build phase

4.20. The build phase of the GSBPM includes activities related to the building and testing of the data collection instrument and dissemination components, and the configuration of the workflow. Building any survey programme requires careful coordination between its different elements. Designing a questionnaire should take into account the desired outputs and be coordinated with the building of the processing system. The processing system should reflect the desired workflow and each stage of the production process should flow smoothly into the next stage. To ensure data quality, there should be an emphasis on testing, once the questionnaire and the collection, processing and imputation applications have been created, to ensure that each system is integrated and functions properly. This phase includes the following activities:

- The collection instrument is generated or built based on the design specifications created in the previous phase. A collection may use one or more modes to receive the data (e.g. in-person or telephone interviews; paper, electronic or web questionnaires). Collection instruments may also be data extraction routines used to gather data from existing statistical or administrative data sets. This subprocess includes preparing and testing the contents and functioning of that instrument (e.g. testing the questions in a questionnaire). If possible, there should be a direct link between collection instruments and the statistical metadata system so that qualitative information from respondents can be captured during the collection phase. The connection between metadata and data at the point of capture can save work in later phases. Capturing the metrics of data collection (paradata) is also an important consideration in this subprocess.

- Building and testing process and dissemination components include, among other things, the detailed description of the subsequent phases of the GSBPM of "Process", "Analyse" and "Disseminate". Once built, the processing system should be tested for functionality to be sure that the collected data will be processed correctly throughout the entire process.
- Designing a process workflow involves configuring the flow of the data through the systems and transformations within the statistical business processes, from data collection through to the archiving of the final statistical outputs. Typically, this form of testing involves a small-scale data collection to test collection instruments, followed by processing and analysis of the collected data to ensure the statistical business process performs as expected.
- Qualitative testing of the questionnaire should be conducted with respondents in the target population. This testing can consist of focus groups or in-depth one-on-one interviews and can include cognitive testing. These methods are used to test question wording, sequencing and format. Cognitive testing involves assessing respondents' thought processes as they respond to the survey and determining whether they understand the questions and are providing accurate results. Qualitative testing may also be used to determine questionnaire content through the evaluation and exploration of key concepts.

4.21. Following the pilot, it may be necessary to go back to a previous step and make adjustments to instruments, systems or components.

4.22. Once testing is completed, the production systems can be finalized.

Important issues to consider in this phase

- All aspects of the newly designed collection and processing systems should be carefully tested.
- Consider two or more phases of questionnaire testing. This will allow testing of any revisions to the questionnaire during development.
- If surveys are conducted with personal interviews, it is good practice to hold debriefing sessions with interviewers after the questionnaire has been tested.
- Thorough testing reduces the risk of errors during the production process that could delay the processing of the survey estimates.
- Develop the quality measures that will be used in subsequent stages of the survey process.

4. Collect phase

4.23. Data collection is any process that acquires data to fulfil a survey objective. During this phase, data are acquired using different collection modes (including extractions from administrative and statistical registers and databases) and uploaded into the appropriate data environment. The collect phase includes the following subprocesses:

4.24. *Creation of frame and selection of sample.* For a survey sample to be properly selected, the frame must be maintained and be as up-to-date as possible to avoid imputations for non-existent establishments, to ensure all new establishments are included, to ensure the use of proper weights, and to avoid misclassification, etc. Once the frame has been established, the sample is selected based on the criteria determined in the design phase. Quality assurance, approval and maintenance of the frame and the selected sample are also undertaken at this stage. This includes the maintenance of underlying registers from which frames for several statistical business processes may be drawn. However, please note that the survey itself can be used to subsequently update and maintain the frame by using the information collected.

4.25. *Setting up of collection*. Setting up the collection stage ensures that people, processes and technology are ready to collect data in all modes as designed. This includes the planning and training necessary to conduct the survey, which are essential to data quality. Training interviewers to ask effective questions and to minimize poor or non-response will ensure that questionnaires are correctly filled out and response rates are maximized.

4.26. *Running of collection*. Once these stages have been completed, the collection process is ready to be implemented, using the different data collection instruments. This includes the initial contact with respondents and any subsequent follow-up or reminder actions. The process must record when and how respondents were contacted, and whether they responded. This activity also includes the management of the respondents involved in the current collection, ensuring that the relationship between the statistical organization and data providers remains positive in recording comments and addressing questions and complaints. For administrative data, this process is brief; the provider is either contacted to send the data, or sends it as scheduled. When the collection meets its targets (usually based on response rates), it is closed and a report on the collection is produced.

4.27. *Finalization of collection.* This subprocess includes loading the collected data and metadata into a suitable electronic environment for further processing. It may include manual or automatic data take-on. It may also include analysis of the associated collection process metadata to ensure that the collection activities met requirements. In cases where there is a physical collection instrument, such as a paper questionnaire which is not needed for further processing, this subprocess manages the archiving of that material.

5. Process phase

4.28. The process phase encompasses the cleaning of data records and their preparation for analysis. It is made up of subprocesses that check, correct and transform the collected data so that they can be analysed and disseminated as statistical output. These subprocesses may be repeated several times.

4.29. During this phase of the GSBPM, important steps in the compilation of the collected data take place, such as the integration of data from different sources. The input data can be from a mixture of external or internal data sources and a variety of collection modes, including extracts of administrative data.

4.30. The review and validation of data, as well as the editing and imputation of missing data, are parts of data compilation and a fundamental step in the finalization of the data files.

Important issues to consider in this phase

- When undertaking the editing process, ensure that all edits are applied consistently.
- Put in place a strategy for selective follow-up. It may be helpful to introduce a score function that concentrates resources on the important units or the most severe errors that require follow-up.
- It is good practice to limit the editing process for fixing errors up to a certain point as the contribution of further editing to error reduction diminishes. While editing is an essential process, it should have a cut-off point so as not to delay other activities.
- Be sure to identify, analyse and correct extreme values if necessary, as they will have a detrimental impact on the survey estimates.
- Guidelines in other sections of the survey process (e.g. editing, dissemination, etc.) are also applicable to administrative data.
- When possible, collaborate with the designers of administrative systems to help in building statistical requirements.
- Investigate and compare the quality of the administrative data with other sources if possible, the concepts, definitions, frequency, timeliness, and procedures underlying the collection and processing by the administrative organization.
- During imputation, evaluate the type of non-response and apply the appropriate imputation method.
- Be sure to carefully develop and test any imputation approach before it is introduced into the survey process.
- Ensure that imputed records are flagged and clearly identify the type of imputation applied to individual records. Evaluate the degree and effects of imputation.
- Consider the degree and impact of imputation during the analysis phase of the survey process.
- Proper estimation must conform to the sampling design. To that end, incorporate sampling weights in the estimation process as necessary.
- As the quality of the non-response adjustment factors used in the weighting process depends on model assumptions, validate the chosen model using diagnostics to minimize non-response bias.
- Outliers often lead to unreliable estimates when variables are continuous. When validating outliers, check for extreme values and outliers caused by excessive weights.

6. Analyse phase

4.31. The analyse phase of the GSBPM is where the statistical outputs are produced, examined in detail and made ready for dissemination. This process enables statistical analysts to understand the statistics produced.

4.32. There are a number of subprocesses defined within this phase which clearly describe the various steps for finalizing the output. They include the validation of the statistical outputs; the in-depth statistical analyses for better understanding of the statistics produced; and the application of disclosure control to ensure that the disseminated data and metadata do not breach confidentiality rules.

Important issues to consider in this phase

- Data should be validated prior to release. Use a variety of tools and processes to analyse the data and flag issues, then take corrective action.
- Determine whether the data are "fit for use". If any data quality issues emerge, assess the overall impact on the reliability of the data.
- Determine the extent of data quality evaluation required for a programme or a product. Factors to consider include the data uses and users, risk of errors, quality variation over the life of the survey programme, and the cost associated with data evaluation and its impact on data quality.
- At the minimum, information on coverage errors, response and imputation rates and sampling errors for key variables should be collected and analysed to assess the quality of the estimates.
- Take the necessary steps to ensure the confidentiality of the data to be released.

7. Disseminate phase

4.33. The disseminate phase of the GSBPM entails the release of the statistical output to the public. It includes all activities associated with assembling and releasing a range of static and dynamic products via a range of channels. These activities support customers to access and use the outputs released by the statistical organization. For statistical outputs produced regularly, this phase occurs in each iteration.

4.34. The dissemination consists not only of the activities naturally associated with data dissemination, such as data publishing, but also activities related to: (a) the promotion of the dissemination products – through, for example, marketing of the products to reach the widest possible audience; and (b) the management of the user support to ensure the prompt response to customers' queries, as well as their regular review, to provide an input to the overall quality management process as they can indicate new or changing user needs.

Important issues to consider in this phase

- Be sure to verify that the released data are consistent with the source data obtained. In the case of regrouped data or derived variables, this means that others should be able to reproduce the same results from the same source data.
- If data are being disseminated electronically, the output should be thoroughly tested to ensure that errors are not being introduced by the programming.
- Provide data quality indicators so that end users are able to assess the quality and limitations of the disseminated data.
- When releasing a dissemination product, be consistent in style and formatting.
- It is good practice to provide a contact person for each release of information for users who have questions or feedback related to the data release.

8. Evaluate phase

4.35. The evaluate phase of the GSBPM entails the regular evaluation of a specific instance of the statistical business process. All statistical activities should be reviewed periodically to adapt to evolving needs, to maintain the relevance and utility of the survey programme, to address quality issues or concerns, and to implement cost savings or strategies to reduce respondent burden, etc. The end of the evaluation phase is generally the agreement on an action plan to improve future iterations.

Important issues to consider in this phase

- Review ongoing statistical activities periodically and adapt to evolving needs to maintain the relevance and responsiveness of the survey programme.
- Elicit feedback from data users about the relevance and quality of the data being produced. Attempt to implement their suggestions if a significant impact on data quality is foreseeable.
- Prioritize the quality issues based on their impact on the aggregate data. It is unlikely that correcting all the issues related to quality is possible in the timeframe allotted, so it is good practice to focus on those issues that have the most impact.
- Use this opportunity to determine if there are processes in the survey programme that can be made more efficient.
- Review any user feedback pertaining to the accessibility of the data. Identify any possible improvements that can be implemented to improve the accessibility of the data to the end user.
- Identify any issues with the data or metadata that hinder accessibility or the end user's ability to analyse the data.
- If possible, compare survey practices with the best practices of other statistical organizations and endeavour to implement quality protocols that will increase the comparability of the data with those produced by other statistical organizations.
- However, note that implementing significant changes can create breaks in time series data. This can affect data comparability over time. Therefore, the need for changes must be balanced with the need for consistency.

Box 4.1:
Implementation of GSBPM for energy statistics in Azerbaijan

Processes for energy statistics in Azerbaijan are managed according to requirements indicated in the "Methodological Conception of Existing Process on Management System on Energy Statistics" and "Standards and Map of Management of Statistical Processes on Industry and Construction Statistics," both adopted by decree of the State Statistical Committee. These methodological materials completely describe the processes on energy statistics relating to quality management.

Both methodological materials were prepared based on the fourth version of the GSBPM. The "Methodological Conception of Existing Process on Management System on Energy Statistics" includes information on the management of energy statistics in 9 phases and 47 subphases, as in the GSBPM, including recording and calculation of 22 types of energy products.

Box 4.1 (continued)

GSBPM on crude oil and natural gas liquids in the State Statistical Committee of Azerbaijan

1.	Determine the need of crude oil statistics (Results: approved projects and plans)
1.1	Identification of need on crude oil statistics
1.1.1	Review decrees on crude oil adopted by higher and central executive government bodies
1.1.2	Review orders and decrees of the State Statistical Committee (SSC) regarding the recording of energy products
1.1.3	Review documents of Eurostat, IEA, Statistics Division and BP on crude oil
1.1.4	Organization of meetings with users on energy statistics and their analysis
1.2	Implementation of discussion with SOCAR and MIE on improvement of crude oil indicators and its coordination and confirmation
1.2.1	Organization of written and oral discussions with the stakeholders presenting reports
1.2.2	Analyse the existence of information
1.2.3	Presentation of draft questionnaires/methodological materials for discussion to the Methodological Council
1.2.4	Presentation of draft questionnaires, forms/methodology and materials for discussion to the Methodological Council
1.2.5	Presentation of draft questionnaires/materials for discussion to the Methodological Council
1.3	Create viable definitions of the variables
1.4	Identify concepts
1.4.1	Indicator of compliance required by the legislation
1.4.2	Compliance indicators with other branches
1.4.3	Ensuring international comparability of indicators
1.5	Determine internal and external sources
1.5.1	Identification of statistical units, with stakeholders
1.5.2	Study of administrative data sources
1.5.3	Methodology analysis of administrative sources
1.5.4	Check metadata availability
1.5.5	Prepare legal comments and letters, find solutions to collect information
1.6	Make a project work plan, determine the budget
1.6.1	Estimation of risks and benefits
1.6.2	Calculation and determination of costs
1.6.3	Approval of budget
2.	Discussion on filling in the report forms (6-extraction (oil, gas) 1-TİG (oil), 1-balance (oil)) on a PC (Results: All designs approved)
2.1	Design outputs and products
2.1.1	Define and set goals on product quality
2.1.2	Define information array
2.1.3	Develop instructions on filling in questionnaires, metadata
2.1.4	Determination of procedure for preservation of energy data and metadata
2.2	Design variable descriptions, surveys, registers
2.2.1	Preparation and harmonization of clarification
2.2.2	Specification of register
2.3	Definition of report coverage

Box 4.1 (continued)

2.3.1	Written request regarding coordination with SOCAR and MIE of the methodology on implementation of oil statistics
2.3.2	Presentation of draft methodology on implementation of oil statistics to Methodological Council for discussion
2.3.3	Confirm methodology on implementation of oil statistics
2.3.4	Execution of methodology
2.4	Determination of frame of survey and plan to conduct survey
2.4.1	Definition of survey frame
2.4.2	Definition of survey methodology
2.5	Design statistical processing methodology
2.5.1	Preparation of methodology of the project and the agreement
2.5.2	Discussion of the methodology in Scientific and Methodological Council
2.5.3	Approval of the methodology
2.6	Including of report forms reflecting crude oil data in statistical work programme
2.6.1	Breakdown of demands on questionnaire about energy statistics
2.6.2	Consideration of issues in the work plan
2.6.3	Review of work plan, timetables
2.6.4	Confirm statistical work programme
3.	Create process components (Results: System in production)
3.1	Build system
3.2	Build process components
3.2.1	Issues on technical assistance
3.2.2	Develop software for online collection of reports
3.3	Preparation of reports
3.4	Test production system
3.5	Test statistical business process
3.6	Finalize production system
3.6.1	Raise the level of staff's skill
3.6.2	Production system application
4.	Data collection (Results: Data ready for processing)
4.1	Including all economic subjects engaged in crude oil extraction, consumption, imports and exports in reporting register
4.2	Set up collection
4.2.1	Increase the level of staff
4.2.2	Provision of all statistical units engaged with extraction, consumption, imports and exports of crude oil with access to reports included in SSC webpage
4.2.3	Including of the software for online collection of reports 6-extraction (oil, gas) 1-TİG (oil), 1-balance (oil) on web page
4.3	Online collection of the reports
4.3.1	Compilation of work plan for LSO, MCC and branch sectors
4.3.2	Online gathering of reports 6-extraction (oil, gas) 1-TİG (oil), 1-balance (oil)
4.4	Finalize collection
5.	Classify and mode, micro edit, control, imputation, calculate weights (Results: Data ready for analysis)
5.1	Link data sources and statistical registers
5.2	Classification and coding
5.3	Review, validity check and editing

Box 4.1 (continued)

5.3.1	Online check of questionnaires
5.3.2	Detection of errors in report 6-extraction (oil, gas) 1-TİG (oil), 1-balance (oil)
5.4	Identification of potential errors and missing data
5.4.1	Identify elements of imputation
5.4.2	Conducting imputation
5.4.3	Holding notes on the imputation
5.5	Derive new variables and statistical units
5.6	Calculate weights
5.7	Calculate aggregates
5.8	Finalize data files
6.	Produce statistics, quality of data, preparation of data, finalize (Results: Analysis ready for dissemination)
6.1	Preparation of initial results based on 6-extraction (oil, gas) 1-TİG (oil), 1-balance (oil) reports
6.2	Validate outputs
6.2.1	Monitoring of the statistical feasibility of questionnaires
6.2.2	Detection of errors and discrepancies in primary data
6.3	Explanation of results based on output data
6.4	Preserve confidentiality of primary data
6.5	Prepare final result
7.	Dissemination of oil data (Results: Statistics disseminated)
7.1	Update of oil data
7.2	Oil data production
7.2.1	Prepare product components
7.2.2	Energy of Azerbaijan publication
7.2.3	Add annual data on oil production, internal consumption, import/export and changes of stocks of oil to web database
7.2.4	Preparation of the analytical report on production, consumption, import/export and change of stocks of the crude oil
7.3	Manage oil release products
7.3.1	Prepare working plan on analysis
7.3.2	Preparation of the data release calendar
7.3.3	Manage release product
7.4	Promote dissemination products
7.5	Manage questionnaires on oil, filling in of questionnaire
8.	Archive (Results: Statistics archived)
8.1	Determine archiving rules
8.2	Determine archiving rules
8.3	Preserve data and associated metadata
8.4	Dispose of data and associated metadata
9.	Evaluation of processes (Results: Evaluation)
9.1	Gather evaluation inputs
9.2	Conduct evaluation
9.3	Agree action plan

BP – British Petroleum office in Baku

IEA – International Energy Agency

JODI – Joint Oil Data Initiative

Box 4.1 (continued)

MCC – Main Computing Centre of SSC
MIE – Ministry of Energy of Azerbaijan
SOCAR – State Oil Company of the Republic of Azerbaijan
SSC – State Statistical Committee of the Republic of Azerbaijan

Chapter V: Data sources and data collection

A. Introduction

5.1. The aim of this chapter is to provide examples of typical data collection methods used for energy statistics. The choice of the specific data collection method depends on a number of factors, including, but not limited to, data availability and human and financial resources. The clear identification of what needs to be collected and what needs to be disseminated, and the review of the necessary methodological concepts (together with the consistency with international and national standards) are important elements that determine the choice of the data collection methods and instruments that best fit the purpose.

5.2. The actual choice of the data collection method is specific to the national situation. A review of the data already available in the country (whether or not in the same institution or in other institutions) is often a first step. If data are already available and can be used for energy statistics, there is little justification to embark on new data collection. It is important, however, to have information on how that data were collected (e.g. administrative data or existing statistical surveys) as it affects the appropriateness of its use for statistical purposes.

5.3. Classical examples include the use of administrative data from public or private registers or data collected through an existing survey designed for different purposes. A typical example of the latter is data on diesel consumption for agricultural purposes collected through an existing agricultural statistics survey. If this is the case, these data should be used to advantage, such as for resource savings or reduction of the response burden.

5.4. If no data are available in the country from other sources and a new data collection method has to be put in place (whether adding questions to an existing survey or designing a new survey), additional work needs to be carried out to identify the respondents and the needed human and financial resources.

5.5. This chapter focuses on the description of four types of data collection methods typically used in energy statistics: administrative data sourcing, statistical surveys (censuses or sample surveys), modelling and in situ measurements. It also provides examples of typical data collection methods used for specific types of energy statistics (e.g. trade, consumption, production, etc.).

B. Data sources and data collection

5.6. Four types of data sources and data collection methods are highlighted in this section: administrative data sourcing, statistical surveys, modelling and in situ measurements.

5.7. Administrative data refer to data derived from an administrative source, that is, by an “organizational unit responsible for implementing an administrative regulation (or group of regulations), for which the corresponding register of units and the transactions are viewed as a source of statistical data."[6]

[6] OECD Glossary of Statistical Terms, http://stats.oecd.org/glossary/.

5.8. Statistical surveys refer to both sample surveys and censuses. They investigate the characteristics of a given population by collecting data from a sample of that population (sample survey), or from the whole population (census), and estimate their characteristics through the systematic use of statistical methodology. It should be noted that a census involves a complete enumeration of the units in the population, which is usually costly; however, in energy statistics, it is often a feasible option in cases where the number of units in the population is low. This can be the case when looking at refineries in countries where there are only a few of them as it may be more accurate to collect the data from all of them rather than from a sample.

5.9. Modelling refers to the estimation of a variable/data item which cannot be measured directly but is estimated based on both measurable and observable data. A typical example is the estimation of the production of electricity through non-metered solar photovoltaic panels. Often, this quantity is estimated based on the number of solar photovoltaic panels in use and average weather conditions.

5.10. In situ measurement refers to techniques for collecting detailed consumption data based on a measuring device which, for example, can be installed at the point of final consumption.

5.11. Each of these data sources has advantages and disadvantages. Ideally, the objective is to collect data by the most efficient means possible. Therefore, if data are already available from an administrative source, or an existing survey (and they can be used for your purposes), then there would be no apparent need to embark on a new data collection process. If no such information of suitable quality is available, then adding questions to existing surveys or creating a new survey could be considered. Often, the production of energy is a licensed activity, so production data may be available within administrative systems, and existing trade data collection may provide suitable information on energy imports and exports.

5.12. There is a clear trade-off between cost and quality; the larger the sample size, the more costly the data collection, but the lower the standard error. Such decisions need to be based on the desired quality, subject to the available budgetary resources. If the population is fairly homogeneous then a smaller sample may suffice, whilst if it is heterogeneous then larger sample sizes will be required. In general, energy surveys are repeated so that the sample sizes can be amended over time, based on levels of variation among companies. In addition to the direct cost to the organization, the response burden of the survey on the businesses concerned should also be considered. This burden should be measured and minimized.

5.13. The frequency of data collection also needs to be considered. For example, if the same statistic is being measured on a monthly and annual basis, a smaller sample to produce the monthly data may be considered, with a larger sample at the end of the year. Also, it may be that only more restricted data, such as total usage of a fuel, are available on a monthly basis, while on an annual basis information may be sought by detailed industrial sectors.

5.14. For surveys on energy supply where, in most cases, the number of respondents is low and the competence of the respondents is high, censuses can be feasible solutions. In contrast, for final energy consumption, where the population of (potential) respondents is large and their knowledge of the topics is mostly relatively low, sample surveys conducted as interviews may be the preferable solution.

5.15. Table 5.1 gives an overview of the strengths and weaknesses of the discussed data sources.

Table 5.1:
Summary of advantages and disadvantages related to statistical techniques

	Advantages	Disadvantages
Statistical surveys	Comprehensive information on all fuels supplied and used Good data quality Can be used directly and as an input for model calculations Good response rates when surveys are covered by legislation	Resource intensive and costly Time consuming High survey burden If voluntary, response rates can be low Data validation required Reporting of non-metered fuels, often purchases not used
Administrative data	Low survey burden Greater number of records allows more detailed breakdowns Avoids duplication by making use of existing data No sample error	Dependency on third parties Definitions and information may not match statistical needs Often requires substantial effort to set up and there may be legal barriers to use
Modelling	Allows quantification of variables that cannot be directly measured or observed Saves resources (money and staff) Low survey burden Quick results Can be used to adapt or improve survey results Can be used to reduce survey frequency	Worse data quality compared to surveys No stand-alone methodology: cannot be calculated without input data Quality of results depends on accuracy of input data and the design of the model
In situ measurements	Detailed information on individual appliances, information on patterns of use of the equipment High quality of the results Input data for surveys and/or modelling	Invasive for respondents: difficulties in finding respondents willing to participate High burden in terms of time and human resources Costly, so often small samples, and less representative Constraints in monitoring equipment: limitation in the number of metering devices and monitoring incidences

5.16. It is often the case that different data sources are used to collect different types of data in energy statistics. Table 5.2 shows the typical data sources used in energy statistics for different groups of data items.

Table 5.2:
Suitable instruments and respondents depending on identified information needs

Information areas	Data collection methods	Data sources	Potential data observed
Energy supply: primary production of solid, liquid and gaseous energy products	Administrative data	Data owners	Coal production
			Crude oil production
	Census/sample survey	Entities in the mining industry (coal, oil, gas)	Natural gas production
		Entities in the forestry, agriculture or other related industries	Biofuel production
Energy supply: primary electricity, primary heat	Administrative data	Data owners	Electricity generation from hydro, wind, tide, etc.
			Geothermal heat
	Census/sample survey	Entities in the energy industries[a]	Heat from chemical processes
		Other energy producers	

Table 5.2: (continued)

Information areas	Data collection methods	Data sources	Potential data observed
Energy supply: solar electricity, solar heat and ambient heat	Administrative data	Data owners	Solar heat production (metered)
			PV electricity generation (metered)
	Census/sample survey	Entities in the energy industries	Electricity generated from solar heat (metered)
		Other energy producers	
	Modelling	Traders, installers	Sales of solar heat and PV panels, and heat pumps[b]
Imports/exports	Customs data	Customs/Ministry of Finance	Imports by country of origin and exports by country of destination
	Census/sample survey	Main importers/exporters	
Energy stocks (levels and flows)	Administrative data	Data owners	Stock levels and flows for coal, oil, natural gas
	Census/sample survey	Entities in the energy industries	Stock levels and flows for biofuels
		Other stock-keeping entities (mining and big industrial entities)	Water content of storage hydropower plants
International bunkers	Census	Traders	Sales to non-domestic ocean carriers and airlines
		Domestic ocean carriers and airlines	Use of fuels for international shipping and aviation
Energy transformation and secondary production (power plants, CHP plants, district heating plants, refineries)[c]	Administrative data	Data owners	Transformation input/losses
	Census/sample survey	Entities in the energy industries	Transformation output
		Other energy producers	
Energy industry own use	Census/sample survey	Entities in the energy industries	Own use of energy products in energy industries
		Other energy producers	
Final consumption	Business data from energy industries	Data owners, resellers/distributors	Final energy consumption (incl. transport)
	Sample surveys	Consumers[d]	Non-energy consumption
Energy prices	Census/sample survey	Suppliers/traders or consumers	Expenditures/costs/taxes
	Modelling		

[a] See IRES, table 5.1.

[b] As input for modelling the respective production.

[c] For full list see IRES, paragraph 5.70.

[d] See IRES, table 5.3

1. Administrative data sources[7]

5.17. Administrative data means all data sets publicly or privately owned that are collected for non-statistical purposes. Examples of administrative data are public registers, such as population or building registers, or private registers, such as billing registers of electricity supply companies. Within energy statistics, many core data points, such as coal production, crude oil exports or natural gas used for power generation, may come from administrative sources in this fashion, as these data are collected for taxation, customs or national security purposes, rather than as energy statistics in the first instance.

5.18. In many countries, the legal framework allows the institutions in charge of statistics to use such registers and apply legal authority to obtain the data, at least from public registers. If legislation allows the use of administrative data, it can be a first-choice data source for the following reasons:

- Data are already available and will be available in the future
- No additional resources for data collection are needed
- No additional burden for respondents

5.19. On the other hand, there are some restrictions which should be noted:

- The definitions used in the register normally do not match the definitions used in energy statistics
- Periodicity may be different from that needed
- Missing metadata

Box 5.1

Country example, United Kingdom: gas production data

The United Kingdom data collection system relies on a range of data sources, some operated by the Government itself, and others by private operators. The principal returns are:

Petroleum Production Reporting System: The main reporting system for data collection is the Petroleum Production Reporting System (PPRS), which is a census requiring all companies involved in extracting and processing oil and gas on the United Kingdom Continental Shelf (UKCS) to submit monthly details of their operations (both oil and gas) to the United Kingdom Government. PPRS is used to report flows from field level through to final input to the country's gas transmission systems.

National grid operational data: Returns from the National Grid, the owner and operator of the United Kingdom's National Transmission System (NTS).

Miscellaneous commercial returns: Various trade data are captured from operators trading gas, either via pipeline or to the United Kingdom's Liquefied Natural Gas (LNG) terminals. Some of these are in the public domain, while others are ad hoc collections run by the Government in cooperation with industry.

Legal basis

All operators of gas facilities on the UKCS, both offshore and onshore, require a licence to operate from the Government in the form of approval from the senior minister within the Department of Energy and Climate Change. As a condition of the license, operators are required to supply any data requested by the Secretary of State. The United Kingdom also rely on various enforcement powers, particularly those in the *Statistics of Trade Act 1947.*

[7] See also IRES, chapter VII, section D, subsection 2 – "Administrative data sources".

Box 5.1 (continued)

Production data

Gas production data are derived through the PPRS system. The PPRS began in 1975 when the United Kingdom began producing oil offshore. In 2000, it was revised and all UKCS operators were consulted with a view to creating a more efficient reporting system that would be beneficial to both the Department of Energy and Climate Change and the companies involved. This resulted in the current upstream PPRS 2000 data collection system. One of the principal shifts between the two systems was the move from the use of field data to terminal data. Prior to 2001, gas production was derived from field level data. Using PPRS 2000, gas production is currently derived from terminal data, supplemented with gas utilized on the oil and gas fields. This has two benefits: it considerably reduces processing time and improves data quality because feeds from the terminals are sales gas, which are more accurate than the pipeline gas reported at field level.

5.20. There are many potential data sources that can be used, if not for producing statistics then at least for providing error-checking information, or for matching with other variables in order to make certain breakdowns. This could include administrative sources such as:

- Buildings register or cadastral register (may contain information on location, age, type of dwelling and heating system)
- Renewable energy subsidy/grant register (often has information on the installed capacity and number of renewable energy installations in a given year)
- Database on energy performance of buildings (may contain detailed information on heating equipment, insulation and other energy efficiency characteristics) such as the one set up for European Union Directive 2002/91/EC on buildings
- Sales records of vendors or manufacturers of appliances or heating equipment
- Records of insulation installed from insulation fitters for homes
- Records of chimneys swept
- Sales records of central heating and heat pump vendors or manufacturers
- Registry of premiums paid for scrapping/recycling of appliances
- Address register
- Population register
- Business register
- Taxation register

5.21. The energy-related variables that, at the time of writing, are collected from administrative sources in European Union countries include:

- Annualized consumption of gas and electricity at household level
- Various aspects of energy performance of buildings (insulation, heating equipment, etc.)
- Housing stock characteristics (number of households, composition, dwelling types, etc.)

5.22. The use of administrative data sources often is combined with modelling to adapt the information obtained from them to the format, reference period, etc. needed.

Box 5.2:
Use of administrative data by Statistics Canada

Using administrative records presents a number of advantages to a statistical agency and to analysts. Administrative data, because they already exist, do not incur additional data collection costs, nor do they impose a further burden on respondents, provided the coverage and the conceptual framework of the administrative data are compatible with the target population or its collection is not terminated.

On the other hand, care must be exercised in using administrative data as there are a number of limitations to be aware of, including (i) the level or the lack of quality control over the data, (ii) the possibility of having missing items or missing records (an incomplete file), (iii) the difference in concepts, which might lead to bias problems, as well as coverage problems, (iv) the timeliness of the data (the collection of the data being out of the statistical agency's control, it is possible that due to external events, part or all of the data might not be received on time), and (v) the cost that comes with administrative data: for instance, computer systems are needed to clean and complete the data in order to make them useful.

The use of administrative records may raise concerns about the privacy of the information when linked to other sources of data, requiring agencies to provide all respondents with information on the confidentiality protection measures, the record linkage plans and the identity of the parties to any agreements to share the information provided by those respondents.

The use of administrative data may require the statistical agency to implement only a subset of the survey steps discussed in the other sections. Furthermore, any decisions related to the use of administrative data must be preceded by an assessment of such records in terms of their coverage, content, concepts and definitions, the quality assurance and control procedures put in place by the administrative programme to ensure their quality, the frequency of the data, the timeliness in receiving the data by the statistical agency and the stability of the programme over time. The cost of obtaining the administrative records is also a key factor in the decision to use such records.

Source: Statistics Canada (2009), *Statistics Canada Quality Guidelines*, Fifth Edition, Ottawa. Available at www.statcan.gc.ca/pub/12-539-x/2009001/administrative-administratives-eng.htm.

Box 5.3:
Country example, Netherlands: Use of client registers for energy statistics – consumption by dwellings and businesses

The basic idea behind the client registers project is to compile statistics on the consumption of electricity and gas by energy consumers for the overall energy balance. The goal is to identify the consumption of households, as well as businesses by branch of industry. The client registers, however, do not contain the information needed to determine if a connection is owned by a household or a business, let alone branch of industry.

In order to reach this goal, the client registers are matched with a register of dwellings and the business register, using the address, postal code, house number, and house number suffix as a matching key.[8] While matching with the business register is obviously required to identify businesses, this also assists in identifying false positives (about 2 per cent of the positive matches in our case) in the population of household connections. From the beginning of the project it was clear that matching the client registers with the business register was the biggest challenge. Only about 50 per cent of the business addresses initially matched one or more addresses of connections in the client registers, both for electricity and for gas.

8 Actually, the data are matched to several business registers to make up for coverage issues.

Box 5.3 (continued)

Matching the records of the client registers of electricity to the dwellings register is less problematic, although not straightforward. Not every match with the dwelling register is a dwelling, and some household connections do not match the dwelling register (about 3 per cent of the connections in our case). The biggest challenge is matching dwellings with unusual house number suffixes, followed by the challenges of identifying dwellings with block heating and businesses within dwellings. These challenges are explained in subsequent sections.

The dwellings register contains variables characterizing the dwellings, for example, dwelling type (whether it is an apartment, a villa, etc.), province, town, town district and neighbourhood. This enables the production of statistics on electricity and gas consumption per dwelling type and quarter/neighbourhood. Statistics Netherlands has been doing this since 2004. *The challenge of suffixes*

Matching addresses in the client files with those in other registers is not as trivial as was thought initially; the matching rate for the dwellings was only about 80 per cent for a number of reasons.

Firstly, the addresses in the client registers may be different (e.g. where the names clearly indicate the same business) from those in the business or dwellings register, or simply wrong (e.g. non-existing in the postal code register).

More importantly, the addresses in the (client) registers are far from standardized. The address variables that uniquely identify an address are the postal code, the house number and the suffix. However, in the (client) registers the suffix is used in many different ways. Common examples are A, B, C, 1, 2, 3, I, II, III, 1st fl., 2nd fl., 3rd fl., 015, 025, 035. Moreover, in the client registers unusual cases also occur like "next", "near", "across", "2 – 3", but also "garage", or "traction", and even worse. These suffixes can hardly be expected to match with more standardized suffixes used in a dwellings or business register.

To overcome these problems, two important steps were taken:

- Standardization of the suffix to the house number, for example, by dropping characters like "/", and changing all lower case characters into upper case, etc.
- Changing the suffix in the client registers into "blank" for those addresses that did not initially match and matching again (only for businesses).

After taking these steps the matching rate was eventually improved to over 96 per cent.

The challenge of block and district heating

For gas consumption, it was a challenge to identify dwellings using block or district heating; each about 4 per cent of all dwellings.

Block heating is the situation where a number of dwellings are heated by a shared heating system. Individual dwellings may either have their own gas connection, which is usually only used for cooking and may therefore have very low gas consumption, or have no gas connection at all. The problem with these dwellings is that the gas consumption of the households living there was underestimated, because gas used for heating was attributed to another connection that might not even be considered a dwelling. To solve the problem, information was used from the Department of Environmental and Spatial Statistics of Statistics Netherlands. They were able to provide information on the precise location of the blocks and the addresses within those blocks (knowing which dwellings make use of district heating is valuable, primarily because it can explain why a particular area has very low gas consumption, which is valuable information for local policymakers). Identifying dwellings making use of district heating was relatively simple, and was achieved by matching the addresses in the register with a list provided by the energy suppliers, including all addresses receiving district heating. However, it was not possible to determine the actual consumption of district heating.

Box 5.3 (continued)

The challenge of identifying businesses in dwellings

In about three per cent of the dwellings there is also a business. Statistics Netherlands excludes this consumption from household consumption and attributes it to businesses. Three separate steps are taken to identify these connections.

First, the customer name for each connection initially attributed to dwellings is scanned with text recognition algorithms. By looking for common combinations of letters or words usually designating a business (such as B.V. in the Netherlands), many connections can be attributed to business users.

Second, each address is matched to another register, one used by local governments for the purpose of building taxation. This register contains information on the value of every building, and crucially also a code designating the legal use/purpose of the building. The latter is used to identify business use. Third, each address is matched to a register with data on shops, bars and restaurants, enabling even more businesses to be identified.

In this way, it is possible to identify blocks of dwellings according to two criteria for each block:

(1) The block includes a number of individual dwellings with zero to very low gas consumption (i.e. gas consumption for cooking).

(2) The block has one address, supposed to be the address of the block heating system, with very high gas consumption.

Plausibility checks

During the studies, doubts arose whether matches with very low or very high electricity or gas consumption should be accepted as consumption by households. It happened, for example, that the lighting of a bus shelter with a low kWhbconsumption matched with an address in the dwellings register. In addition, a very high consumption of electricity could be due to a business started in an ordinary house. It was therefore assumed that electricity consumption needed to be in a certain range in order to be considered as household consumption. The top one per cent of consumption in each dwelling type category was labelled as possibly a business. Following advanced matching with business registers, it was finally determined whether it was a dwelling or a business. In this procedure, no lower limit is used. For natural gas, consumption lower than 400m^3 will mostly be excluded if found for several dwellings in a larger building, except for dwellings identified as users of district heating (these low consumptions were only excluded for the average gas consumption figures. However, they are included in the total national consumption by dwellings).

Businesses

Gas and electricity connections identified as belonging to businesses are grouped by address level and matched with business registers. The information from these registers in combination with text recognition algorithms on the name variable results in a 1-digit NACE classification of more than 95 per cent of the gas and electricity deliveries. Gas and electricity deliveries to the services sectors have been published since 2010; and the figures for other branches since 2014. Also, figures on a regional scale (province, municipality) will then be made available.

An extended version of this example is available on ESCM country practices website at https://unstats.un.org/unsd/energy/escm/country_examples/.

2. Surveys[9]

5.23. The term "survey" covers a wide range of activities, starting with small-scale data collection with only one respondent and very few data items, for example, the production of coke oven coke in a country with only one coke oven operator, or a comprehensive survey on final energy consumption of all households and businesses.

5.24. In designing a new survey, a number of considerations should be taken into account, such as formulating simple questions, identifying a clear structure for the questionnaire, providing incentives to potential respondents to join the survey voluntarily (even if there is a legal obligation), and possibly including data validation checks in the questionnaire (e.g. to ask for physical quantities and the corresponding expenditures), etc. While all these aspects are useful in the discussion of data sources for energy statistics, they are not addressed in this chapter; there is a vast range of literature on how to design a questionnaire and the reader is encouraged to review that literature.

5.25. For new surveys, special attention has to be given to the units used. Units used in billing that are familiar to the respondents are preferred, e.g. in the case of gasoline, litre instead of kilogram. In cases where different units are commonly used in the market (e.g. stere or kilogram in case of fuelwood), the respondent should be able to choose the unit, such as by ticking a respective box. The same method can be used to distinguish between gross and net expenditures for energy commodities, as demonstrated in the paper questionnaire used by respondents in Austria, and shown in the country example at the end of this chapter.

5.26. For cases in which surveys are necessary, two possibilities exist:

- Adding additional questions to an already existing survey
- Implementing a new survey

5.27. There are also some common rules which energy statisticians should recognize:

- Ask only the **must have** and eliminate the **nice to have.**
- Keep the questionnaires and questions **as simple as possible.**
- Ask only questions the respondents can answer.
- Give the respondents a **comprehensive explanation** of why and for what the data are needed.
- Be comprehensive **to fulfil all information needs**, e.g. to compile energy balances and energy accounts.
- Use **international classifications** as far as possible (ISIC, SIEC, PRODCOM, CPA, CN).
- Implement elements for **data validation**, e.g. ask for quantities and expenditures or revenues, and heated and cooled areas in m².

5.28. Supply surveys are normally fuel specific surveys, which means they focus on a specific fuel, such as electricity or on a closely linked fuel group characterized by a specific supplier (refineries for oil products or coal mines for coal).

9 See also IRES, chapter VII, section D, subsection 1 – "Statistical data sources", which provides a survey breakdown by respondents.

Box 5.4:
Country example, United States of America: Annual report of the origin of natural gas liquids production (EIA)

The collection of basic, verifiable information on the reserves and production of natural gas liquids (NGLs) at the national level in the United States is mandated by law and is required by the United States Department of Energy to assist in the formulation of national energy policies.

The data collected by the United States Energy Information Administration (EIA) include the annual volumes of natural gas received, and NGLs extracted at gas processing plants by areas of origin. It also includes the total gas shrinkage resulting from NGLs extracted and the annual volume of natural gas utilized as fuel at the gas processing plants.

Gas shrinkage volumes reported by natural gas processing plant operators are used with natural gas data collected on a "wet after lease separation" basis through the *Annual Survey of Domestic Oil and Gas Reserves* to estimate "dry" natural gas reserves and production volumes, regionally and nationally. The shrinkage data are also used, along with plant liquids production data and lease condensate data collected through the two surveys, to estimate regional and national gas liquids reserves and production volumes.

Respondents are natural gas processing plant operators. To facilitate the processing of data, the use of EIA forms is requested (either hard copies or Excel spreadsheets). Each operator is required to complete a separate form for each gas processing plant. The completed forms may be submitted by Secure File Transfer, email, fax or mail. pleted forms may be submitted by Secure File Transfer, email, fax or mail.

Disclosure limitation procedures are applied to the statistical data published from the survey information to ensure that the risk of disclosure of identifiable information is very small.

In the survey instructions available on the EIA website, several elements are clearly indicated:

The purpose of the data collection and its mandatory nature

Who are the respondents

The contents of the survey

The timing and the procedure to follow for data submission (including specific instructions for using the Secure File Transfer System)

The sanctions for those who fail to comply with the reporting requirements

Details on the disclosure of the information collected.

In addition, respondents can find practical guidelines on data standards (data are to be reported on a total operated basis; what units of measurement are to be used; and rules of rounding) and specific instructions on how to fill out each item of the form and a glossary with all the relevant definitions.

Adapted from http://www.eia.gov/survey/form/eia_64a/instructions.pdf.

5.29. There are at least three good reasons to ask suppliers for data, whenever it is legally possible and the information requested can be provided:

- The respondents are familiar with the topic
- Data are often available in business registers
- The small population makes surveys cheaper

5.30. One problem associated with supply surveys is the danger of double counting due to resellers. Therefore, questions on sales have to focus clearly on sales to final consumers.

5.31. If more detailed information (e.g. consumption by economic sectors or by purpose) is needed and suppliers – normally energy industries – cannot provide sufficient data, consumer surveys have to be conducted.

5.32. Consumption surveys typically focus on all energy carriers used by the specific target group of the survey focus (e.g. manufacturing industry or private households). Exceptions are detailed surveys on the consumption purposes and the consumption pattern of electricity in a specific sector. Such detailed data collection is restricted because it is extremely costly due to its complexity and the significant effort needed for respondents care.

5.33. Information on energy use by households may come from energy suppliers (oil, electricity distributors, etc.), or from household surveys. Results from the sample surveys could be grossed up by using population figures relating to households (if that exist) or the total number of persons. In some countries, household energy surveys are carried out as an additional survey to household budget studies. Households often have reliable information on energy costs, but not on their energy use in terms of physical units (e.g. kWh). It might therefore be necessary to request, in the survey, for receipts from energy suppliers to be displayed.

5.34. If sector-specific energy consumption surveys are planned, it is helpful to know the main end uses driving consumption and the most used fuels of the respective sector. These drivers can widely differ from sector to sector and region to region. For example, in a central European country, some 80 per cent of energy consumption in the manufacturing industry is used for production, while a similar share in the residential sector is used for space heating. In a country with tropical climate, the industry sector consumption may be similar but the main drivers for private households are completely different.

Box 5.5:

Country example, India: Energy sources for cooking in ruraland urban areas

The National Sample Survey Office (NSSO) conducts nationwide household consumer expenditure surveys at regular intervals as part of its "rounds", with each round normally being of a year's duration. The NSS surveys are conducted using household interviews from a random sample of households selected through a scientific design and covering practically the entire geographical area of the country. The household consumer expenditure survey (CES) is generally one of the main subjects of the NSS survey conducted at intervals of five years. The sixty-sixth round of the survey (July 2009–June 2010) was the eighth survey of this series.

One of the key indicators of household consumption is energy sources used for cooking by Indian households. Results show a significant difference between households living in rural and urban areas:

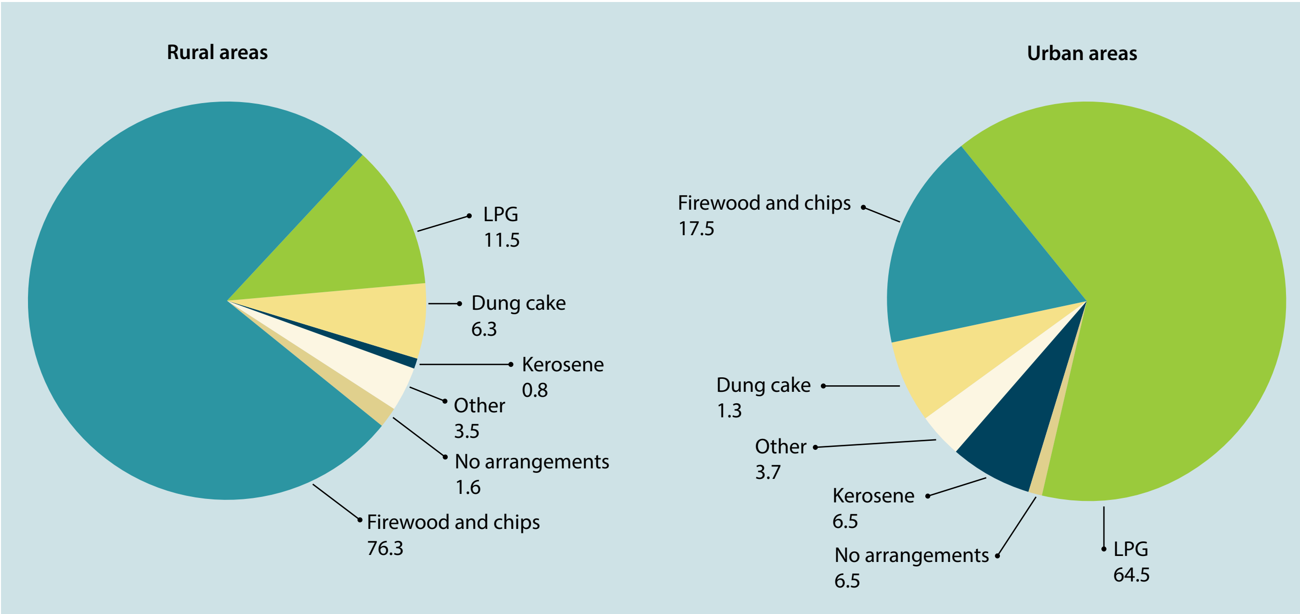
In rural India, in 2009–2010, firewood and chips were used as the principal energy source for cooking by more than three quarters (76.3 per cent) of households, LPG by 11.5 per cent and dung cake[10] by 6.3 per cent. About 1.6 per cent of rural households did not have any arrangement for cooking. The remaining households used other sources.

In urban India, LPG was used by 64.5 per cent of households, firewood and chips by 17.5 per cent, and kerosene by 6.5 per cent. As many as 6.5 per cent of urban households did not have any cooking arrangement, while dung cake was used by 1.3 per cent of households.

Figure 5.1 shows a graphical representation of the survey results.

10 Dung cake is included in the Standard International Energ-Product Classification under 513, "Animal waste".

Figure 5.1:
Energy sources used for cooking in rural and urban areas in India, 2009–2010



Source: Energy Sources of Indian Households for Cooking and Lighting, NSS 66th round, Ministry of Statistics and Programme Implementation

5.35. If results from both supply and consumption surveys are available, their comparison is a good instrument for data validation and verification. This normally is the case for all traded energy carriers and is of outstanding importance for compiling energy balances and energy accounts. Fossil fuels, such as oil, coal and natural gas, as well as electricity and traded heat, can be checked sufficiently in this way.

5.36. For energy carriers that are non-traded or traded on informal (grey) markets and used by private households (usually biomass, such as wood, bagasse or dung), or energy residuals from the production processes of manufacturing industries that can be used for energy purposes, information is available only from consumer surveys. Another disadvantage of these kinds of fuels is their heterogeneous composition and water content, which can result in widely differing energy contents. This means that the collection of data with sufficient quality for these fuels is much more difficult.

Box 5.6:
Food and Agriculture Organization – A Guide for Fuelwood Surveys

The Food and Agriculture Organization of the United Nations (FAO) published *A Guide for Woodfuel Surveys* in the frame of the EU-FAO Partnership Programme (2000–2002). The *Guide* provides reference material for formulating and solving problems that arise in reviewing, verifying, collecting, compiling, analysing, interpreting and presenting information on biofuel supply and demand. It also proposes a uniform methodological basis for obtaining data that are compatible, comparable and consistent at different levels and for diverse sectors of users, producers and suppliers of woodfuels.

For the most significant variables, their usefulness and data collection techniques are described in the FAO *Guide*, along with the analysis of the design of samples to acquire data on target variables, the different stages of survey design and the various ways of presenting the final survey data.

Box 5.6 (continued)

The main variables are presented under three major thematic headings: demand, supply and provision. The principal reasons for such grouping are that the methodologies for acquiring and processing information differ for each group and the information sources are also different.

However, four supplementary variables are also needed for expressing, in International System (SI) units, certain highly important variables, such as consumption, prices, stocks, productivity, and physical and economic flows. The accurate determination of these variables is of utmost importance and is needed in all woodfuel studies if comparable results are to be obtained. These variables, for which the *Guide* contains general rules and recommendations on how to collect the information at the local level, are:

Local units and their SI equivalents: Local units are the conventional or commonly used woodfuel units in a specific area that differ from SI units. In the case of woodfuel, local units can be head load, bicycle load, donkey load, back load, maze, billet, stick, cord, as well as the stere cubic metre. There is an enormous diversity of local units and the same term can often denote different quantities in different locations or regions.

Specific weight: Specific weight is the weight of the material contained within a set volume. For present purposes, specific weight needs to be expressed in dry weight and in SI units (kg/m^3).

Moisture content: Moisture content is the proportion of water physically in the woodfuels. Its determination is essential for calculating dry weight, which must be the basic unit of measurement for all woodfuels, and also for estimating its heating value "as burned", i.e., its effective energy content. **Heating value**: The heating value (HV), also named calorific value, is the amount of energy released by each unit of mass (or volume) burnt. It is expressed in kcal/g or Joules/g, and is determined in the laboratory using calorimeters. It is an essential variable for calculating the energy contained in woodfuel, and considerable time and effort are devoted to its determination. Although the calorific power of dry biomass is fairly constant, some important variations derive from the moisture content and, secondly, from the ash content.

As this *Manual* was being finalized, the FAO released a new guide, *National Socioeconomic Surveys in Forestry*, available at

www.fao.org/3/a-i6206e.pdf?utm_source=twitter&utm_medium=social+media&utm_campaign=faoforestry.

5.37. A detailed description of the problems associated with fuelwood and potential solutions are provided in the *Quality standard for statistics on wood fuel consumption of households*,[11] developed by the European Union's Working Group 2: Methodology of the Concerted Action on Renewable Energy Sources (CA-RES).

11 See https://circabc.europa.eu/sd/a/6adb0814-2638-4b87-a403-9dfc3b431fcc/Quality%2520standard%2520for%2520statistics%2520on%2520wood%2520fuel%2520consumption%2520of%2520households_CA-RES_2012.pdf.

Box 5.7:

Country example, Brazil: Integrating renewables and wastes in the energy balance

Compiling annual energy balances is a complex task that demands standardized data processing that is uniform over time, thus preserving statistical coherency. Compiling the Brazilian energy matrix, which, 40 per cent of which was, in 2013 made up of renewable sources and wastes,[12] requires the assessment of approximately 2,000 data points.

Source data for the energy balance are collected both systematically (from regulatory agencies, producers, distributors and industrial associations) for products subject

12 Resenha Energética Brasileira, Exercício de 2013 – Ministry of Mines and Energy.

Box 5.7 (continued)

to stronger market regulation and, in a more complementary way, for products more loosely regulated. Data for the former are already available in the desired format from their sources and do not need further processing (other than possible use of aggregation/disaggregation rules). Energy data for the latter are collected explicitly through specialized surveys, as well as implicitly from more general surveys, which might only provide rough indications on energy use. For complementary information, the explicitly collected energy data may need grossing up, interpolation or extrapolation. The use of expert statistical methods and models, such as the use of specific consumption indicators, correlation with other energy products, and correlation with the value added of the economic activity, may be necessary to estimate relevant energy statistics from more generic data.

Data obtained systematically cover about 99 per cent of the supply, transformation and sectoral consumption of crude oil, natural gas, oil products, uranium, coal, hydroelectricity, ethanol, biodiesel, biogas and wind electricity. The remaining one per cent refers to the disaggregation of sales to autoproducers of electricity into final consumption and transformation input, in which case the information is obtained by survey.

For data obtained complementarily, the most important survey is the annual autoproducer survey carried out by Empresa de Pesquisa Energética (EPE) (the Government agency that compiles the energy balance). The survey form includes questions on production and consumption of electricity and other sector-specific energy products. It covers 100 per cent of the producers of steel, pulp and paper, pelletization of iron ore, aluminium, alumina, copper, zinc, lead, iron alloys, petrochemicals and cement; 30 per cent of the producers of sugar and ethanol; and 40 per cent of the producers of charcoal-based pig iron. For sectors where coverage is incomplete, specific consumption indicators are applied to obtain total figures. As regards specifically renewables and wastes, the survey obtains information about the complete chain of energy products derived from sugar cane (bagasse, as well as ethanol and its feedstocks: cane juice and molasses); the consumption of charcoal and agricultural residues in cement production; consumption of charcoal in the production of pig iron; the consumption of fuelwood, black liquor and other residues in pulp and paper production; and the consumption of fuelwood and charcoal in the production of iron alloys. Data on the use of other, less relevant, renewable sources are also obtained for other industries, such as the chemical and textile industries.

Other occasional surveys and research have complemented the information necessary for the renewable part of the energy balance compilation:

(a) The Brazilian energy matrix project (1968–1970), which consisted of a generalized energy survey to assess all aspects of energy production and consumption in Brazil that allowed, together with expert information, the compiling of the first national energy matrix in 1970.

(b) National Household Surveys (PNAD) carried out by the Instituto Brasileiro de Geografia e Estatística (IBGE) (the Brazilian National Statistical Office), which, among others, collect information on predominant fuels for cooking (fuelwood, gas, charcoal) and carry a supplement on electrical appliances.

(c) The FAO/IBAMA[13] regional survey (1988–1992) on fuelwood, charcoal and LPG, covering the residential and agricultural sectors, as well as large consumers of renewables.

(d) The EPE national fuelwood survey (2011) covered the use of fuelwood and charcoal in rural areas nationally. The results corroborated the figures in the energy balance, showing that they were accurately measuring these quantities.

(e) The IBGE decennial agricultural censuses, which include questions on the consumption of fuelwood, charcoal and gas-diesel oils.

[13] IBAMA is the Brazilian Institute of Environment and Renewable Natural Resources, the administrative body of the Brazilian Ministry of the Environment

Box 5.7 (continued)

(f) The IBGE industrial censuses, which before 1990 used to include a question on the consumption of fuelwood.

(g) The IBGE annual family budget surveys (POF), which provide the number of households that consume fuelwood and charcoal.

(h) Other state-level surveys, which are carried out by state governments at irregular intervals and cover the consumption of energy from renewable sources.

These surveys provide good benchmarks for the years they are carried out. Data interpolation or extrapolation in the years the surveys are not conducted is done by using the correlation with other energy sources and products, as well as with the economic growth of the industry/sector in question. For example, the residential consumption of fuelwood and charcoal is obtained by using the correlation with LPG and natural gas. For agriculture and specific industries, consumption of fuelwood is estimated by using its correlation with economic growth, as well as other information derived from data obtained systematically. The fuelwood input into charcoal production is estimated from charcoal consumption by applying an average transformation efficiency inferred from information provided by the producers. Other less relevant energy flows are estimated by using less rigorous methods which, because of the small magnitude of the results, do not affect the quality of the general results of the energy balance. As an example of the kind of information brought about by the state-level surveys, an industrial survey carried out in 2012, in the State of Santa Catarina, indicated that the fuelwood consumption in the ceramic, food and other industries could be undervalued by as much as 30 per cent, according to preliminary analysis. That signals the need for revisiting such figures in the national energy balance, and could lead to revisions based on subsequent survey results in the future.

5.38. In gathering data on stocks, it is important to note that stocks may be held by producers, importers, governments, industries and other consumers. However, it may not be possible to collect stock data from all consumers. Since it is often difficult to measure stock level and stock changes held by consumers, stock changes should at least be collected from large producers, such as oil companies and refineries. Information on stocks for coal, coke and petroleum coke could also be collected from large industrial consumers of coal.

5.39. When surveying data on energy prices, note should be taken if prices have to be reported on a net or gross basis, and which price components have to be reported separately, if necessary (e.g. fuel specific taxes, recoverable taxes, network fees and metering fees). Furthermore, it makes sense to have a distinct focus, depending on the specific energy carriers (e.g. for electricity or natural gas on the respective supply companies), or groups of fuels by origin (e.g. for oil products, on refineries), or even use (e.g. for transport fuels, on filling station operators).

5.40. Last but not least, one should be aware that a survey has to be conducted at least three times to obtain reliable results. The first data collection obviously gives good results because no comparable information is available at that point. The second survey cycle – especially in the case of sample surveys – often delivers some contradictory results, and at least a third cycle is needed to resolve those problems. In the case of the household energy consumption survey in Austria (see country example below), four survey cycles were needed to establish a satisfactory data validation procedure.

Box 5.8:

Country example, Austria: Survey on energy consumption in manufacturing industries

No of registered vehicles	until 3,5 t		Annualy driven kilometers		No of unregistered vehicles (e.g. diggers, forklifts etc.)	
	above 3,5 t		Annualy driven kilometers			

Transport fuels	Unit	Amount	Value in Euro	gross [1]	net	Purposes [1]: Car	Purposes [1]: Others [2]
Gasoline	litre						
Diesel	litre						
LPG	kg [1] litre						
Natural gas	m³						

Other fuels used	Unit	Amount	Value in Euro	gross [1]	net	Purposes [1]: Space heating	Purposes [1]: Cooking, water heating	Purposes [1]: Others [2]
Electricity [3]	kWh							
Natural gas [3]	m³ [1] kWh							
District heating [3]	kWh							
Fuel oil and gas oil	kg [1] litre							
LPG	kg [1] litre							
Pellets, Wood briquettes	kg							
Wood chips, bark	stere [1] kg							
Others, please specify [4]	(Unit)							

Heat pumps	Installed capacity in kW		Space heating	Others [2]
Solar pannels	Surface area in m²		Space heating	Others [2]
PV pannels	Surface area in m²			

Overall area of the establishment (building areas only) in m²	of which heated during winter in m²	and airconditioned during summer respectively in m²	
Mainly used heating system [1]	Central heating system	Stove	Other
Age of the heating system [1]	below 5	5 until below 10	10 until below 20	20 and older

[1] Tick appropriate
[2] e.g. lightning, computing, process heat, power sets etc
[3] Please use your last annual bill; Value = Overall price including all fees and taxes
[4] e.g. hard coal, lignite coal briquettes, coke, fuel wood etc.; **please specify the unit!**

(a) Additional questions to existing surveys

5.41. This solution is a good option if the information needed is specific, with restricted data volume, the complexity is low (self-explaining questions), or a specific target group is to be surveyed, such as the residential sector.

5.42. It is cheaper than an exclusive survey and the respondent burden is normally lower. Typical widespread surveys that can be used to implement specific energy consumption related questions are household expenditure surveys or agricultural censuses.

Box 5.9:
Country example, Austria: Implementation of the voluntary household energy consumption survey as an integrated module in the obligatory labour force survey

The household energy consumption survey is an independent module with voluntary response appended to the labour force survey, which is mandatory every second year. This, in turn, has been collected since 2006 by computer-assisted telephone interview (CATI) only. The energy data, supplemented with basic data from the labour force survey, such as size of dwelling and building, construction period of the building, heating system and number of persons living in each household, are transmitted in electronic form upon completion of the survey.

The advantages of integrating the household energy consumption survey into the labour force survey are the higher response rate (50 to 60 per cent) and the lower survey costs compared to an independent household energy consumption survey.

(b) Implementing a new survey

Surveys on energy supply and transformation

5.43. Censuses are the best choice for surveys on energy supply and energy production because, on the one hand, the number of actors is normally manageable and, on the other hand, the actors often are not comparable, for example in the case of oil supply refineries with blenders, where grossing up procedures would produce the wrong results.

Energy consumption surveys

5.44. As mentioned above, sample surveys normally are the choice for surveying energy consumption because of the large number of potential respondents.

Box 5.10:
Country example, Azerbaijan: Household energy consumption survey

The State Statistical Committee of the Republic of Azerbaijan conducts sample surveys of 18,500 households for the purpose of studying the processes in the country's energy sector, reflecting real energy consumption, particularly in households.

The sample survey questionnaire consists of 38 questions on the quantity and value of electricity, natural gas, liquefied petroleum gases, wood (for heating purposes), wood charcoal, kerosene, diesel fuel and motor gasoline. Local boiler houses and other item consumed for centralized heating are also covered.

The survey covers proportionally all regions of the country based on quarterly random sampling. The data for the past three months are indicated separately in the questionnaire for each month.

The survey results are used for the compilation of the energy balance. The provided information on consumption was calculated based on the current survey and used for balance compilation.

Box 5.11:

Country example, Germany: panel survey – energy consumption of private households

In an ongoing effort to establish an energy panel in Germany, the most recent documented survey covers the energy consumption of approximately 7,000 households nationwide for the years 2009 and 2010, with earlier survey waves covering the years since 2003. Data for the years 2011 and 2012 were gathered at the outset of 2014, with additional improvements on the questionnaire and survey design currently in progress.

The participating household samples are drawn from the population of German-speaking residents aged 14 to 69, with the aim of assembling a representative panel.

Each participating household is equipped with an interface that projects the questions onto the respondent's television screen and facilitates completion of complex questionnaires through filter techniques and visual assistance. Respondents may answer the questions at a time of their choosing; they are not under time pressure and are not obligated to complete the questionnaire in a single session. The forsa tool, which immediately saves the collected information on a server, further allows for automatic consistency and validity checks during the data input by the participant. This ensures that the collected data is of high quality.

The survey gathers information about the energy consumption of various fuels by private households, as well as the corresponding expenditures. In addition, households also disclose the use of renewable energy, socioeconomic characteristics, and their ownership of electrical appliances.

With a participation rate of approximately 70 per cent, the often-encountered problem of low participation is not an issue. However, there exist two groups of survey participants for which Germany seeks to improve the number of observations: low-income households and households that predominantly use district heating.

3. Modelling

5.45. Besides using existing information, modelling provides another opportunity to reduce costs, especially with energy consumption surveys, either by lowering survey frequency or by reducing the extent and complexity of a data collection exercise by using additional sources to obtain all the required information.

5.46. Because of the complexity of energy consumption, responses have to be validated intensively and carefully. Such data validation/adjustments are normally based on the default values used to check the reliability of the responses. Therefore, the survey is the first step, and a model-based data validation is the necessary second step to get realistic consumption figures.

5.47. Other typical cases for modelling besides data validation are:

- *Consumption breakdown by purpose*, if a fuel is used for more than one purpose (e.g. electricity).
- *Fuel shares*, if more than one fuel is used for a single purpose (e.g. for space heating).

Box 5.12:

Country example, Austria: Model-based data validation procedure for the household energy consumption survey

The new approach for model-based data validation was implemented in 2004. The total reported energy consumption is checked against the modelled consumption for each individual household, which is calculated from household characteristics (floor space, number of people in household) by applying pre-set parameters for the individual types of use (space heating, water heating and cooking). If quantity-value pairs do not match, some quite complicated plausibility routines are employed in order to select the most probable quantity-value pairs.

The procedure described below was implemented because the survey results showed unreliable consumption patterns for several fuels and were inconsistent with other information available. The default values are derived from households, using a single fuel for the respective purposes (water heating, cooking) and from Internet research. For space heating, they are based on research at the Vienna University of Technology. This model is in continuous development and the default values are adapted to changing circumstances. The results are currently in a reliable range for all fuels and can be used, together with other survey results, to compile annual energy balances for Austria and its Laender (regions).

The data validation procedure is conducted in Microsoft Access-VBA programming. The processing is based on the original untreated dataset. For checking the plausibility of single datasets, the following set of tables with factors are used:

- Fuel-specific conversion factors with corresponding calorific values
- Fixed upper limits of fuel amounts and corresponding expenses
- Laender-specific prices for natural gas, non-interruptible and interruptible electricity
- Average prices for other fuels
- Average annual energy consumption for water heating and cooking
- Average annual energy consumption for space heating per m^2 depending on the dwelling´s age and size

1. Single-data level

At first, common plausibility checks and data validation in accordance with data correction are conducted at the lowest level – that means all data reported on the household level are checked:

- Using upper thresholds for the annual energy consumption of every fuel; unrealistic amounts and respective costs are estimated down.
- Checks on impossible fuel – purpose combinations and corrections if necessary (e.g. district heat, solar energy and heat pumps used for cooking).
- Item non-response in case of missing amount, or price: completion of pairs of varieties by use of average consumption amounts/prices. Item non-response in case of missing amount and price: calculation of pairs of varieties by use of average consumption depending on the purpose, the fuel and the household size.

2. Completion and creation of pairs of variables

Completion of pairs of variables in case of missing amounts or costs

Incomplete pairs of variables are completed by using average market prices as follows:

- Laender-specific prices for non-interruptible/interruptible electricity and natural gas are based on information provided by the energy suppliers (local players) and information derived from annual bills or further sources (e.g. Internet publications)

For other fuels such as coal or LPG, the average prices are calculated using the extreme value-adjusted median derived from completely reported pairs of varieties of the survey itself.

Box 5.12 (continued)

Creation of pairs of varieties in case of missing amounts and costs

In the case of information on the amount and costs of a fuel used in the household being unknown, the fuel amount is calculated by summing up assumed default amounts for the announced end-use categories as follows:

Calculation of the share for water heating and cooking

If a fuel is used for water heating/cooking and both the amount and costs are unknown, the household need for both is calculated as a proportion of the average annual consumption based on the household size. In case of the use of more than one fuel for water heating/cooking, the calculated overall amount is divided by the number of all fuels reported for this end-use category.

Calculation of the share for space heating

In contrast to the energy needs for water heating and cooking, the energy consumption for space heating does not depend on the number of persons but is highly related to the dwelling´s size and age, and the living space. The figures used are split into three construction periods and three classes of dwelling size.

While the number of floors per dwelling, as well as the size and the age of the building are provided for every household through the labour force survey, the figures for space heating needs are directly applicable to each single data record. Compared with the former method of using Laender-specific averages of fuel-specific energy needs for space heating, the new approach provides much more precise results.

In the case of two or more fuels used for space heating, the calculation of the overall share for space heating in the household is much more complex and will not be dealt with here.

After computing the single shares, the overall fuel consumption and corresponding costs are calculated in the following order: water heating, cooking and space heating. If a particular fuel is allocated to more than one end-use category, the individual shares are added up and the costs calculated by using the average market prices to arrive at the final cost.

3. Calculation of the annual final energy consumption (FEC)

The consumed amounts reported in various physical units are converted (standardized) to kWh by means of calorific values. Default values of the average annual consumption per end-use category (space heating, water heating, cooking) enable the allocation of consumption shares for fuels where the neither the amount nor the costs have not been reported. Furthermore, default values are used to calculate the assumed FEC (= average energy need for space heating + average energy need for water heating + average energy need for cooking) of a household.

5.48. Other (frequent) reasons for modelling are:

- Conversion of the surveyed period to calendar years
- Climate adjustment
- Extrapolation with additional (administrative) data
- Matching survey results
- Data validation
- Estimation of non-metered consumption or non-purchased fuels like solar and ambient heat
- Estimation of consumption of non-standardized biofuels, mainly fuelwood
- Calculation of useful heat

5.49. Depending on the purpose, different modelling parameters can be applied. Typical model assumptions used by a large number of countries are:

- Heating degree days are normally used in climate correction and extrapolation of heating shares of overall energy consumption. Presently, a common approach does not exist. In the box below, the 15°/18° approach used by Eurostat is provided as an example.

Box 5.13:
15°/18° Approach used by Eurostat

Heating degree day (HDD) = Sum of differences between a given indoor temperature (base tempeature = 18°C) and the daily average outdoor temperature, if the latter is lower than an assumed limit for heating of 15°C (15/18 model).

Heating degree total = Sum of HDD of a given period.

This 15/18 model means that days with an average temperature (in a 24-hour period) over 15.0°C are not taken into account, but days with an average up to and including 15.0°C are included, with their difference to 18.0°C.

The following examples for four days are displayed:

1st day: Ø temperature of -6.0°C: 24.0 points

2nd day: Ø temperature of +1.5°C: 16.5 points

3rd day: Ø temperature of +15.0°C: 3.0 points

4th day: Ø temperature of +15.1°C: 0 points

- Due to technological developments and the increasing penetration of air conditioning, cooling degree days are of growing importance.
- Default consumption quantities for space heating (by m² and depending on insulation measures, as well as building age), water heating (by person) and cooking (by family size) can be used to break down energy consumption by purpose.
- Assumptions on the typical (regional/national) fuelwood mixture or on its water content can be used for modelling consumption quantities and energy content of fuelwood.
- Penetration of energy saving appliances, the composition of heating systems, and typical consumption behaviour (depending on socioeconomic parameters) can be used for modelling energy efficiency indicators.

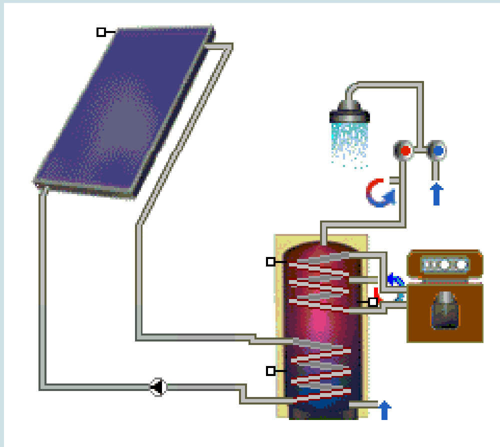
5.50. All these model assumptions can be applied on their own, combined with each other, or with additional national or regional assumptions. This opens a broad field for modelling and can help energy statistics compilers obtain realistic estimations when surveys reach their limits.

5.51. In the worst case scenario – if no information can be obtained from administrative sources, or by conducting surveys – modelling, based on experts' assumptions, may be the only possible way to make at least a "best estimate". A typical case is the International Energy Agency (IEA)/European Solar Thermal Industry Federation (ESTIF) methodology[14] for estimating solar heat production (see example in box 5.14). The advantage of using this estimation methodology is that the installed overall area or installed capacity can be used to calculate heat production. Both (or at least one) of these figures normally are well known by respondents. (As an example of how to collect the data needed, see the implemented questionnaire in the country example from Austria, in para. 5.40.)

[14] See www.iea-shc.org/common-calculation-method.

Box 5.14:
IEA/ESTIF methodology for estimating solar heat production

The only data needed are the installed collector area or collector thermal capacity, the type of system in the case of glazed collector types, and the average annual global radiation of the region.



The two possible calculation options are shown below:

A. Solar heat production as a function of the installed solar collector area:

Unglazed collectors:	$0.29 * H_0 * A_a$
Glazed collectors in domestic hot water (DHW) systems:	$0.44 * H_0 * A_a$
Glazed collectors in combi systems:	$0.33 * H_0 * A_a$

B. Solar heat production as a function of the installed collector nominal thermal power:

Unglazed collectors:	$0.42 * H_0 * P_{nom}$
Glazed collectors in DHW systems:	$0.63 * H_0 * P_{nom}$
Glazed collectors in combi-systems:	$0.47 * H_0 * P_{nom}$

where:

- H_0: Annual global solar irradiation in kWh/m^2.

  A_a: Collector aperture area (in m^2), but used in the calculation without a unit like a constant as shown in the example below.

  P_{nom}: Nominal thermal power output of collector in (kW), but used in the calculation without a unit like a constant as shown in the example below.

The following example uses the global radiation of Graz (1126 kWh/m^2) and typical solar panel parameters for a single-family house (glazed collector 5.9 m^2 installed panel area or 4.12 kW installed capacity) and gives the annual yield of solar heat (SHann) of this installation in kWh for DHW and combi systems.

DHW systems:

SHann = 0.44 * 1126 * 5.9 = 2923 (kWh/a)	using installed panel area
SHann = 0.63 * 1126 * 4.12= 2923 (kWh/a)	using installed capacity

Combi systems:

SHann = 0.33 * 1126 * 5.9 = 2192 (kWh/a)	using installed panel area
SHann = 0.47 * 1126 * 4.12 = 2192 (kWh/a)	using installed capacity

Box 5.15:
Country example, Brazil: Primary equivalent for direct residential use of solar heat

In 2005, the Brazilian Electricity Efficiency Programme (PROCEL) carried out a study which showed that 24 per cent of residential electricity use was for water heating. A decade later, in 2015, a study update saw this percentage lowered to 15 per cent, due mainly to the increased use of personal computers, television sets and other electronic appliances. This result was used by the Brazilian Ministry of Mines and Energy to help select a method, from among several, for assessing the direct use of solar heat in households.

Brazil has approximately 12 million m^2 of solar heat collectors, with 80 per cent being used for residential water heating and 20 per cent for swimming pool heating and other end uses. Considering 3 m^2 of collectors per household, 3.2 million households (4.7 per cent of the total)[15] are estimated to be using solar heating.

The Ministry of Mines and Energy considered two methods for assessing direct residential use of solar heat: (1) assessing the avoided electricity consumption (gauging the consumption side); and (2) applying solar collector suppliers' technical specifications to gauge production (comparable to the IEA/ESTIF methodology).

The first method employs the figure of 100 kWh of annual avoided electricity consumption per square metre of solar heat collectors, based on empirical studies carried out by the Ministry of Mines and Energy. The second method assumes that 500 kWh are generated in a year by each square metre of solar heat collectors, based on a range of between 420 and 750 kWh/m^2 provided by the manufacturers/suppliers. The table below compares the two methods using statistics for the year 2015 (yearly data).

Parameters	Unit	Method 1: Using avoided electricity consumption	Method 2: Using suppliers' specifications
Total solar collector area	000s m^2	12 000	12 000
Total residential solar collector area (80%)	000s m^2	9 600	9 600
Solar collector area per household	m^2	3	3
Households with solar collectors	000s	3 200	3 200
Total number of households	000s	68 000	68 000
Average electricity consumption per household	kWh	1 930	1 930
Residential electricity consumption used for heating	%	15	15
Households with solar collectors			
Households with solar collectors (% of total)	%	4.7	4.7
Direct use of solar heat per collector area	kWh/m^2	**100**	**500**
Direct use of solar heat per household with solar collector	kWh	**300**	**1 500**
Solar heat compared with residential electricity consumption	%	**15.5**	**77.7**
Entire country			
Direct residential use of solar heat	GWh	**960**	**4 800**
Domestic electricity consumption	GWh	161 267	131 267
Electricity consumption for residential heating (15%)	GWh	19 690	19 690
Solar heat as compared to electric heat	%	**4.9**	**24.4**

15 Recent studies (July 2016) and assessments by PROCEL indicated that around 5 per cent of households made use of solar heating.

Box 5.15 (continued)

The first method (assessing the avoided electricity consumption) estimates solar heat production/use as 300 annual kWh per household with solar collectors, which compares as 15.5 per cent of the annual electricity consumption of a typical Brazilian household. This figure is very close to the 15 per cent of residential electricity consumption being used for heating purposes found by the above-mentioned 2015 study by PROCEL. As for the whole country, the first method estimates solar heat residential use, comparing as 4.9 per cent of the country's total residential electricity consumed for heating purposes, which is well in line with the 4.7 per cent of households estimated as possessing solar collectors. By comparison, the second method (employing the suppliers' and/or manufacturers' specifications) yields percentages five times higher, at 77.7 per cent and 24.4 per cent, respectively, which seem unrealistic.

For these reasons, to estimate direct residential use of solar heat, Brazil's Ministry of Mines and Energy chose the method of assessing avoided electricity consumption, rather than relying on manufacturers' or suppliers' specifications.

4. In situ measurement

5.52. This data collection methodology can be expensive but, in the case of grid connected energy carriers like electricity, natural gas and district heating/cooling, it is often the only possibility for collecting data on consumption by purpose. Because of its high costs the sample is normally small.

Box 5.16:
Country example, Austria

In Austria, the measurement action is part of the "Survey on electricity and natural gas consumption of households by purpose". In contrast to other countries, the measurement is done by the households themselves.

The following general information was provided, together with the meter to the households, and was generally well understood.

Electricity cost meter

The electricity consumption has to be metered for at least one hour (e.g. for TV, computers), at least 24 hours (e.g. for cooler, freezer) or by load (e.g. washing machines, dishwashers). For some devices, it is necessary to note the power (e.g. vacuum cleaner) additionally. Please follow the requirements mentioned in the questionnaire for the respective appliances.

The advantage of this methodology is that it is inexpensive (approx. €10 per meter) and although the accuracy of the meter itself is limited and most respondents were inexperienced in handling such meters, the quality of the results was good. In 2008 and 2012, 254 and 264 households, respectively, joined the action.

5. Country cases – multipurpose approaches

5.53. This section lists some country examples to show how different approaches can be combined to obtain all information needed with minimum resources.

Box 5.17:
Country example, Austria: Modelling of electricity consumption in private households according to type of usage

This project developed a method for linking data records relating to electricity consumption in private households from different surveys by means of statistical matching.

In addition to information about the heating system, overall electricity consumption, electricity consumption for space heating, water heating and cooking, the model also included socioeconomic criteria (household size and legal relationship to the dwelling), property-related criteria (age of property, living area, number of dwellings in the property) and regional criteria (urban versus rural).

Methodology

The basis for the method is the following two surveys:

Household Energy Consumption: a voluntary survey with a gross sample size of approximately 19,000 households and a response rate of over 60 per cent for the years 2004, 2006, 2008 and 2010.

Electricity and Gas Journal: a voluntary survey with a gross sample size of approximately 500 households and a response rate of almost 51 per cent that was first conducted for 2008, and is to be repeated in future at intervals of several years.

Despite the relatively small sample size (254 responding households), the data quality of the Electricity and Gas Journal is good. The average annual power consumption of 4,417 kWh per household shows a good correspondence with the figure of 4,533 kWh for annual power consumption from the Household Energy Use survey (12,399 responding households).

As a first step, the information from the Electricity and Gas Journal survey was linked at the individual data level with the results of the Household Energy Consumption survey, using the 11 variables displayed below, which were present in all data records:

- V1: Household size (5 categories)
- V2: Number of dwellings in the property (5 categories)
- V3: Federal province (9 categories)
- V4: Age of property (8 categories)
- V5: Legal relationship to the dwelling (6 categories)
- V6: Use of solar heating (2 categories)
- V7: Primary heating system (5 categories)
- V8: Space heating with electricity (2 categories)
- V9: Water heating with electricity (2 categories)
- V10: Living area (m^2)
- V11: Overall power consumption (kWh/a)

The quantitative variables V10 and V11 were standardized for the matching in all data records, i.e. they were changed so that average=0 and variance=1. For each data record in the Household Energy Consumption Survey, a match was sought from the data record of the Electricity and Gas Journal that was closest to it. Where there was more than one data point with the same distance, one data point was randomly selected. In terms of the distance function, the variables V1 to V9 were given the value 0 or 1, respectively, whether

Box 5.17 (continued)

or not they were a match. In the case of V10 and V11, the absolute difference between the standardized values of V10 and V11 was included in the distance function. This means that the 11 variables were included in the distance function with the same weighting.

distance function, the variables V1 to V9 were given the value 0 or 1, respectively, whether or not they were a match. In the case of V10 and V11, the absolute difference between the standardized values of V10 and V11 was included in the distance function. This means that the 11 variables were included in the distance function with the same weighting. The linked data records were then recalculated in accordance with the annual series of overall final consumption of electrical energy for private households in Austria up to 2003, and a forecast was generated for the year 2011, from which time series were generated for the detailed uses of electricity in private households.

In the concluding processing phase, an analysis was performed on the generated time series, with thermal and detailed non-thermal power consumption categories for 2004, 2006, 2008 and 2010, according to socio-economic and regional parameters at the household level.

Box 5.18

Country example, United Kingdom: National Energy Efficiency Data-Framework (NEED)

The National Energy Efficiency Data-Framework (NEED) was set up by the United Kingdom Department of Energy and Climate Change (DECC) to provide a better understanding of energy use and energy efficiency in domestic and non-domestic buildings in Great Britain.

NEED is a framework for combining data from existing sources (administrative and commercial) to provide insights into how energy is used and the impact of installing energy efficiency measures. The address information in each dataset is used to assign a unique property reference number (UPRN) to each record. Data from different sources can then be matched to each other via the UPRN (see figure below). The principle is the same for both the domestic and non-domestic sectors, though different data sources are used.

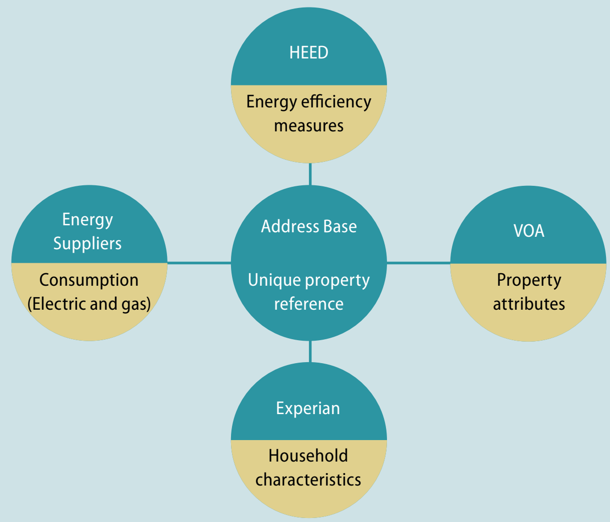
Four key data sources, shown in the figure below, have been used to analyse domestic energy. In addition to these four main property level data sources, a number of other indicators have been assigned to the property based on its geographic location (for example, the index of multiple deprivation assigned based on the Lower Layer Super Output Area (LSOA)).[16]

[16] Lower layer super output areas (LSOAs) are part of a hierarchical geography designed to improve the reporting of small area statistics in England and Wales. There are between 400 and 1,200 households in each LSOA. See www.ons.gov.uk/ons/guide-method/geography/beginner-s-guide/census/super-output-areas--soas-/index.html

Box 5.18 (continued)

Structure of domestic NEED

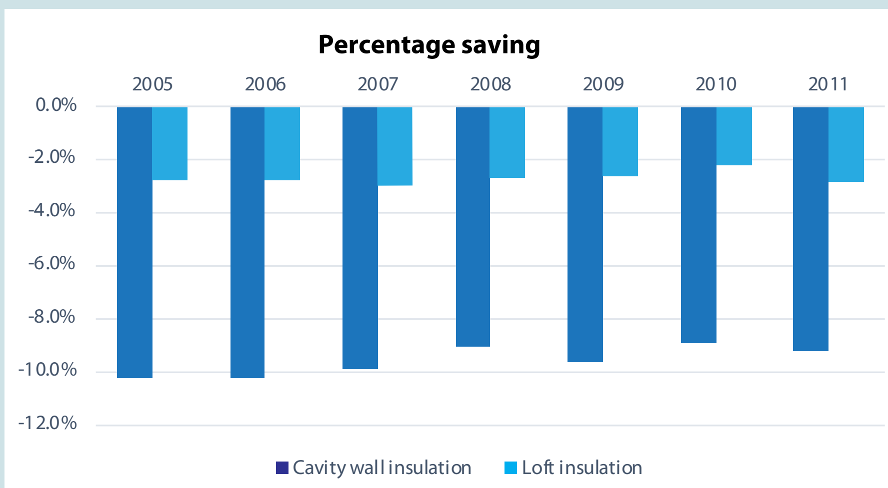
NEED has sup-



ported a number of DECC policies, with important outcomes (for example, by estimating the reduction in consumption for households installing energy efficiency measures).

Summary of observed savings for energy efficiency measures (median, not weighted)

Furthermore, the publication of an anonymized version of domestic NEED in 2014 has further increased the utility of the data by facilitating analysis of record level data by external researchers.



More information on NEED, including outputs, is available at www.gov.uk/government/collections/national-energy-efficiency-data-need-framework.

Chapter VI:
Commodity and energy balances

A. Introduction and purpose

1. Introduction

6.1. Energy statistics expressed in natural units, such as metric tons for oil, GWh for electricity, or terajoule or cubic metres for natural gas, can be presented in the form of commodity balances which show the supply and use of energy products. A commodity balance provides a check on the completeness of the data and is a simple means for assembling the main statistics of each energy product so that key data can be easily obtained. However, because fuels are mainly bought for their heat-raising properties and can be converted into different fuel products, it is also helpful to present the supply and use data in common energy units, such as terajoule or ton of oil equivalent (toe). That makes it possible to sum across the various products to obtain a value for total energy and to compare the relative contributions of the energy products against each other. The format adopted is termed the energy balance.

6.2. An energy balance is a complete framework for showing the production, trade, transformation and consumption of all energy products in the national territory. It allows users to see transformation efficiencies and the relative importance of the different fuel supplies in their contribution to the overall energy supply. The statistician also uses the energy balance as a high-level check on data accuracy as apparent energy gains in conversion processes or large losses can indicate data problems.

6.3. This chapter supplements the *International Recommendations for Energy Statistics* (IRES), chapter VIII, on energy balances and commodity balances. It presents practical guidelines on how to compile commodity balances and use them to compile energy balances. It is difficult to give guidance on energy balances without mentioning some main principles and definitions that are also described in IRES. However, as a rule, reference is made only to those IRES principles and definitions, instead of repeating the complete text, to avoid duplication to the extent possible. However, some repetitions have been included for practical reasons and to help the reader.

6.4. The chapter is aimed at countries that already have some experience with balances, as well as those that are in the initial phase of this work. In particular, it will describe how to use the data items presented in chapter VI of IRES on statistical units and data items. The chapter also discusses some of the options available to countries and international organizations on how to present their data.

6.5. The chapter is divided into four main parts. Section A presents a general discussion on the importance of energy balances and their main uses; Section B provides general information on the scope, frequency and level of detail of commodity balances and energy balances; sections C and D describe how to compile a commodity balance and how to convert it to an energy balance; and section E provides country-specific examples. Throughout the chapter, information is given on various principles that are needed in order to actually calculate an energy balance, and on estimating certain required data. Data checks on the commodity balance and the energy balance are also discussed.

2. Importance of energy balances

6.6. When compiled and published on a regular basis, an energy balance gives essential information for analysis and energy policy purposes, and for greenhouse gas emission calculations. Energy balances are also a starting point for the construction of various energy indicators (e.g. consumption per capita) and provide the basis for monitoring energy efficiency (although additional detail is required for this purpose).

6.7. Energy balances can also be used to monitor the progress of energy policy targets, such as increasing the share of renewable energy, reducing the growth of energy consumption or reducing dependency on energy imports. Policymakers are often interested in the overall growth (or decline) of a territory's total energy supply to see how much energy is being used. This figure is often combined with gross domestic product (GDP) or population data to get an energy intensity indicator. Comparing energy supply with a territory's production figure will also give a relative indication of energy security, as this shows whether a country is a net exporter or importer. Moreover, if a single product (either domestically produced or imported) accounts for a large share of total energy supply, this can also flag energy security concerns.

6.8. Moving on from basic indicators, the consumption data, broken down by economic sector in the energy balances, at least allows for each industry's total use of energy. But to really measure energy efficiency, these data can be combined with other socioeconomic data, such as number of vehicles and the size of housing stock.

6.9. Energy balances are also essential for greenhouse gas inventories. The Intergovernmental Panel on Climate Change (IPCC) guidelines state that energy balances, rather than commodity balances, should be used because data in a common energy unit are likely to vary less than data in physical terms (one ton of sub-bituminous coal can contain a wide range of carbon values, while one terajoule of sub-bituminous coal will have a carbon content that is much less variable). The necessity for energy balances not to double count production allows for everything in greenhouse gas data to be counted only once also.

6.10. In addition, energy balances may serve as input to physical energy flow account tables (see the *System of Environmental-Economic Accounting for Energy* (SEEA-Energy) publication), or as a basis for verifying value figures for energy in the national account framework (energy data with definitions comparable to national accounts).

6.11. Energy balances can be set up in a variety of ways. The level of detail and layout will depend on the specific objectives to be met. This chapter provides several examples of different layouts for energy balances. However, problems may arise when different principles and definitions are applied across countries, institutions and organizations. The user of such energy balances has to be cautious when comparing figures using the different layouts. This chapter aims to describe and explain the methodology for setting up energy balances as recommended in IRES, although different user needs and layout preferences are taken into account. For these reasons, when publishing an energy balance, it is always crucial to mention which methods and layout principles have been used, so that the data user understands how each figure was calculated.

B. General information pertinent to both commodity balances and energy balances

6.12. Complete energy balances are based on a variety of data sources. The data from those sources are usually collected in the energy products' natural unit, such as metric tons or litres for oil products, GWh for electricity, etc. It is recommended to start by compiling the commodity balance for each energy product, or group of products, in its natural unit, before converting it to an energy balance in a common unit. Commodity balances are valuable accounting tools used to compile and check the national energy statistics of a country. Once compiled, commodity balances provide the building blocks for calculating an energy balance.

1. Scope of commodity balances and energy balances

6.13. A commodity balance and an energy balance follow the territory principle. This means that the balance follows the flow of energy within the country. The figures include energy produced, traded and used within the country, regardless of the user's residence status or the nationality of the enterprise producing, transforming and selling the energy product. This principle is described in several places within IRES. This implies different figures between the energy balances and energy accounts (see SEEA-Energy), especially for international shipping and aviation.

6.14. The scope of the balances in terms of products is defined by the scope of the Standard International Energy Product Classification (SIEC). Therefore, all energy products of fossil origin listed in SIEC (except peat) should be accounted for in the balances regardless of whether or not they are used for energy purposes. This means that bitumen used in asphalt, lubricants in engines, or liquified petroleum gas (LPG) used in the production of plastics or chemicals, should be included, so that the inputs and outputs from refineries (or other transformation activities) can be properly accounted for, even when the final product falls outside the energy sector.

6.15. Data on waste, biofuels and peat used for energy purposes are also included in the scope of the balances. Thus, their inclusion in total energy production depends on their use; that is, it is derived from use-side information (i.e. waste, biofuels and peat are covered in balances only to the extent they are used for energy production and not when used for other purposes); a corollary to this is that the non-energy use for these products is always zero by definition.

6.16. As described in IRES, the energy balance does not include passive energy, such as heat gains in buildings and solar energy falling on land, energy resources and reserves, or waste, biomass and peat not used for energy purposes. An energy balance should also include energy from nuclear fuels. The SIEC lists "Uranium and plutonium", and "Other nuclear fuels" as energy products in section 9. However, in an energy balance, nuclear heat is normally the first form of energy accounted for (see section D, subsection 4 of this chapter).

6.17. The SIEC gives a hierarchical classification of energy products to be included in the balance and should be used as a reference to determine which products to include.

6.18. However, the energy market is fast evolving and statisticians should constantly consider whether new technologies or energy sources should be added to the energy balance, along with appropriate revisions of the SIEC. For example, until the 1960s, almost no electricity was produced from nuclear energy; more recently wind and solar energy have started to develop rapidly; biofuels have been quickly increasing in relevance; and in the near future, we might see a rapid development of hydrogen and fuel cells. As a consequence, there is an obvious need for statistics and statisticians to follow, if not to anticipate, the rapid evolution of the energy market.

6.19. There are also some minor products, not mentioned in the SIEC, that are explicitly included by a few countries only in their energy balance, such as energy from heat pumps and district cooling. This is further described in section D, subsection 12 of this chapter.

2. Level of detail

6.20. The main parts of the balances have been largely described in IRES. Energy products are described in detail in chapter III of IRES on SIEC, and the flows in chapter V on energy flows. The format and principles of commodity and energy balances are described in chapter VIII.

6.21. In general, the level of detail that appears in the disseminated commodity and energy balances (for the energy products and the energy flows) depends on the energy situation and user needs in each country, data availability and confidentiality.

6.22. Commodity balances should be constructed at the national level for every energy product in use. In practice, minor commodities are often aggregated for collection purposes. Care should be taken, however, when aggregating products; combining products too quickly in the process may inhibit making good estimations and carrying out accurate checks at a later stage.

6.23. Another reason for combining products may be to protect the disclosure of individual information (confidentiality). This usually takes place at a later stage of the compilation of the balances. While detailed information for each energy product should be kept in a database for working and analytical purposes, the final layout and presentation of the balance may require aggregating some products and/or flows to avoid the release of confidential information to the general public.

6.24. The SIEC gives a list of about 80 energy products. Within an individual country, fewer energy products may be relevant, but it may still be necessary to aggregate some products before publishing data in general. The same holds true for the energy balance as far as the calculation is performed on disaggregated data. Examples of different levels of aggregation used in countries are provided in this chapter.

3. Frequency of balances

6.25. Balances are usually prepared and presented on an annual basis, and IRES recommends that countries compile and disseminate them annually, at a minimum, because detailed data sources are often available on an annual basis only.

6.26. However, user needs, such as the need to see seasonal trends and variations, and available information and resources, form the basis for deciding if commodity balances and energy balances should also be compiled on a quarterly and/or monthly basis. Because of the high frequency of compilation, monthly and quarterly balances are compiled on a more aggregate level than annual balances, and generally focus on selected flows of supply (such as production, imports, etc.) and some aggregated figures for consumption. While supply and sales data may be available on a monthly basis, detailed consumption data are usually not available as frequently due to the costs of collecting that data. Box 6.1 describes some examples of more frequently compiled commodity balances.

Box 6.1:
Examples of monthly/quarterly commodity balances

There are a number of examples of balances compiled on a monthly and/or quarterly basis. In the United Kingdom, quarterly energy balances have been compiled for several years, and at a fairly detailed level of aggregation, showing, for example, a breakdown of final consumption (iron and steel, other industries, transport, other final users and non-energy uses), as well as a breakdown of transformation. The quarterly energy balances are published three months after the end of the quarter (see table 6.12).

An example of a monthly commodity balance is provided by the Joint Organizations Data Initiative Oil (JODI Oil), which consists of an international monthly data collection for a commodity balance on oil statistics (in tons or barrels), focusing on major flows (such as production, import, exports, stock changes and use) for oil and oil products (such as crude oil, NGL, other hydrocarbons, LPG, naphtha, gasoline, kerosenes, residual fuel oil and other petroleum products). More than 90 countries provide monthly data and the JODI Oil database serves as a basis for understanding the short-term oil market. For more information, see www.jodidata.org.

C. How to compile commodity balances

1. Introduction

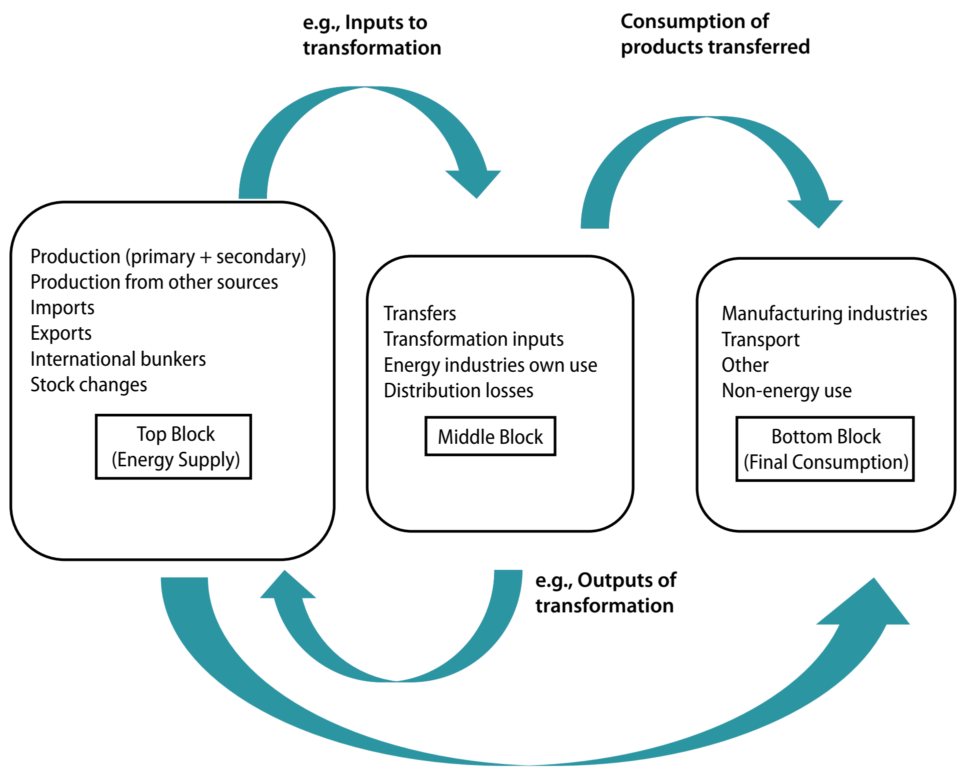
6.27. A commodity balance can be compiled for one single energy product, or for all energy products in a country. In the latter case, it will resemble the energy balance, except that there will be no "Total" column, and the layout, especially the transformation sector, will usually be different. As the main components in the commodity balances are the same as in the energy balances, most of the descriptions of those components are valid for the energy balance as well. The differences in format are further described in section D of this chapter on compiling energy balances.

6.28. A complete commodity balance for all energy products is usually compiled by using multiple data sources. One of the challenges of using various data sources is getting comparable data for both supply and use. If statistics were perfect, supply data would correspond to use data, but this is not always the case due to inaccurate or inconsistent data. Another common problem is that data from different sources for the same variable are not always equal, often due to inaccuracy in the data, or respondents' different interpretation and understanding of definitions, as well as differing time frames. These issues are further described below in section C, subsection 5 on statistical differences.

6.29. The different flows in the commodity balance are described in the following sections, starting from the top with the supply side.

6.30. The main sections of the commodity balance are production, trade, transformation, energy industries own use, final energy consumption and non-energy use. One of the reasons that energy producing industries and transformation are so clearly separated from final consumption in the balance is to avoid double counting between primary and secondary products. A large proportion of the energy used as final consumption is secondary energy that has been transformed from the primary energy so as to be more useful for final consumption purposes; for example, gasoline and electricity (rather than crude oil or lignite). Technical descriptions of energy flows can also be found in chapters V and VIII of IRES.

Figure 6.1:
Flows in the commodity balance



6.31. Figure 6.1 shows the flows of energy within a commodity balance. When a primary energy product is transformed, the output of the transformation process will be presented in the top block (of the output product) since production includes both primary and secondary production. There can also be movement from the middle to the bottom block if a product is transferred in and then consumed in the bottom block.

2. Energy supply

6.32. Supply in the commodity balance consists of the following components. This is referred to as the "Top block" in IRES.

> Energy supply =
>
> Production (of primary and secondary products)
>
> + Production from other sources
>
> + Imports of energy
>
> – Exports of energy
>
> – International (aviation and marine) bunkers
>
> – Stock changes

6.33. It should be noted that only selected components of the balances are contained in this section in order to discuss relevant issues of concern. However, detailed definitions of these components are not repeated in this chapter as they are already included in chapter V of IRES.

(a) Production

6.34. **Primary production**: This is the capture or extraction of fuels or energy from natural energy flows, the biosphere and natural reserves of fossil fuels within the national territory in a form suitable for use. Inert matter removed from the extracted fuels and quantities reinjected, flared or vented are not included. Examples of primary production are hard coal, crude oil and fuelwood.

6.35. **Secondary production**: This is the manufacture of energy products through the process of transformation of primary or secondary fuels, or energy. The quantities of secondary fuels reported as production include quantities lost through venting and flaring during and after their production. Examples of secondary products are gasoline, kerosene and blast furnace gas.

6.36. Annex A of IRES gives a complete list of primary and secondary products. The distinction between primary and secondary energy production is essential for energy balances; however, commodity balances essentially treat the two forms of production identically. One main difference from the energy balance is that in a commodity balance both primary and secondary products are normally included in the top row under production. In the energy balance, only primary production is included in the top row, while secondary production appears as output of transformation processes to avoid double counting supply. Those concepts will be described in the last part of section D, subsection 9 of this chapter.

6.37. **Production from "other sources"**: Production from other sources represents an addition to the supply of an energy product already accounted for in the production of another energy form, or, in the case of coal, can also be recovered slurries, middlings and other low-grade coal products that cannot be classified according to type of coal and includes coal recovered from waste piles and other waste receptacles. Examples include the blending of petroleum gases with natural gas (see more information in IRES, paras. 6.44 and 6.47.) In energy balances, these additions to a product's supply are normally accounted for as a transformation output.

(b) Imports and exports

6.38. This includes imports and exports of both primary and secondary energy products. A main data source is a country's external trade statistics, which are usually based on custom declarations. Another source could be separate trade information from the producing units. The applicability of this data may depend on a variety of factors, such as inconsistent declaration of energy products due to difficulties distinguishing between, for instance, oil products with a near-identical chemical composition.

6.39. A particular problem with using external trade statistics is related to the classifications used – the Harmonized Commodity Description and Coding System (HS) and, in a few cases, the Standard International Trade Classification (SITC); while for energy statistics purposes, SIEC is used. Although a conversion table for these different classifications is given in IRES, a simple conversion of data is often not possible, mainly because the HS does not sufficiently distinguish among the various oil products (in fact, motor gasoline, jet fuel and white spirit are all located in the same HS subheading). As a result, in many cases, a nationally adapted version of the HS is used, which may provide more detail that aligns better with the energy products in SIEC. Still, even in such cases, caution has to be used in matching the relevant products, as similar names do not always carry similar definitions.

(c) International bunkers

6.40. This should include both marine and aviation bunkers, as defined by IRES. Bunkers have generally been subtracted from a country's energy supply since the principle of their exclusion from a country's greenhouse gas inventories has become the international standard.

6.41. It should be noted that there may be an overlap in the export data of transport oils for shipping/aviation and data included in the bunkers category. If it is possible to identify bunker fuels in the export figures, then these should only be included in the bunkers category and the export figures should be adjusted.

6.42. It may be problematic to separate national and international bunkering, both for aviation and shipping, as national bunkering in aviation and shipping should be accounted for as final consumption in transport. If the data sources are oil distributors, they do not always know the exact breakdown of sales to the different user groups, such as fishing, national sale and international sale. Another problem is that much of the oil is sold to resellers of oil. A solution may be to purchase a customer list with identification numbers from the oil companies, and to identify the right industry from registers, and/or contact some of the main resellers to get the distribution of oil by user group. Some countries also collect these kinds of data from aviation or shipping companies directly.

6.43. Other possible data sources could be oil companies that sell oil products, if they have information about their sales to different customer groups. Furthermore, some statistical offices have transport statistics that cover energy consumption in aviation and shipping. Chapter V of IRES gives more information on the treatment of bunkers.

(d) Stock changes

6.44. Except for electricity and heat, energy is sometimes not consumed immediately after its production or provision. Since energy can be stored, the supply of energy will not correspond to consumption unless it is adjusted to account for stock changes. To ensure comparability of energy statistics with the accepted practice in other areas of economic statistics, such as national accounts, stock changes are defined in IRES as closing stock minus opening stock in the reference period. In keeping with this convention, a positive value of stock change is a stock build, representing a reduction in the supply available for other uses, while a negative value is a stock draw, representing an addition to the supply for other uses. Any deviation from this convention, which is defined in IRES, should be clearly indicated in a footnote.

6.45. The stock changes flow should not be used as a balancing item in lieu of the statistical difference flow.

3. Transfers, transformation, energy industries own use and losses

6.46. This part of the energy balance is referred to as the "Middle block" in chapter VIII of IRES. At first glance, the components in the "middle block" are similar between the energy balance and the commodity balance, but the methodology of the transformation part is different.

(a) Transfers

6.47. The transfers row in IRES is a statistical tool for moving energy between products when, for example, they are reclassified or have changed names. Transfers comprise products transferred and interproduct transfers (both of which are defined in

IRES). In theory, the commodity balance transfers row should sum to zero but may not, if recycled products are included.

(b) Transformation

6.48. Transformation processes are an important part of the balance and show changes from a primary or secondary fuel, by physical and/or chemical means, into a secondary or tertiary energy product which is better suited to the uses for which it is intended (see definition of primary and secondary products in section C, subsection 2 (a) of this chapter).

6.49. The commodity balance shows transformation input as a separate category. This is energy products used as input to produce secondary energy products, such as waste used in the production of district heating, or crude oil converted to oil products. Energy used as own use (to support operation, e.g., for heating, etc.) is not included here but in the category "energy industries own use". The distinction between quantities of energy transformed into a new product and quantities used to maintain the process is an important one, as it allows better calculation of efficiencies in the transformation process.

6.50. The transformation category should show separately the input in the main energy producing industries. However, the level of detail depends on data availability. The output – the production of secondary energy products – is often presented in the top row in the commodity balance, together with primary production, as illustrated in tables 6.1 and 6.2. This is in contrast to the energy balance in which the transformation flows show both input and output, with input as negative numbers and output of secondary energy products as positive numbers. An exception to this is when the output of the transformation activity is included in the commodity balance of a different family of energy products, such as blast furnace gas blended with natural gas. While this is still treated as transformation in energy balances, in commodity balances these quantities appear under "production from other sources."

6.51. The various processes of fuel conversion and energy production are described in detail in chapter V of IRES (section D on energy industries and section E on other energy producers), and also in the energy balance table 8.1 in IRES. Examples include the manufacture of coke oven coke from coking coal in coke ovens or the generation of electricity from steam produced by burning fuels. Below are examples on how to treat some specific, less straightforward transformation processes.

6.52. *Heat plants*: The production of heat in heat plants is also the direct result of combustion and is identical in nature to heat raising by final consumers. Heat produced by district heating plants, whose main purpose is to produce and sell heat to other consumers, is considered "main activity production," and should be included in the transformation flow in the balance. In addition, the production of heat (steam) for sale by an autoproducer (another energy industry, a manufacturing industry or other final consumer where the main activity is not heat production) is also considered a transformation activity. By including the heat sold by the autoproducer in the transformation sector, the heat will appear in the total heat supply and its consumption by final users recorded. The fuel used to produce the heat sold must also be included as input in the transformation category.

6.53. If this practice is not adopted, then heat produced and sold by autoproducers would not appear in the balance, with the consequence that the fuel consumption by the autoproducer would be overstated and the heat used by final consumers understated (see more about the definitions and treatment of autoproducers and main activity producers of heat, electricity and combined heat and power (CHP) in IRES, chap. V, paras. 5.45 to 5.48.)

6.54. *Oil refineries*: Crude oil and other oils needed by oil refineries for the manufacture of petroleum products are shown under this heading. Examples on how to treat backflows to refineries and oil products used as feedstock and blendstock in refineries are given in section D, subsection 6 of this chapter.

6.55. *Blast furnaces*: Blast furnaces produce blast furnace gas as a by-product when making pig iron from iron ore. It is sometimes difficult for reporting companies to distinguish between the amount of the coke oven coke transformed into blast furnace gas and the amount that should be considered part of final consumption in the iron and steel industry. While a blast furnace's principal activity is the manufacture of pig iron, the importance of correctly measuring blast furnace gas production for both energy and greenhouse gas (GHG) leads blast furnaces to be considered an energy industry for the purposes of energy statistics.

6.56. If a country knows the actual transformation efficiency, then it is good practice to use this specific efficiency for coke used in blast furnace gas production. However, often the amount of coke oven coke (or other) input needs to be calculated on the basis of the output, generally blast furnace gas, by assuming a theoretical transformation efficiency. Efficiencies used vary between countries and organizations and this may lead to inconsistencies. For example, when 100 per cent efficiency is used, the estimated energy amount of all input to transformation will equal the output of transformation. If 40 per cent efficiency is applied, a value which takes into account considerations about carbon conservation, the input as measured in energy units becomes 1/0.4 = 2.5 times higher than the output. The chosen efficiency depends usually on traditions in the country. Determining which efficiency to use should be considered carefully, since it affects the final consumption level. The estimated amount of coke used in blast furnace gas has to be deducted from final coke consumption, and balances prepared by different institutions or organizations may not be comparable due to differences in efficiencies used.

6.57. *Petrochemical plants*: Petrochemical plants refer to plants which convert hydrocarbon feedstock into organic chemicals, intermediate compounds and finished products, such as plastics, fibres, solvents and surfactants. Feedstock used by the plant is usually obtained from the refinery and includes naphtha, ethane, propane and middle distillate oils (for example, gas oil). While not an energy industry like blast furnaces, the inputs and outputs are considered in the balance transformation block because the petrochemical by-products are often used as energy products, either returning to a refinery or combusted at the petrochemical plant for heat raising.

(c) Energy industries own use

6.58. This represents energy industries' own use of energy products for direct support of their operations, such as energy used for heating, lighting, and for compressors and cooling systems. Energy transformed to secondary energy products should not be included here.

(d) Losses

6.59. Losses occur during the transmission, distribution and transport of fuels, heat and electricity. Transmission losses occur, for instance, for electricity and district heating and can be estimated from technical-based information. For electricity, losses are mainly dependent on the length of power lines, voltage of transmission and distribution and quality of network. In some countries, theft may be a large part of losses; this is sometimes called non-technical losses. Theoretically, this should be noted in the metadata for transparency. For district heating, losses can also be estimated or calculated as the difference between measured energy delivered into the pipelines and the total energy measured by the consumers' meters.

6.60. Losses for fossil fuels can include quantities of oil products that have evaporated, or coal dust lost in transit. The flow "losses" may also include small quantities of secondary gases flared by energy distributors or end consumers. It does not include flaring during extraction in oil/gas fields or at oil/gas terminals since these quantities should not be included in production in the commodity balance.

6.61. In general, losses should be compared to the magnitude of supply. For electricity, losses account for anywhere between 7 and 15 per cent of electricity supply as a general rule; large transmission distances or significant electricity theft (non-technical losses) can result in larger losses. For heat, losses account for about 15 per cent.

4. Final consumption

6.62. This part of the balance is referred to as the "Bottom block" in IRES.

(a) Final consumption

6.63. Final consumption is the energy used directly for energy or non-energy purposes by households, industry, transport, etc. Energy used for non-energy purposes is presented as a separate subgroup in the energy commodity balance. It is a part of final consumption, but not a part of final energy consumption. The subgroups specified for final energy consumption have traditionally been the following:

- Manufacturing, construction and non-fuel mining industries (broken down by main industries)
- Transport (broken down by road, rail, domestic aviation, domestic navigation, pipeline transport and transport not elsewhere specified)
- Other (broken down by household, commerce/public services, agriculture/forestry, fishing and not elsewhere specified, including defence activities)

6.64. The groups in the commodity balance should, in principle, follow the most updated *International Standard Industrial Classification of All Economic Activities* (ISIC) codes as laid out in IRES. However, there are some exceptions. The consumer group "Transport" does not represent energy used by transport industries, although most transport industries are also included in this group. Rather, it includes fuels used for transport purposes regardless of the economic sector to which it is contributing. Examples include fuels used in private cars for households and vehicles used for transport in manufacturing industries.

6.65. Oil products used, for instance, in fishing boats, tractors or military vehicles are, however, not considered as energy used in transport because the purpose is fishing, agriculture and defence rather than transport. Furthermore, energy consumption by airports and other transport infrastructure should be recorded under commerce/public services rather than transport.

(b) Non-energy use

6.66. Energy products used for non-energy purposes include bitumen used in asphalt, lubricants in engines, LPG used in the production of plastic or chemicals, or petroleum coke used to produce carbon anodes. This is not combusted and utilized as energy, but rather remains broadly stored in the material produced.

6.67. Non-energy use is a part of final consumption but not a part of final energy consumption. Energy products used for non-energy purposes are placed in a separate category in the commodity and the energy balances, and are not included in the consumption of the sector where it was actually used. The split between energy and non-energy use is particularly important for CO_2 emission calculations.

5. Statistical difference

6.68. In a commodity balance, the last row to be considered is that of "Statistical difference". It represents the difference between data collected on supply and use. National statisticians should investigate large statistical differences in order to establish whether data may be incorrect or incomplete. Unfortunately, it will not always be possible to correct the data and, in those cases, the statistical difference should not be changed but left to illustrate the size of the problem.

6.69. Deciding whether a statistical difference should be investigated with the reporting enterprise(s) is a matter of judgement. The percentage difference which might be considered acceptable would depend on the magnitude of the supply of the energy product. For major supplies, like natural gas or electricity, efforts should be made to keep the statistical differences below one per cent if possible. On the other hand, for a minor product like coal tar from coke ovens, a 10 per cent error could be tolerated.

6.70. According to the *International Energy Agency (IEA) Yearbook* on energy balances, the statistical difference is lower than one per cent for many countries, measured as a percentage of total supply. However, the relative size of statistical differences depends on which flow in the balance it is compared with. If a country is a large energy producer and exports most of its production, the statistical difference could be much greater when measured as part of domestic supply than when measured as part of production.

Box 6.2:
Norway's check on statistical differences

In Norway, energy production is about six times higher than domestic supply, since large quantities of oil and gas are exported. As a result, statistical differences are much higher when measured as part of supply than when measured as part of production. In some years, the statistical difference for natural gas in Norway has been about one per cent of production, but represented up to 17 per cent of total energy supply. This indicates that there is a problem between the supply and use data, and calls for an investigation of the data sources. A number of different reasons may cause this kind of misalignment. It may be that some of the use is not being captured or some of the consumption is being double counted. Alternatively, the production may incorrectly include some reinjected gas (this was the case in Norway some years before it was detected) or exports may include quantities that are being attributed to a different reference period. A concerted effort should be made to reconcile supply and use. In the case of Norway, a closer examination of the statistical differences showed that one of the reasons for the difference was that certain oil/gas fields were missing from the Norwegian export figures.

6.71. It is also important to look at the trends in statistical difference for each product. If, for instance, the supply of an energy product is always higher than its consumption, this could indicate a systematic error that has to be investigated and corrected. If the statistical difference fluctuates around zero, it is more likely that it comes from random measuring or data reporting errors, such as imports/exports reported in the previous/next period and unaccounted stock changes.

6.72. When commodity balances are constructed, they may also show a statistical difference which is zero (a "closed" balance). This apparently ideal position should be regarded with suspicion as, in almost all situations, it indicates that some other statistic in the commodity balance has been estimated to close the balance. This usually occurs when the data come from a single reporter (e.g. a refinery or an iron and steel plant) possessing all the data making up the commodity balance and is therefore able to adjust figures to close the balance. For information and an appreciation of the data problems encountered by the enterprise concerned, the statistician should find out what element(s) has (have) been estimated to balance the account and encourage the data reporter to submit "true" data, even with numbers that do not completely agree.

6. Possible elements for reconciliation

6.73. A common problem when constructing commodity balances is having different figures for the same variable, but from different data sources that are not equal. It is important to determine which number is the more reliable. The method used for reconciliation should be based on what might be the most likely result, or which figures are judged to be of the best quality. These best figures should be retained and the statistician could consider modifying data from the other source to fit logically into the commodity balance system.

6.74. Consumption by certain user groups may also need to be estimated as a residual. If information is available on total consumption of heating oil in the country, and the share that goes to manufacturing industries is also available, it may be appropriate to assume that the rest goes to services and households. Distribution keys, such as population, employees or production figures, can be used to distribute the figure further among industries, services or households, preferably combined with survey information on energy consumption per produced unit, energy per employee, etc., by industry, because this can vary significantly among industries.

6.75. For oil and gas producing countries, additional problems may arise if:

- Offshore fields are on the border between the national and international continental shelf;
- Crude oil/condensate and natural gas liquids (NGL) are transported through the same pipelines and broken down into different products at terminals abroad;
- Inconsistencies occur in how production and exports are shared between domestic and foreign oil companies;
- Inconsistencies between production and import/export figures, for instance, because the crude oil/NGL split is classified differently for production data and foreign trade statistics, and data are based on different data sources;
- Foreign trade statistics do not cover all oil exporting fields;
- The measuring point for oil and gas is different in external trade statistics and production statistics; and
- Products are classified and reported differently by producers and importers.

7. Country examples

6.76. Several countries, such as the United Kingdom and Azerbaijan present separate commodity balances for each energy product, with time series for each product. Table 6.1 shows what the tables look like for wood, crude oil and motor gasoline for Azerbaijan (on their website, time series with commodity balances are published separately for each product). For instance, Norway (see table 6.2) – as well as IEA and Eurostat – prepares and publishes more complete commodity balances with all energy products gathered in the same balance, almost like in the energy balance. Table 6.3 shows the example of Canada's commodity balances combining all energy products.

Table 6.1:
Commodity balances for fuelwood, motor gasoline and crude oil for Azerbaijan

Energy product, unit	2011 Fuelwood, 1 000 tons	2011 Motor gasoline, 1 000 tons	2011 Crude oil, petajoule
Production	192.9	1 295.5	1 966.2
Imports	-	1.0	-
Exports	-	-96.5	-1 681.9
International bunkers	-	-	-
International marine bunkers	-	-	-
International aviation bunkers	-	-	-
Stock changes	-0.5	-257	-2.6
Total supply	**192.4**	**1 174.3**	**281.8**
Statistical difference	**-**	**27.9**	**6.3**
Transfers	**-**	**-**	**-**
Transformation input	**-**	**-**	**273.1**
Energy industries own use	**-**	**-**	**0.4**
Losses	**-**	**-**	**2.0**
Final consumption	**192.4**	**1 146.4**	**-**
Final energy consumption	**192.4**	**1 146.4**	**-**
Industry and construction	**0.2**	**-**	**-**
Iron and steel	-	-	-
Chemical and petrochemical	-	-	-
Non-ferrous metal	-	-	-
Non-metallic minerals	-	-	-
Transport equipment	-	-	-
Machinery	-	-	-
Mining and quarrying	-	-	-
Food and tobacco	0.1	-	-
Paper, pulp and printing	-	-	-
Wood and wood products	-	-	-
Textile and leather	-	-	-
Construction	0.1	-	-
Non-specified		-	-
Transport	**0.5**	**1 134.3**	**-**
Road	-	1 134.3	-
Rail	0.5	-	-

Domestic aviation	-	-	-
Domestic navigation	-	-	-
Pipeline transport	-	-	-
Non-specified	-	-	-
Other fields of economy	**191.7**	**12.1**	-
Agriculture, forestry and fishing	2.2	11.1	-
Community and public services	36.9	-	-
Households	152.6	1.0	-
Not elsewhere specified	-	-	-
Non-energy use	-	-	-

Table 6.2:
Aggregated energy commodity balance for Norway, 2011 (preliminary figures)

Unit	Coal 1 000 tons	Crude oil, NGL, condensate 1 000 tons	Oil products 1 000 tons	Natural gas Million Sm3	Other gases 1 000 toe	Biofuels and waste 1 000 toe	Electricity GWh	Heat GWh
Production	1 386	94 093	15 846	105 697	1 029	1 565	128 144	6 113
Production of primary energy products	1 386	94 093		105 697	-	1 565	-	-
Production of secondary energy products			15 846		1 029		128 144	6 113
From other sources	-	-	-	-	-	-	0	-
Imports	1 166	1 140	3 624	1		161	11 255	
Exports	1 504	78 312	9 659	99 842		86	14 329	
Int. marine bunkers[d]			465					
Int. aviation bunkers			413					
Stock change	206	-450	-132					
Total energy supply	**1 254**	**16 471**	**9 652**	**6 335**	**1 029**	**1 639**	**125 070**	**6 113**
Statistical difference	147		-69	-49				
Transfers								
Transformation input	120	14 460	2 426	745	82	416	700	10
Elec/CHP/heat plants			44	745	82	416	700	10
Oil refineries/other		14 460	2 382					
Energy industries own use[e]			234	4 145	560		8 679	884
Losses				15	77		10 292	1 316
Final consumption	**986**	**841**	**6 211**	**999**	**311**	**1 223**	**105 400**	**3 903**
Final energy consumption	**909**		**6 211**	**408**	**311**	**1 223**	**105 400**	**3 903**
Manufacturing, mining and quarrying	909		567	282	299	466	42 831	392
Manufacture of industrial chemicals	905		119	103	276	21	7 324	133
Production of metals	295		35	86	16	3	23 777	10
Manufacture of paper and paper products	481		80	2	3	315	4 362	0
Mining and quarrying and other industries	-		333	91	5	126	7 368	249
Transport	129		4 470	75	0	107	684	-
Other sectors[f]	-		1 173	51	11	650	61 885	3 510
Non-energy use[g]	**4**	**841**		**591**	-	-	-	-

[a] Including waste oil, paint and varnish.

[b] Including blast furnace gas, refinery gas and fuel gas. According to international recommendations, there is no column for "other gases" in the energy balance, but this is a national modification for Norway in order to make the use of these gases more visible and not only as a part of other products. However, it may be changed in the future.

[c] Including, wood, wood waste, pellets, black liquor, biodiesel, bioethanol, methane (from biogas), charcoal, etc.

[d] Deliveries in Norwegian harbours to ships in foreign trade, irrespective of the ship's nationality.

[e] In addition, 479 million Sm^3 natural gas was flared off in oil fields and terminals (not included here).

[f] Households, private and public services, construction, agriculture, fishing and defence.

[g] A high share of the coal and coke is used for non-energy purposes/reductants, but this is considered as energy consumption because these industry processes generate heat/gases that might be utilized for energy purposes.

Table 6.3:

Primary and econdary energy, natural units – Canada, 2011

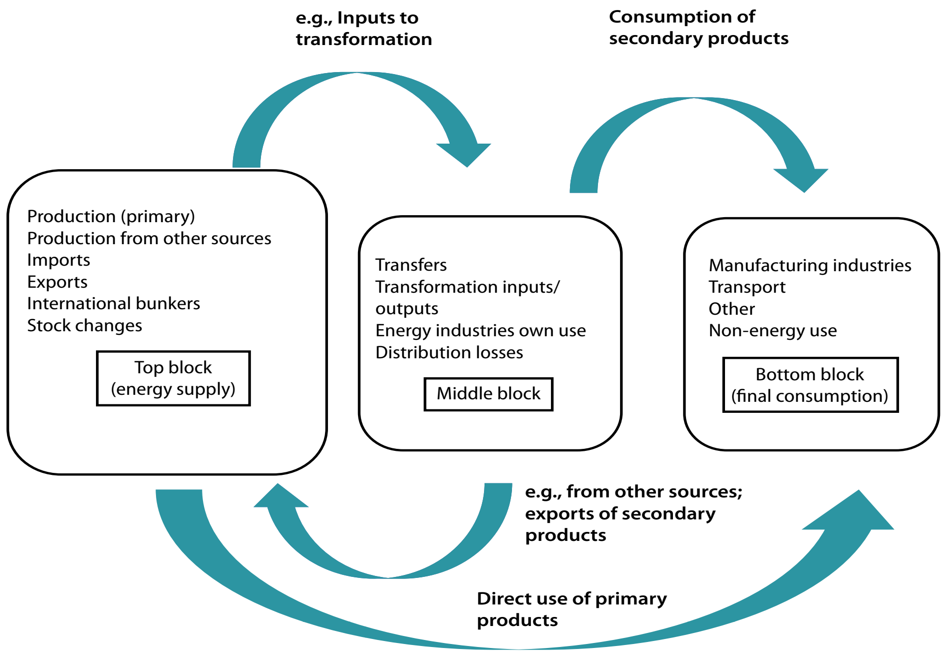
	Primary energy						Secondary energy			
	Total coal	Crude oil	Natural gas	Gas plant natural gas liquids (NGLs)	Primary electricity, hydro and nuclear	Steam	Coke	Coke oven gas	Total refined petroleum products	Secondary electricity, thermal
	kilotons	megalitres	gigalitres	megalitres	GWH	kilotons	gigalitres	megali-tres	GWH	
Supply and demand characteristics										
Production	68 675.4	200 842.4	157 455.8	29 804.1	496 974.6	.	2 479.6	1 285.0	x	138 231.4
Exports	x	152 027.9	82 449.8	6 891.3	67 246.3	.	..	.	29 322.7	.
Imports	x	43 203.8	26 698.4	613.5	11 370.7	.	664.0	.	x	.
Interregional transfers	x	0.0	0.0	0.0	0.0	.	.	.	0.0	.
Stock variation	-465.2	8.9	-3 566.2	762.6	.	.	-137.4	.	-596.3	.
Interproduct transfers	.	.	-2 632.2	.	.	.	.	.	-3 360.7	.
Other adjustments	936.0	13 484.9	-690.3	440.9	.	.	..	.	22 489.1	.
Availability	**41 372.3**	**105 494.2**	**101 948.2**	**23 204.6**	**441 595.1**	**..**	**3 280.9**	**1 285.0**	**120 132.9**	**138 231.4**
Stock change, utilities and industry	..	.	.	.	.	.	..	.	..	.
Transformed to other fuels										
Electricity by utilities	35 833.7	.	9 645.1	.	.	.	.	.	1 370.2	.
Electricity by industry	..	.	5 398.9	.	.	.	.	15.9	505.5	.
Coke and manufactured gases	3 390.3	.	.	.	.	.	.	.	.	.
Refined petroleum products	.	105 904.4	1 038.7	2 786.1	.	.	.	.	.	.
Steam generation	1.9	.	810.6	.	.	-9 657.8	.	2.2	4.0	.
Net supply	**2 146.4**	**.**	**85 054.9**	**20 418.5**	**441 595.1**	**9 657.8**	**3 280.9**	**1 266.9**	**118 553.1**	**138 231.4**
Producer consumption	0.2	.	16 332.6	680.2	57 606.5	.	.	.	14 260.1	.
Non-energy use	256.7	.	3 777.4	16 979.2	.	.	439.8	.	13 900.0	.
Energy use, final demand	**1 838.5**	**.**	**71 650.3**	**22 470.3**	**480 149.3**	**9 657.8**	**2 596.8**	**1 266.9**	**87 706.0**	**.**

D. Compilation of energy balances

6.77. As mentioned earlier, an energy balance is an accounting framework for the compilation and reconciliation of data on all energy entering, exiting and used within the national territory of a given country during a certain reference period. The energy balance shows all forms of energy in a common accounting unit, such as joule, ton of oil equivalent or GWh, and shows the relationship between the inputs and the outputs from the energy transformation industries.

6.78. The benefit of the energy balance is that all data are recorded in the same unit, which makes it possible to compare different energy products with each other and to calculate total energy figures. This also makes possible the measuring of energy trends on a macro level. Exact figures for changes in total annual energy consumption or production, which are of great importance for, among others, energy analysts and politicians in the country, are calculated in an energy balance. Furthermore, as mentioned previously, energy balances provide input to basic energy indicators, such as energy independence, energy efficiency or energy intensity.

Figure 6.2:
Flows in the energy balance



6.79. Figure 6.2 shows the flows of energy within an energy balance. As mentioned earlier in this chapter, one of the major differences, compared to the commodity balance, is that the outputs of the transformation process appear in the middle block as positive figures, while transformation inputs are negative figures. A large proportion of primary energy is usually transformed to secondary products, such as gasoline and electricity, which makes it more suitable for different purposes but some primary energy is also used directly as final consumption. This is the case of coal used directly in some manufacturing industries, or for heating purposes. Fuelwood is also used for heat raising without any transformation. Furthermore, secondary energy products are not always used in final consumption but can be exported or used in energy producing industries for further transformation.

1. Level of detail

6.80. In section B of this chapter, the level of aggregation was discussed in terms of data collection. The same considerations need to be made for the level of aggregation in the published energy balance. While an extended balance with individual products can be made available online, attention is often drawn to a country's aggregated energy balance that should fit on one or two pages.

6.81. A country should strive to show separately each major part of its energy supply in an aggregated form. Dedicating a whole column to manufactured gases would make sense for a country with a large iron and steel industry, for example, while a country using large quantities of biomass may want to give a specific form of biomass, such as bagasse or coconut husks, its own column. On the contrary, a country with no nuclear or hydro production would not want to "waste" a column of the balance on either of these. In general, and subject to the energy supply specifics of each country, an energy balance may want to include a column for coal (possibly split between primary and secondary products, and possibly including peat), primary oil and secondary oil (possibly combined if primary oil is not large), natural gas, biofuels and waste, electricity and heat. While primary electricity and heat (such as nuclear heat or hydroelectricity) can be accounted for directly in the figures for electricity and heat production, it is often the case that these are given their own columns, with transformation or transfers into the electricity and heat columns, in order to explicitly highlight them (rather than as an aggregate).

2. Calculating an energy balance

6.82. When calculating an energy balance, there are certain principles that should be adhered to and some presentational aspects that need to be decided. To highlight the different aspects under discussion, practical examples have been included throughout the chapter.

6.83. First, an energy balance requires the use of a common unit to sum together the contributions of the various different products collected in different physical units. The common unit needs to be an energy unit, such as joule, ton of oil equivalent (toe), GWh, etc. Therefore, it is necessary to start with the commodity balances and convert them to a joint energy unit by using calorific values for each energy product. Calorific values represent the amount of energy per unit of mass or volume of a given product. This could be done using Microsoft Excel spreadsheets, with the complete commodity balance and calorific values for the product being included in separate sheets. The commodity balance can then be converted automatically to an energy balance by multiplying the figures in the commodity balance with the calorific values for each energy product, and at the same time making the necessary adjustments in the layout. The layout adjustments imply changes in the transformation category. In particular, secondary energy production is represented as positive numbers in the transformation and not any longer in the top row of the balance to avoid double counting, as the corresponding primary energy is already included under production. The energy input remains in the transformation category but as negative figures. The advantage of such representation is that the sum of the input and output will represent the losses of the transformation process.

6.84. Furthermore, the primary energy form of electricity and heat produced from non-combustible sources, such as nuclear, hydro and wind, has to be calculated. This is explained in section D, subsections 5 and 6 of this chapter.

6.85. The energy balance could also be compiled in a more advanced database, where basic energy statistics are prepared, adjusted and inserted in automatically generated balances tables. A database solution has some advantages. It may, for instance, be easier to extract long time series, but technically more challenging to build such a database.

6.86. Using a few selected flows of the commodity and energy balances, the diagram below provides a simple example of how a commodity balance expressed in thousand metric tons (kt) can be converted into an energy balance in terajoules (TJ).

Figure 6.3:
Illustration of Calculating an Energy Balance from a Commodity Balance

Commodity Balance

	Crude oil (kt)	Motor Gasoline (kt)
Production	100	30
Import		
Export	10	24
Supply	90	6
Oil Refineries	88	
Final Consumption	2	6

Energy Balance

	Crude oil (TJ)	Motor Gasoline (TJ)
Production	4230	
Import		
Export	423	1063
Supply	3807	-1063
Oil Refineries	-3722	1329
Final Consumption	85	266

Motor gasoline in kt x 44.3 TJ/kt = Motor gasoline in TJ

Crude oil in kt x 42.3 TJ/kt = Crude oil in TJ

6.87. Figure 6.3 shows how crude oil and motor gasoline are presented in both the commodity balance format and the energy balance format. Under the commodity balance, production includes both the primary production of crude oil and the secondary production of motor gasoline (i.e. the refinery output), so that any total production figure (or indeed a total supply figure) calculated from a commodity balance will lead to double counting. On the contrary, the energy balance only accounts for the primary production. Therefore, in the supply line, because in this example the country exports most of its secondary production, a negative supply of motor gasoline is shown (which means that summing all products' energy supply will not result in double counting anything). Finally, in the oil refinery line of the commodity balance, there are the refinery inputs but no outputs, thus an input/output check is not possible, as in the energy balance. (Although in the above example the other refinery outputs, such as diesel and fuel oil, are not shown.)

3. Calorific values

6.88. When energy data are collected, they are usually reported in physical units for fossil fuels, such as tons for oil and coal, million cubic metres or terajoules (on a gross basis) for natural gas, and tons or cubic metres for biofuels. To convert these to energy units, the number in physical units is multiplied by the theoretical energy content in the fuels, which can be on a gross or net basis. In chapter IV of IRES, it is recommended that energy balances be presented on a net calorific value (NCV) basis. This is the amount of heat that is actually available from the combustion process, in practice, for capture and use. This contrasts with the gross calorific value (GCV), which includes latent heat that is locked up in the evaporation of water present in the fuel before combustion. NCV is recommended because, when measuring energy products against each other, there is interest in comparing the useful energy of each source.

6.89. When reporting natural gas data in energy units, such as petajoule (PJ), oil and gas companies often use GCV (by convention). The difference between gross and net calorific values for natural gas is about 10 per cent, which means that natural gas measured in GCV can be converted to NCV by multiplying by 0.9. Similarly, the difference between NCVs and GCVs for coal and oil is typically 5 per cent. That being so, for coke oven coke, which has a higher calorific value than primary coals, the difference between GCVs and NCVs is probably closer to 1–2 per cent (see table 4, annex B of IRES for the difference between gross and net calorific values for various energy products.)

6.90. The calorific values in different energy products vary between countries in time and in some cases by flow. For example, the calorific value of nationally produced crude oil may differ greatly from that of imported crude. The calorific value of imports may change in time depending on the origin or, for bituminous coal, on the proportion of the amounts used for electricity production and within industry, whose calorific values will be generally very different. It is recommended that country-specific (and in some cases flow-specific) calorific values be used if they are available. In the event that the actual values in the country are not available, and no regional-specific value is available either, the default calorific values as presented in table 4.1 of IRES may be used. Uncertainties in calorific values will impact the quality of the energy balance, and of the derived indicators, such as, for example, the estimates of CO_2 emissions.

4. Choice of the primary energy form

6.91. Electricity and heat, as defined by SIEC, can be produced by numerous sources other than thermal combustion. For example, electricity can derive from wind energy, hydro energy, solar energy, wave energy, tidal energy, other marine energy, geothermal energy and nuclear energy; heat can derive from geothermal energy, nuclear energy and solar thermal energy.

6.92. For each of these non-combustion sources, it is necessary to define the form of the primary energy to be considered. For example, in the case of hydro energy, a choice must be made between the kinetic energy of falling water (or even the potential energy of the water before this) and the electricity produced. For nuclear energy, the choice is between the energy content of the nuclear fuel, which depends on the technology employed in the nuclear power plant, the heat generated in the reactors and the electricity produced. For photovoltaic electricity, the choice is between the solar radiation received and the electricity produced.

6.93. The principle adopted by IRES is that the primary energy form should be the first energy form downstream in the production process for which multiple energy uses are practical on a medium to large scale. The application of this principle leads to the choice of the following primary energy forms:

- Heat from nuclear, geothermal and solar thermal; and
- Electricity from hydro, wind, tide, wave and other marine sources, and solar photovoltaic.

5. Calculation of the primary energy equivalent

6.94. There are essentially two methods that can be used to calculate the primary energy equivalent of nuclear, hydro, geothermal, solar, etc.: the partial substitution method, and the physical energy content method (for further description of these methods, see section D, subsection 6 of this chapter). The principle currently adopted by most countries and international organizations is the "physical energy content" method in which the physical energy value of the primary energy form is used for the production figure. Generally, energy balances only record the physical energy content of the source (i.e. electricity produced in the case of wind, hydro, etc.).

6.95. The partial substitution method is usually used to compare how much conventional fossil energy is displaced by the use of renewable energy, such as wind or hydro.[17] The physical energy content method is described below (further information on the differences between these two methods can be found in section D, subsection 6 of this chapter).

6.96. *Nuclear heat production:* Estimation of the heat content of steam generation from nuclear plant reactors is used only if actual values are not available. However,

[17] It can also enhance cross-country comparability of indicators of energy use per capita defined on the basis of the total primary energy supply as a measure of economic and social development and/or energy poverty.

many countries do not possess this information. As a result, the primary heat production has to be imputed. IEA imputes the primary heat production value for nuclear plants from the gross electricity generation, using a thermal efficiency of 33 per cent. Where some of the steam coming directly from the reactor is used for purposes other than electricity generation, the estimated primary production value must be adjusted to include it, and indicate where it is used, so that the supply and use will balance. As the primary energy form is the heat produced, this implies that the nuclear heat generated will have an efficiency of 100 per cent (similar to hydroelectricity), using the physical energy content method. If countries have estimates of the primary heat production, these should be used in the energy balance, as newer plants will generally have higher efficiencies than older ones. In the United Kingdom, estimates are produced by the operators of the steam generation plants, and published annually in the *Digest of United Kingdom Energy Statistics*.

6.97. The principle of using the steam from nuclear reactors as the primary energy form for energy statistics has an important effect on any indicator of energy supply dependence. Under the present convention, the primary nuclear heat appears as an indigenous resource. However, the majority of countries using nuclear power import their nuclear fuel. If this fact is taken into account, it will reflect those countries' increasing supply dependence on other countries.

6.98. *Geothermal heat production:* Primary heat from geothermal sources is also used in geothermal power plants and a similar back-calculation of the heat supply is used where the quantities of steam supplied to the plant are not measured. In this case, if country-specific information is not available, the default thermal efficiency used by IEA is 10 per cent for geothermal electricity and 50 per cent for the geothermal heat distributed by heat plants. The figure is only approximate and reflects the generally lower-quality steam available from geothermal sources. It should be stressed that if data for the steam input to geothermal power plants are available, they should be used in determining geothermal heat production.

6.99. *Solar thermal heat production:* Solar thermal energy is treated similarly to geothermal energy in that it can also be used both to produce electricity and directly as heat. If it is used to produce electricity and the quantities of steam supplied to the plant are not measured, then IEA, by default, assumes a 33 per cent thermal efficiency. Heat that is distributed by heat plants is accounted for, assuming an efficiency of 100 per cent.

6.100. *Solar PV, hydro, wind, wave, tidal and other marine energies:* For hydro, wind, wave, solar PV, tidal and other marine energies, electricity is the primary energy form. The primary energy is calculated as the physical energy content of the electricity generated by these energy sources, which implies that input equals output.

Box 6.3:
Adapting to IRES

In the past, some countries assumed that some of the potential hydro energy is lost in production and estimated hydro energy as higher than the actual electricity production. For example, Norway assumed earlier that 15 per cent of the hydro energy was lost. However, this practice was changed in accordance with new international recommendations and because the estimated losses are very uncertain. It was also determined that it would not be practical to define falling water and wind not utilized for energy purposes as practical use of energy.

6.101. Within an energy balance, before hydro energy is calculated, the electricity from pumped storage has to be deducted, as shown in figure 6.4.

6.102. It should be noted that these back-calculations for the primary energy equivalent are only necessary if the energy balance structure identifies these non-combusted power sources in their own columns. In some energy balance presentations, these quantities are shown under primary electricity and primary heat, and while this means, for example, that hydro or nuclear are not specifically identifiable, no further methodological assumptions are necessary.

6. Physical energy content method and the partial substitution method

6.103. There are essentially two methods that can be used to calculate the primary energy equivalent of the above-mentioned energy sources: the partial substitution method and the physical energy content method.

6.104. *The physical energy content method:* The principle currently adopted by most countries and international organizations is the "physical energy content" method, in which the physical energy value of the primary energy form is used for the production figure. For primary electricity, this is simply the gross generation figure for the source. Care is needed when expressing the percentage contributions from the various sources of national electricity production. As there is no transformation process recognized within the balances for the production of primary electricity, the respective percentage contributions from thermal and primary electricity cannot be calculated using a "fuel input" basis. Instead, the various contributions should be calculated from the amounts of electricity generated from the power plants, classified by energy source (coal, nuclear, hydro, etc.). In the case of electricity generation from primary heat (nuclear and geothermal), the heat is the primary energy form. As it can be difficult to obtain measurements of the heat flow to the turbines, an estimate of the heat input is often used.

6.105. *The partial substitution method:* In the earlier days of the use of the energy balance methodology, a partial substitution method was often used to value primary energy production. In this method, the primary energy equivalent of the electricity produced from solar, hydro, wind, and other sources represented the amount of energy that would be necessary to generate an identical amount of electricity in conventional thermal power plants. The advantage of this method was to limit the variations in total national energy supply due to changes in primary electricity production in countries where a significant part of their electricity production was from combustible fuels. In years of inadequate rainfall, for example, hydro production would fall and be compensated for by a corresponding amount of electricity produced from fuels which were either produced or imported for that purpose. However, because of the lower efficiency of thermal power generation, a far greater amount of energy in the form of fuels is required to compensate for the electricity lost from hydro plants. This imbalance was overcome by substituting for the hydro production an energy value nearly three times (1/0.36) its physical energy content. The principle was abandoned because it had little meaning for countries, like Norway, where hydroelectricity generation accounted for the major supply of energy and because the actual substitution values were hard to establish as they depended on the marginal electricity production efficiencies. Partial substitution also had unexplained effects on the energy balance as transformation losses appeared which had no physical basis.

6.106. Since these two types of energy balance methodologies differ significantly in the treatment of electricity from solar, hydro, wind and other sources, the share of renewables in total energy supply will be very different depending on the method used. As a result, when looking at the percentages of various energy sources in total supply, it is important to understand the underlying conventions used to calculate the primary energy equivalent.

7. Hydro-pumped storage plants, backflows to refineries, and oil used as feedstock and blendstock in refineries

6.107. In addition to the calculations and adjustments already mentioned in the previous sections, there are a few additional adjustments and issues of concern which are highlighted below.

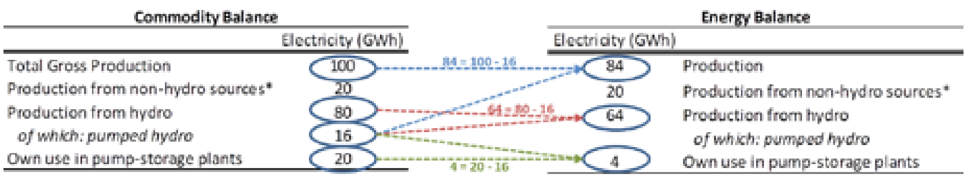
6.108. *Hydro-pumped storage plants:* Hydroelectricity may also be produced from water flow taken from special reservoirs filled by pumping water from a river or lake at lower levels. At pumped storage plants, electricity (taken from the national grid) is used during periods of low demand (usually at night) to pump water into reservoirs for release during times of peak electricity demand when the marginal cost of electricity generation is higher. Less electricity is produced than is used to pump the water to the higher reservoir. However, the procedure is economical when the costs avoided by not using less efficient thermal power stations to generate a similar amount of electricity exceed the cost of the pumped storage procedure. While "regular" hydroelectricity is normally considered renewable energy, the same cannot necessarily be said for pumped hydroelectricity, where the original energy comes from various sources, often fossil fuels.

6.109. As the electricity required to pump water is generated by using energy products recorded in indigenous production or imports elsewhere in the energy balance, any inclusion of pumped storage generation with natural flow hydroelectricity would double count the energy content of the pumped storage generation in "total energy supply". As a result, the pumped storage generation is usually treated differently in energy balances than in commodity balances. In the commodity balance, total electricity, including pumped hydroelectricity, is normally shown, as interest lies in gross generation. The pumped storage figure is normally shown as a memo item of production, and the value under pumped storage plants in energy industries own use will represent all the electricity taken from the grid to pump the water.

6.110. In the energy balance, however, including electricity generation from pumped storage will double count electricity production, overstate energy industries own use and could also affect calculation of the proportion of renewables in the country. Therefore, energy balances exclude electricity produced from pumped storage from production figures (which is possible if recorded in the basic energy statistics as a memo item) and the energy industries own use value will represent the energy lost through pumping.

6.111. Figure 6.4 is a simple illustration of how adjustments are made in energy balances. As shown, in an energy balance, the figures for electricity production from pumped hydro are removed from the figures for total production, production from hydro and own use in pump-storage plants.

Figure 6.4:
Adjusting for electricity generation from pump-storage plants in an energy balance



* This is not a flow in the commodity or energy balances, and it is used in this example only to illustrate how pumped hydro is adjusted between the commodity and energy balances.

6.112. *Backflows to refineries*: These are flows from the petrochemical industry back to refineries. In energy balances, they can be treated as inter-product transfers from various products (e.g. from naphtha or gas/diesel oil) to feedstocks, which are then displayed as transformation inputs into refineries. In Eurostat balances, there is a separate row for these backflows to differentiate them from “normal” inter-product transfers; the backflows of products from the petrochemical industry are shown as negative numbers in this row, while for feedstocks the numbers are positive.

6.113. *Oil used as feedstock and blendstock in refineries*: Oil can be used as feedstock and blendstock in refineries. Feedstock is the oil refineries heat and fractionate, while blendstock is a cold material that is added to get the right specifications of the product. LPG is, for instance, used in the blending of gasoline. Some countries consider both refinery feedstock and blendstock as “refinery input” in the transformation flow in their balances. However, some organizations, like IEA, do not include blendstock as input in refineries in the transformation category; the use of blendstock is instead deducted from the figures for refinery output. In this case, only refinery feedstock is included in transformation inputs in refineries.

8. How to calculate the renewables column

6.114. The detailed energy balance in IRES, chapter 8, “Energy balance”, includes a column entitled “of which: Renewables”. This section describes a proposal for calculating this column. It must be noted that the method described here cannot be used to calculate the share of renewable energy as required in European Union Directive 2009/28/EC on renewable energy, but it gives an indication of the share of renewable energy in the country. As annex A to IRES lists those energy products that should be considered renewable and those that should be considered non-renewable, for the most part, the calculation is fairly straightforward.

6.115. However, a complicating factor arises for electricity and heat since some (but usually not all) of the electricity and heat are produced from renewable sources but, when looking at the consuming sectors, it is not possible to separate out the renewable electricity and heat from the non-renewable electricity and heat. For practical purposes, the easiest solution is to calculate the percentage of electricity produced by renewable sources and to apply this to the electricity consumption and likewise to apply the percentage of heat from renewable sources to the heat consumption.

6.116. However, some countries may have a high level of imports or exports of electricity. If a country, for instance, imports 90 per cent of its electricity, this approach could give an incorrect picture of the renewable part of the electricity consumption. If the technology used in the production of imported electricity is known and available, this could be used to estimate a more realistic share of the renewable electricity consumption in the country. Alternatively, all imports could be regarded as non-renewable. However, in the example given above of a country with 90 per cent imports, this would give a rather low renewable share, which could be incorrect if the imported electricity was renewable.

6.117. The method chosen should therefore depend on the importance of the electricity trade in the country. If imports of electricity are negligible, the best approach would be to calculate the renewable share on the basis of production. Otherwise, imports and exports should be taken into account when calculating this share.

6.118. To illustrate the proposal, table 6.4 details electricity and heat production in Austria by source, with clear indication of which sources are considered renewable and which are considered non-renewable according to IRES. With that information, the shares of renewable electricity and heat production (respectively, 66.4 per cent and 38.2 per cent) are calculated. These shares are applied to the original columns "Electricity" and "Heat" in table 6.5 to yield columns "Electricity *0.664" and "Heat *0.382" in the same table. The latter two columns represent a proxy for the renewable electricity and heat consumption, respectively, and are added to the column "Renewable fuels" to derive the consumption side (bottom block) of the column "Of which: Renewables".

Table 6.4:
Percentage of electricity and heat production from renewable sources for Austria

	Electricity output in GWh	Heat output in TJ
Non-renewable	**22 823**	**49 217**
Coal	6 704	3 067
Peat	0	0
Oil	1 275	7 459
Natural gas	14 346	35 050
Nuclear	0	0
Industrial waste	273	706
Municipal waste (non-renewable)	207	2 780
Heat from chemical sources	18	155
Other sources	0	0
Renewable	**45 114**	**30 427**
Hydro (excl. pumped storage)	38 406	0
Geothermal	1	538
Solar photovoltaics	89	0
Solar thermal	0	0
Tide, wave and ocean	0	0
Wind	2 064	0
Municipal waste (renewable)	300	2 394
Primary solid biofuels	3 574	26 813
Biogases	650	511
Biogasoline	0	0
Biodiesels	0	0
Other liquid biofuels	30	171
Total	**67 937**	**79 644**
% renewable	**66.4%**	**38.2%**

Table 6.5:
Calculation of the column "Of which: Renewables" for Austria

Unit: PJ (to 3 decimal places)	Electricity	Heat	Renewable fuels	Electricity *0.664	Heat *0.382	Of which: Renewables
Primary production	0	0.155	360.063	0	0[a]	360.063
Imports	71.646	0	38.011	NA	NA	38.011
Exports	-63.253	0	-18.489	NA	NA	-18.489
International bunkers	0	0	0	NA	NA	0
Int. marine bunkers	0	0	0	NA	NA	0
Int. aviation bunkers	0	0	0	NA	NA	0
Stock change	0	0	0.217	NA	NA	0.217
Total energy supply	**8.393**	**0.155**	**379.801**	**0**	**0**	**379.801**
Statistical difference	0	0	0.033	NA	NA	NA
Transfers	0	0	0	NA	NA	NA
Transformation processes	244.552	79.319	-223.555	NA	NA	NA
Elec/CHP/heat plants	244.552	79.494	-223.494	NA	NA	NA
Oil refineries	0	0	0	NA	NA	NA
Energy industries own use	-20.102	0	0	NA	NA	NA
Losses	-12.069	-6.450	0	NA	NA	NA
Final consumption	**220.839**	**73.024**	**156.279**	**146.637**	**27.895**	**330.811**
Final energy consumption	**220.839**	**73.024**	**156.279**	**146.637**	**27.895**	**330.811**
Manufacturing, construction, non-fuel mining industries	96.152	10.563	45.876	63.845	4.035	113.756
Iron and steel	12.952	0.623	0.002	8.600	0.238	8.840
Chemical and petrochemical	14.079	2.698	2.101	9.348	1.031	12.480
Other industries	69.122	7.242	43.773	45.897	2.766	92.437
Transport	12.458	0	19.774	8.272	0	28.046
Road	0	0	19.590	0	0	19.590
Rail	6.474	0	0.146	4.299	0	4.445
Dom. aviation	0	0	0	0	0	0
Dom. navigation	0	0	0.037	0.000	0.000	0.037
Other transport	5.984	0	0.001	3.974	0	3.975
Other	112.229	62.461	90.629	74.520	23.860	189.009
Agriculture/forestry/fishing	2.837	0.410	9.672	1.884	0.157	11.713
Households	65.017	32.975	74.862	43.171	12.528	130.560
Other	44.375	29.076	6.095	29.465	11.175	46.736
Non-energy use	**0**	**0**	**0**	**0**	**0**	**0**

[a] This value is not obtained using the mentioned formula, since data in table 6.4 show that all primary heat comes from non-renewable sources.

9. Examples of an energy balance

6.119. As discussed above, various countries and organizations have slightly different formats for their commodity and energy balances. However, the general principles remain the same. Tables 6.5 and 6.6 present aggregated energy balances for Austria and Azerbaijan, respectively, in the format recommended by IRES. In addition, some of the more disaggregated underlying calculations made to produce the Austrian energy balance are given to illustrate some of the concepts presented above.

6.120. In this example, the production of crude oil and NGL is 43 PJ, and is shown in the production row of crude oil. Table 6.8 shows the calculation underlying this figure. The crude oil and NGL are converted separately from kt to PJ using product-specific NCVs. The total primary energy (42.97 PJ) can be seen at the bottom of the table.

6.121. The inputs of crude oil into refineries (343 PJ) figure as a negative number in the oil refineries row of table 6.7.

6.122. The production of secondary oil products is 340 PJ and can be seen as output (a positive number) in the oil refineries row of table 6.7. Again, the underlying calculation of the secondary products can be seen in table 6.8.

6.123. The difference between inputs of crude oil and the output of secondary oil products represents the transformation losses (-3 PJ), which can be seen in the total column in the oil refineries row of table 6.7. Losses of up to about one per cent, as in this case, can be considered reasonable. If losses are significantly higher, that may indicate problems in the data, either in physical units, or in the calorific values used to for the conversion, as discussed later in section D, subsection 10, paragraph 6.127.

6.124. In this example, the primary energy equivalent of hydro is 138 PJ and for geothermal/solar/wind/etc. it is 16 PJ, 7 PJ of which are used directly, and 9 PJ transformed. These two numbers (138 PJ and 9 PJ) can be seen as negative numbers in the electricity/CHP/heat plants row in table 6.7. The underlying calculations for these two numbers can be found in table 6.9.

6.125. As seen in the examples, although some countries may appear different in terms of presentation, the general frame of the energy balance remains consistent.

Table 6.6:
Aggregated energy balance for Azerbaijan, 2011 (toe)

Unit: PJ	All products Total	Crude oil	Petroleum products Total	Natural gas	Renewables	Heat	Electricity	Other fuel products
Primary production	62 541.5	46 949.4	-	15 265.2	326.9	-	-	-
Imports	45.3	-	26.5	-	-	-	11.0	7.8
Exports	-48 783.5	-40 160.5	-2 193.2	-6 360.6	-	-	-69.2	-
International bunkers	-517.2	-	-517.2	-	-	-	-	-
Int. marine bunkers	-79.2	-	-79.2	-	-	-	-	-
Int. aviation bunkers	-438.0	-	-438.0	-	-	-	-	-
Stock change	308.8	-60.9	23.5	345.1	-0.2	-	-	1.3
Total energy supply	**13 594.9**	**6 728.0**	**-2 660.4**	**9 249.7**	**326.7**	**-**	**-58.2**	**9.1**
Statistical difference	257.2	150.7	72.2	23.1	-	-	11.2	-
Transfers	-	-	-	-	-	-	-	-
Transformation processes	**-3 160.0**	**-6 521.0**	**6 177.5**	**-4 439.7**	**-230.1**	**109.4**	**1 745.3**	**-1.4**
Electricity plants	-1 623.2	-	-4.7	-2 529.6	-230.1	-	1 141.2	-
CHP plants	-1 219.7	-	-81.4	-1 774.4	-	32.0	604.1	-
Heat plants	-47.2	-	-1.3	-123.3	-	77.4	-	-
Gas works	54.0	-	66.4	-12.4	-	-	-	-
Blast furnaces	-1.4	-	-	-	-	-	-	-1.4
Oil refineries	-349.0	-6 521.0	6 172.0	-	-	-	-	-
Petrochemical plants	26.5	-	26.5	-	-	-	-	-
Other transformation processes	-	-	-	-	-	-	-	-
Energy industries own use	1 033.1	8.6	429.0	380.1	-	22.1	193.3	-

Losses	1 235.9	47.7	-	832.7	-	13.8	341.7	-
Final consumption	**7 908.7**	**-**	**3 015.9**	**3 574.1**	**96.6**	**73.5**	**1 140.9**	**7.7**
Final energy consumption	**7 290.4**	**-**	**2 430.1**	**3 545.7**	**96.6**	**73.5**	**1 140.9**	**3.6**
Industry and construction	949.0	-	87.2	691.2	0.2	0.3	170.1	-
Transport	1 984.7	-	1 934.0	-	0.2	-	46.9	3.6
Road	1 789.3	-	1 789.3	-	-	-	-	-
Rail	44.2	-	2.8	-	0.2	-	37.6	3.6
Domestic aviation	106.5	-	106.5	-	-	-	-	-
Domestic navigation	35.3	-	35.3	-	-	-	-	-
Pipeline transport	9.4	-	0.1	-	-	-	9.3	-
Non-specified	-	-	-	-	-	-	-	-
Other fields of economy	**4 356.7**	**-**	**408.9**	**2 854.5**	**96.2**	**73.2**	**923.9**	**-**
Agriculture, forestry and fishing	428.0	-	337.2	26.1	1.1	-	63.6	-
Commerce and public services	519.5	-	9.2	116.2	18.5	24.1	351.5	-
Households	3 409.2	-	62.5	2 712.2	76.6	49.1	508.8	-
Not elsewhere specified	-	-		-	-	-	-	-
Non-energy use	**618.3**	**-**	**585.8**	**28.4**	**-**	**-**	**-**	**4.1**

Table 6.7:
Aggregated energy balance for Austria, 2010

(Supplied by IEA following IRES recommendations)

Unit: PJ	Coal and peat	Crude oil	Oil products	Natural gas	Nuclear	Hydro	Geoth., solar, wind, etc.	Biofuels and waste	Electricity	Heat	Total
Primary production	0	43	0	62	0	138	16	233	0	0	492
Imports	126	298	280	427	0	0	0	38	72	0	1 241
Exports	-1	-2	-91	-171	0	0	0	-18	-63	0	-346
International bunkers	0	0	-28	0	0	0	0	0	0	0	-28
Int. marine bunkers	0	0	0	0	0	0	0	0	0	0	0
Int. aviation bunkers	0	0	-28	0	0	0	0	0	0	0	-28
Stock change	17	4	10	26	0	0	0	0	0	0	57
Total energy supply	**142**	**343**	**171**	**344**	**0**	**138**	**16**	**253**	**9**	**0**	**1 416**
Statistical difference	-1	0	0	0	0	0	0	0	0	0	-1
Transfers	0	-1	1	0	0	0	0	0	0	0	0
Transformation processes	-93	-343	314	-119	0	-138	-9	-86	245	79	-150
Elec/CHP/heat plants	-59	0	-18	-119	0	-138	-9	-86	245	79	-106
Oil refineries	0	-343	340								-3
Energy industries own use	-27	0	-23	-19	0	0	0	0	-20	0	-89
Losses	-1	0	0	0	0	0	0	0	-12	-6	-19
Final consumption	**21**	**0**	**464**	**206**	**0**	**0**	**7**	**166**	**221**	**73**	**1 158**
Final energy consumption	**20**	**0**	**405**	**190**	**0**	**0**	**7**	**166**	**221**	**73**	**1 082**
Manufacturing, construction, non-fuel mining industries	18	0	27	101	0	0	0	63	96	11	316
Iron and steel	10	0	0	14	0	0	0	1	13	1	39
Chemical and petrochemical	1	0	1	14	0	0	0	7	14	3	40

Other industries	7	0	26	73	0	0	0	55	69	7	237
Transport	0	0	298	6	0	0	0	20	12	0	336
Road	0	0	294	0	0	0	0	20	0	0	314
Rail	0	0	3	0	0	0	0	0	6	0	9
Dom. aviation	0	0	1	0	0	0	0	0	0	0	1
Dom. navigation	0	0	0	0	0	0	0	0	0	0	0
Other transport	0	0	0	6	0	0	0	0	6	0	12
Other	2	0	80	83	0	0	7	83	113	62	430
Agriculture/forestry/fishing	0	0	10	1	0	0	0	10	3	0	24
Households	2	0	58	56	0	0	4	71	65	33	289
Non-energy use	**1**	**0**	**59**	**16**	**0**	**0**	**0**	**0**	**0**	**0**	**76**

Table 6.8:
Production of primary and secondary products

(Calculation used for table 6.7; the calorific values are specific to Austria)

Production	Quantity (kt)	Net calorific value (kJ/kg)	Equation	Primary products (PJ)	Secondary products (PJ)
Crude oil	877	42 500	=877*.0425	37.27	
NGL	134	42 500	=134*0.0425	5.70	
Refinery gas	392	49 500	=392*0.0495		19.40
Ethane	0	49 500	=0*.0495		0.00
LPG	87	46 000	=87*0.046		4.00
Motor gasoline	1 353	44 000	=1353*0.044		59.53
Kerosene-type jet fuel	476	43 000	=476*0.043		20.47
Kerosene	14	43 000	=14*0.043		0.60
Gas/diesel oil	3 274	42 600	=3274*0.0426		139.47
Fuel oil	699	40 000	=699*0.04		27.96
Naphtha	892	44 000	=892*0.044		39.25
White spirit	70	43 600	=70*0.0436		3.05
Lubricants	96	42 000	=96*0.042		4.03
Bitumen	292	39 000	=292*0.039		11.39
Paraffin waxes	0	40 000	=0*.04		0.00
Petroleum coke	62	32 000	=62*0.032		1.98
Non-specified oil products	231	40 000	=231*0.04		9.24
Total				**42.97**	**340.38**

Table 6.9:
Calculation of primary energy using the physical energy content method

(Calculation factors are specific to Austria)

Non-combusted products	Electricity output (GWh)	Equation	Primary energy equivalent in PJ
Nuclear	0	=(0*.0036)/0.33	0
Hydro	41 596	=(41596-3190)*0.0036	138.26
(of which pumped storage)	3 190		not included in energy balance

Solar PV	89	=89*0.0036	0.32
Wind	2 064	=2064*0.0036	7.43
Geothermal	1	=(1*0.0036)/0.1	0.04
Solar thermal	0	=(0*0.0036)	0
Non-combusted products	**Heat output (TJ)**	**Equation**	**Primary energy equivalent in PJ**
Geothermal	538	=(538/1000)/0.5	1.08
Solar thermal	0	=(0/1000)/0.33	0

10. Checking the energy balance

6.126. The energy balance can be used to provide additional checks on the data quality of the commodity balances, as well as allow expert estimation of missing data in some instances.

6.127. *Transformation losses*: Transformation losses provide an extremely useful check on the accuracy of the energy balance. Should the losses be too high or too low (or positive) this may highlight problems in either the basic energy data from the commodity balance or in the conversion equivalents (mainly the calorific values) used to prepare the energy balance. Box 6.4 below provides an example of transformation gains in Norway and how the data can be fixed in such instances.

Box 6.4:
Transformation gains in Norway

For example, for oil refineries, losses of about 0.5 per cent are acceptable. If the figure is much higher (for example over 2 per cent), or is positive (a transformation gain), then data should be checked. In Norway, the calorific values for crude oil and oil products are somewhat uncertain. They are based on historical values and not checked regularly (which often is the case with many countries). This sometimes leads to refinery gains on a heat content basis, despite a correct balance on a mass basis. If the data compiler judges that the calorific value of the oil products is more reliable than the corresponding value for crude oil, one solution is to estimate the average calorific value of the crude oil, based on the output of oil products (after applying the individual calorific values to products).

6.128. For other transformation processes, losses are much larger as the processes are inherently inefficient. The following ranges for transformation losses by process are given as an indication. However, not all countries will fall within these ranges.

Table 6.10:
Efficiency in different transformation processes (except electricity and heat generation)

Transformation process	Typical plant efficiency
Coke ovens	80–95%
Patent fuel plants	95–100%[a]
Brown coal briquette plants	80–95%
Coal liquefaction plants	10–85%
Gas works (and other conversion to gases)	60–90%
Blast furnaces	≈ 40%
Peat briquette plants	40–100%
Natural gas blending plants	80–95%
Gas to liquid (GTL) plants	40–55%
Oil refineries	98–99.5%
Petrochemical plants	90–98%
Charcoal plants	25–50%

[a] Calculated efficiencies for patent fuel plants can often exceed 100 per cent due to the addition of a binding agent.

6.129. *Statistical difference*: Statistical difference by product is another useful check for the energy balance. While the statistical difference should be checked at the commodity balance level, an energy balance can provide further insights. For a given product (e.g. lignite), a country may be using flow-dependent calorific values because the quality of lignite going to electricity production may not be the same as that used in industry. Similarly, for supply, domestically produced lignite may not have the same calorific value as imported lignite. For exporting countries, only higher-end lignite may be exported. For any given product, if the statistical difference in the energy balance represents a much higher percentage of supply than in the commodity balance, this thus indicates a problem in the calorific values.

6.130. *Generation efficiencies*: Generation efficiencies are another excellent way of using energy balances to reconcile inputs to and outputs from each transformation activity (as discussed above), but particularly those of electricity and heat generation. Inputs and outputs for each product should be compared and any anomalies investigated. For electricity plants and heat plants, this is fairly straightforward. For CHP plants, the inputs to the plant should be compared to the sum of electricity and heat produced, bearing in mind that the efficiency will often vary depending on the proportion of the output that is electricity or heat. Again, not all countries will show efficiencies within the ranges of the typical efficiencies presented below in table 6.11.

Table 6.11:
Efficiency in transformation in thermal power plants and heat generation

Generation process	Typical plant efficiency
Electricity plants	
Coal	32–40%
Oil	35–40%
Natural gas	35–60%
CHP plants	
Coal	50–75%
Oil	50–75%

Natural gas	60–90%
Heat plants	
Coal	80–90%
Oil	80–90%
Natural gas	85–95%

11. Presentational issues in commodity balances and energy balances

6.131. Energy balances may differ slightly in their treatment of some minor aspects, three of which merit explanation here.

Sign conventions

6.132. From a purely presentational aspect, energy balances can appear different because of how the signs for certain flows, such as exports, bunkers and stock changes, are treated. This varies across organizations and countries, but IRES clearly states that stock changes are calculated as closing stocks minus opening stocks, and exports and bunkers are subtracted to obtain total energy supply (therefore the sign for exports and bunkers should always be positive).

Treatment of primary electricity and heat

6.133. Each organization's energy balance must relate the figures in the column for primary electricity produced (e.g. hydro) to those of the electricity column so that its disposal, together with all other electricity, can be accounted for according to the sectors of consumption. Once primary electricity enters a national transmission system, it becomes indistinguishable from electricity produced from all other sources.

6.134. The IEA energy balance presents primary electricity in a separate column and then enters these figures into the transformation matrix as an input with a negative sign, and an identical amount is included in the transformation output of the "Electricity" column. The Statistical Office of the European Union (Eurostat) uses the transfer row to establish the same relation. This distinction is unnecessary in those balances in which these quantities first appear under the "Production of electricity" column.

Inclusion of the primary energy form in the energy commodity balance

6.135. Some organizations, such as Eurostat, also include the primary form of electricity and heat (such as nuclear and hydro, as described in section D, subsection 4 of this chapter) in their commodity balance. The advantage of doing this is that there is no need for reformatting when compiling the energy balance. On the other hand, the primary form of electricity and heat cannot directly be defined as energy commodities, and, in this sense, does not belong here.

12. New technologies

6.136. This section aims to explain those technologies or energy sources that are still not yet commonly included in the energy balance, but which ought to be considered if relevant for a particular country.

Heat from heat pumps

6.137. A heat pump is a device for transferring heat from areas with lower ambient temperatures to areas with higher ambient temperatures and is used, for example, to extract heat from a medium (air, ground, water) outside a building to be used to warm the building's interior. They often use electricity (although some are also designed to use natural gas) to drive motors to perform this function and provide an efficient means of heating. According to current international methodology, energy generated from heat pumps is only included in heat production if the heat is distributed to a third party, for instance from heat plants. Own use of heat from heat pumps is usually not captured. Heat pumps in buildings (e.g. households, businesses, etc.) are, in most countries, considered final consumers, and only the electricity delivered to the consumer is known, not how much electricity was used by the heat pump or how much heat was produced and then consumed. There are some exceptions to this practice. A few countries, for instance Finland and the United Kingdom, do include residential heat pumps in their energy balances.

6.138. Energy from heat pumps may make a significant contribution to a country's energy supply. Furthermore, energy from heat pumps with a coefficient of performance (COP) above a certain value is regarded as renewable energy in the European Union's Renewable Energy Directive. The value depends on the efficiency in electricity generation in the European Union; the required COP in 2010 was 2.5. Energy generated from heat pumps has to be calculated in those countries that have implemented this directive.

6.139. If energy from heat pumps is not included in the energy balance, it can be useful to have supplementary tables showing estimated energy generation from all heat pumps and not just from those covered by traditional energy statistics. The estimation of energy from heat pumps is, however, somewhat unreliable because it cannot be measured accurately and is based on assumptions for determining the COP and the time period the heat pumps are used for heating. The COP depends on the inside and outside temperatures and how and where the heat pumps are installed, among other factors, while the time in use varies with such factors as the outside temperature and the price of alternative energy products. It is difficult to take all these factors into account when estimating the heat generated from heat pumps. The uncertainty and the fact that ambient energy used in heat pumps seldom has alternative uses as energy are arguments for not including heat generated from them in the energy balance. On the other hand, if heat pumps are widespread in a country, a correct estimation of the actual energy consumption will not be obtained, if they are excluded.

Proposal for estimating ambient heat from heat pumps

6.140. To estimate the ambient heat production from heat pumps, one option is to take account of the energy content of the electricity used directly by the heat pump, since this information will mostly likely be available, and to apply a COP above 1.0. The COP is usually between 1.5 and 4 depending on the efficiency and capacity of the heat pump, the temperature of the ambient heat, the air humidity, etc. Exact methods or meters for monitoring the energy generated from all heat pumps in the country are seldom available but calculations can be made on the basis of information about the capacity, type and COP of the different heat pumps, and assumptions about the times per year they are used for heating. However, as mentioned before, these kinds of calculations are somewhat unreliable because they also depend on the inside and outside temperatures, how and where the heat pumps are installed, among other factors. This information is seldom available for heat pumps in the whole country.

6.141. Another option is to follow the European Union's guidance. The European Commission has established guidelines for its Member States for calculating renewable energy from heat pumps using different heat pump technologies, pursuant to Article 5 of Directive 2009/28/EC (on renewable energy). The guidance was also published in the *Official Journal of the European Union*, OJ L 62/67, 6.3.2013 (2013/114/EU). A short outline of the methodology follows:

Box 6.5

Calculating renewable energy from heat pumps using different heat pump technologies

In accordance with Annex VII to the European Union Directive, the amount of renewable energy supplied by heat pump technologies (E_{RES}) shall be calculated with the following formula:

$E_{RES} = Q_{usable} * (1 – 1/SPF)$

$Q_{usable} = H_{HP} * P_{rated}$

Where:

— Q_{usable} = the estimated total usable heat delivered by heat pumps [GWh],

— H_{HP} = equivalent full load hours of operation [h],

— P_{rated} = capacity of heat pumps installed, taking into account the lifetime of different types of heat pumps [GW],

— SPF = the estimated average seasonal performance factor ($SCOP_{net}$ or $SPER_{net}$).

"$\mathbf{Q}_{usable}$" means the estimated total usable heat delivered by heat pumps, calculated as the product of the rated capacity for heating (P rated) and the annual equivalent heat pump hours (H_{HP}), expressed in GWh;

"**Annual equivalent heat pump hours**" (H_{HP}) means the assumed annual number of hours a heat pump has to provide heat at rated capacity to deliver the total usable heat delivered by heat pumps, expressed in h;

"**Rated capacity"** (P_{rated} **)**" means the cooling or heating capacity of the vapour compression cycle or sorption cycle of the unit at standard rating conditions;

"SPF" shall mean the estimated average seasonal performance factor, which refers to the "net seasonal coefficient of performance in active mode" ($SCOP_{net}$) for electrically driven heat pumps or "net seasonal primary energy ratio in active mode" ($SPER_{net}$) for thermally driven heat pumps.

(Default values for H_{HP} and conservative default SPF values are set out in tables in the document describing the methodology.)

District cooling

6.142. District cooling, which is cold water delivered from district plants to buildings for cooling purposes, is not currently a part of reporting requirements to international organizations and most countries do not include this in either their commodity balances or their energy balances. If, at a future date, this type of plant becomes more important, then consideration of whether (and how) to incorporate the information into the balances will need to be examined. An example of one country that includes district cooling in its balances is Norway. Norway collects district cooling, together with district heating statistics from the plants producing them and adds them to district heating in the energy balance. Figures for production, losses and deliveries of district cooling in GWh are added to the district heating figures in the energy balance. The input, which is mainly electricity used in heat pumps for district cooling, is added to the transformation sector in the balance.

Electricity in hybrid cars and electric motor cars

6.143. Hybrid cars have a battery and produce their own electricity from fossil fuels. Own electricity production from fossil fuels in the car should not be included in energy balances. Certain hybrid cars (plug-in hybrid cars) can also be charged, as well as produce their own electricity. Electricity used for charging plug-in hybrid cars and electric motor cars should be estimated and included as electricity consumption in "road transport" in the energy balance. The figures can be estimated on the basis of mileage (for instance, from the vehicle registry) and information about electricity charged per km driven (usually around 0.2 kWh/km). Furthermore, surveys will be required in order to indicate how much electricity is produced when plug-in hybrid cars are charged versus when they produce their own electricity. The surveys must also contain information on where the cars are charged. If the cars are charged at home 30 per cent of the time, the corresponding quantity of electricity produced has to be deducted from the household electricity consumption to avoid double counting. It is expected that the number of electric cars will increase in the future and reliable statistics for electricity used in those cars will become more important.

E. Country-specific examples

6.144. To further demonstrate the different aspects under discussion, practical examples from some countries have been included in this section. Many countries and organizations have slightly different formats for their commodity and energy balances. However, the general principles remain the same.

6.145. A short description of the practices used by Norway and Austria to compile and calculate their energy balances are added below. In addition, table 6.12 shows the format used by the United Kingdom to present its quarterly energy balance. A full balance, with an expected number of rows for energy industries own use and with a greater sector breakdown, is published annually.[18]

[18] Examples of more comprehensive country practices are available on the United Nations' website at https://unstats.un.org/unsd/energy/escm/country_examples/.

Box 6.6:

Country example, Norway: Statistics Norway's energy balance

Statistics Norway publishes its first preliminary annual balances (both commodity balances and an energy balance) about four months after the reference year. The statistics used for these preliminary figures are, among other things, monthly statistics from the previous year for the production and use of electricity, annual figures from the Norwegian Petroleum Directorate for crude oil and natural gas, imports and exports from foreign trade statistics, monthly sales of petroleum products, energy products used for non-energy purposes in manufacturing industries, monthly refinery statistics, and some preliminary figures for energy consumption in some manufacturing industries.

Consumption statistics, e.g. for households and different industrial groups, are often gathered through surveys. As it takes time to obtain these figures, the preliminary balances are often limited in the level of detail of the consumption data.

There may be a delay of up to 18 months or more after the reference year before all the relevant information from surveys is available. Depending on the nature of the publication plan, revised preliminary figures can therefore be published when new information becomes available. For Statistics Norway, new information on district heating, among other things, and preliminary results for the annual survey on energy use in manufacturing are available and are published in a revised preliminary balance in the autumn. For the final balance, published in late autumn, nearly two years after the reference year, final

Box 6.6 (continued)

results from the manufacturing survey and from the yearly electricity statistics are also used. The electricity statistics give information on the use of electricity in services and other groups, and are important for the final energy consumption figures by industrial group. Table 6.2 of this chapter shows an aggregated commodity balance for Norway. In the autumn, Norway publishes energy balances, commodity balances, energy accounts and long time series for energy use. Statistics Norway has a system, in which energy balances, commodity balances and energy accounts are compiled simultaneously in the same compilation system on the basis of more or less the same input data.

Box 6.7:
Country example, Austria: Statistics Austria's energy balance

The data required to prepare the energy balances come from varied sources with different survey aims and therefore have unavoidable inconsistencies.

Data preparation

Data preparation consists of comparing and, if necessary, correcting data from various sources. In addition, there are parallel surveys of exports with, in some cases, inconsistent results (e.g. foreign trade statistics collected by Statistics Austria in accordance with the Federal Statistics Act and the survey by the Federal Ministry of Economics, Family and Youth, in accordance with the Oil Stockholding Act (FORM III)). In other areas, gaps in the data, which would require enormous time and effort to fill (if at all possible), require qualified estimates. Resolving such inconsistencies and/or the selection of a particular source necessitate the use of internal or external expertise that corresponds to the particular requirements.

Since the data situation varies greatly between different energy sources and balance items, there is no uniform procedure. Instead, the procedures specific to each energy source are documented in detail in the description of energy sources (see fuel definitions) and balance items. Some of the data sources are: material input statistics, short-term statistics, electricity and natural gas surveys, household energy consumption surveys, surveys on energy consumption in the service sector and in small- and medium-sized industries, and biomass lighted heating plants.

Publication

The energy balances are currently published solely on Statistics Austria's website – Energy. The detailed Austrian energy balances are available for download for the entire time series. The regional balances are available as synoptic tables only. They are prepared for each calendar year at the regional (Laender) and national levels (total of all Laender). The single balances were released as a time series from 1970 to Y-1 for Austria and from 1988 to Y-1 for the Laender. The publications currently include 27 commodity balances (individual fuels as smallest displayed units), four fuel groups aggregated from the individual energy sources (coal, oil, gas and renewable energies) and the total balance. The individual energy sources are shown both in physical units (tons, 1000 m^3 and MWh) and energy units (terajoules). The fuel groups and the overall balance are only shown in terajoules (TJ).

Box 6.8:
Country example, United Kingdom: Aggregated energy balance

The United Kingdom publishes energy balances on a quarterly basis. These are produced three months after the end of the quarter and are published on the web. An example is shown in table 6.12. The quarterly energy balance is based on monthly and quarterly surveys of energy producers and distributors, which have smaller sample sizes, compared to the more comprehensive surveys conducted on an annual basis. The quarterly balance is a contraction of the full annual energy balance, with final consumption figures comprising only five rows, compared to 23 rows in the full annual balance, with no breakdown of the data on industrial consumption. Other simplifying adjustments are used to estimate data collected only on an annual basis.

The United Kingdom publishes its full commodity and energy balances annually, in its *Digest of United Kingdom Energy Statistics*, on the last Thursday in July – seven months after the reference year. It follows international practice, with a couple of exceptions. It uses gross calorific values as its main method for converting data from commodity balances to energy balances. Information on international aviation is still published in the air transport row within final consumption. However, energy balances produced using net calorific values are published as an annex to its *Digest*. The publication also states that six per cent of aviation turbine fuel is consumed domestically, so balances fully compatible with IRES can be produced. Commodity balances are produced for each fuel and then combined into an overall energy balance. At that stage, consistency checks are performed which may result in further versions of the commodity balances before the final energy balances are produced.

Table 6.12:
Aggregated energy balance for the United Kingdom (third quarter, 2012)

Unit: ktoe	Coal	Manufactured fuel[d]	Primary oils	Petroleum products	Natural gas[e]	Bioenergy and waste[f]	Primary electricity	Electricity	Heat sold	Total
SUPPLY										
Indigenous production	2 555	0	10 678	0	7 909	1 426	4 576	0	0	27 144
Imports	7 099	16	16 631	7 199	8 064	488	0	371	0	39 867
Exports	-74	-26	-8 660	-7 513	-3 740	-87	0	-20	0	-20 119
Marine bunkers	0	0	0	-552	0	0	0	0	0	-552
Stock change[a]	-971	43	700	-106	-725	0	0	0	0	-1 058
1. Primary supply	8 609	33	19 349	-972	11 509	1 827	4 576	351	0	45 282
Statistical difference[b]	10	0	0	47	11	0	0	5	0	74
Primary demand	8 599	33	19 348	-1 018	11 497	1 827	4 576	346	0	45 208
Transfers[c]	0	1	-251	318	-1	0	-491	491	0	67
TRANSFORMATION	-8 195	396	-19 097	18 820	-4 567	-1 406	-4 085	6 416	272	-11 446
Electricity generation	-6 958	-151	0	-207	-4 166	-1 406	-4 085	6 416	0	-10 558
Heat generation	-61	-13	0	-17	-401	0	0	0	272	-219
Petroleum refineries	0	0	-19 097	19 044	0	0	0	0	0	-53
Coke manufacture	-929	837	0	0	0	0	0	0	0	-92
Blast furnaces	-195	-326	0	0	0	0	0	0	0	-521
Patent fuel manufacture	-52	50	0	0	0	0	0	0	0	-2
Energy industry use	0	179	0	1 171	1 079	0	0	544	24	2 997
Losses	0	18	0	0	215	0	0	495	0	729

FINAL CONSUMPTION	404	232	0	16 949	5 635	421	0	6 214	253	30 108
Iron and steel	9	135	0	1	97	0	0	87	0	329
Other industries	270	44	0	1 096	1 623	99	0	2 038	210	5 379
Transport	3	0	0	13 602	0	199	0	88	0	13 892
Domestic	115	53	0	455	2 806	73	0	1 957	5	5 464
Other final users	7	0	0	323	936	50	0	2 044	38	3 399
Non-energy use	0	0	0	1 471	173	0	0	0	0	1 644

a Stock fall (+), stock rise (-).

b Primary supply minus primary demand.

c Annual transfers should ideally be zero. For manufactured fuels, differences occur in the rescreening of coke to breeze. For oil and petroleum products, differences arise due to small variations in the calorific values used.

d Includes all manufactured solid fuels, benzole, tars, coke oven gas and blast furnace gas.

e Includes colliery methane.

f Includes geothermal, solar heat and biofuels for transport; wind and wave electricity included in primary electricity figures.

Chapter VII: Quality assurance frameworks and metadata

A. Introduction

7.1. In the *International Recommendations for Energy Statistics* (IRES), countries are encouraged to manage the quality of their energy statistics most effectively through the use of: (a) national data quality assurance frameworks, as overarching frameworks to provide context for quality concerns, activities and initiatives and explain the relationships between the various quality procedures and tools; (b) quality measures and indicators for describing, measuring, assessing and monitoring over time the quality of their outputs; (c) quality reports to communicate the quality of their outputs on the basis of quality dimensions, such as relevance, accuracy, reliability, timeliness and punctuality, coherence and comparability, accessibility and clarity; (d) quality reviews for evaluating the performance of their energy statistics programmes; and by (e) compiling and presenting metadata to users in amounts and ways that best help them understand and make use of the data.

7.2. This chapter is structured as follows: the first part (sections B to E) covers quality concepts and outlines several quality frameworks to illustrate their structures and the elements they should contain. It defines and describes the different dimensions of product quality, as well as quality measures and indicators for measuring quality, and advice is given on how to balance different quality dimensions against each other, in particular timeliness against accuracy. Quality reports are then described, followed by a summary of the types of quality reviews that can be undertaken to evaluate statistical programmes. In the second part of the chapter (sections F and G), examples of various data quality and validation checks for energy statistics being implemented by the International Energy Agency (IEA) are presented, followed by a summary of country feedback on experience and good practices in assuring output quality, and a brief description of the quality of official statistics in a decentralized statistical system. The chapter ends with a discussion of metadata in section H.

B. Data quality, quality assurance and quality assurance frameworks

1. Data quality

7.3. While the word “quality” can have different meanings depending upon the context in which it is used, *data quality* is commonly defined in terms of its “fitness for use”, or how well the statistical outputs meet their users’ needs. The definition is thus a relative one that allows for various perspectives on what constitutes quality depending on the intended purposes of the outputs.

2. Quality assurance

7.4. Quality assurance comprises all planned and systematic activities that can be demonstrated to provide confidence that products or services are adequate or fit for their intended uses by clients and stakeholders. It entails anticipating and avoiding problems, with the goal of preventing, reducing or limiting the occurrence of errors, as in a survey, for example. It is worth noting that quality assessment is a part of quality assurance that focuses on assessing or determining the extent to which quality requirements have been fulfilled.

7.5. Activities or measures for ensuring that attention is paid to the quality of the data cover not only the final outputs, but also the organization producing the outputs, and the underlying processes that lead to the outputs. The outputs or products are typically described in terms of quality dimensions, such as relevance, accuracy, reliability, timeliness, punctuality, coherence, comparability, accessibility and clarity. The organization or agency exhibits high quality when it maintains a professionally independent, impartial and objective institutional environment, a commitment to quality, and the guarantee of confidentiality and transparency, and provides adequate resources for producing the outputs. For these processes to be considered as being of high quality, the organization should employ sound statistical methodology and cost-effective procedures that minimize the reporting burden.

7.6. Thus, quality is approached along three lines: statistical product (or output) quality; process quality; and the quality or characteristics of the environment in which the office/agency operates. As will be seen below, many of the quality frameworks are structured clearly along those three lines. The focus in this chapter will be on assuring product (or output) quality.

3. Quality assurance frameworks

7.7. Systematic data quality management in a statistical office typically takes the form of a quality assurance framework. Examples include the generic National Quality Assurance Framework (NQAF) template developed by the United Nations Statistics Division (UNSD) Expert Group on NQAF, the European Statistics Code of Practice, the International Monetary Fund (IMF) Data Quality Assessment Framework (DQAF), and the Statistics Canada Quality Assurance Framework, which are summarized below. Although these quality frameworks may differ from each other, they share most of the same common aspects and provide comprehensive structures for taking stock of quality concerns, activities, requirements and initiatives, and for fostering standardization and systematization of quality practices and measurement within statistical offices and across countries. [19]

7.8. In light of the benefits of having a quality assurance framework in place, in IRES, countries are encouraged to develop their own national quality assurance frameworks if they have not already done so. They are invited also to review the frameworks mentioned above and to consider either adopting one directly or structuring a framework of their own against one or some of them in a way that best fits their country's practices and circumstances. It is to be noted that the NQAF is based on and aligned with the other four frameworks.

7.9. Tables 7.1 through 7.4 present the template for a generic NQAF and summaries of the European Statistics Code of Practice, the IMF DQAF, and the Statistics Canada Quality Assurance Framework.

[19] A mapping of each of the above-mentioned frameworks to the NQAF Template, as well as an extensive collection of other quality documentation, are available on the Statistics Division NQAF website at http://unstats.un.org/unsd/dnss/QualityNQAF/nqaf.aspx.

Table 7.1:
Template for a generic National Quality Assurance Framework

Template for a generic National Quality Assurance Framework (NQAF)
(Developed by the United Nations Expert Group on NQAF)

1. Quality context
- 1a. Circumstances and key issues driving the need for quality management
- 1b. Benefits and challenges
- 1c. Relationship to other statistical agency policies, strategies and frameworks and evolution over time

2. Quality concepts and frameworks
- 2a. Concepts and terminology
- 2b. Mapping to existing frameworks

3. Quality assurance guidelines

3a. Managing the statistical system
- [NQAF 1] Coordinating the national statistical system
- [NQAF 2] Managing relationships with data users and data providers
- [NQAF 3] Managing statistical standards

3b. Managing the institutional environment
- [NQAF 4] Assuring professional independence
- [NQAF 5] Assuring impartiality and objectivity
- [NQAF 6] Assuring transparency
- [NQAF 7] Assuring statistical confidentiality and security
- [NQAF 8] Assuring the quality commitment
- [NQAF 9] Assuring adequacy of resources

3c. Managing statistical processes
- [NQAF 10] Assuring methodological soundness
- [NQAF 11] Assuring cost-effectiveness
- [NQAF 12] Assuring soundness of implementation
- [NQAF 13] Managing the respondent burden

3d. Managing statistical outputs
- [NQAF 14] Assuring relevance
- [NQAF 15] Assuring accuracy and reliability
- [NQAF 16] Assuring timeliness and punctuality
- [NQAF 17] Assuring accessibility and clarity
- [NQAF 18] Assuring coherence and comparability
- [NQAF 19] Managing metadata

4. Quality assessment and reporting
- 4a. Measuring product and process quality – use of quality indicators, quality targets and process variables and descriptions
- 4b. Communicating about quality – quality reports
- 4c. Obtaining feedback from users
- 4d. Conducting assessments; labelling and certification
- 4e. Assuring continuous quality improvement

5. Quality and other management frameworks
- 5a. Performance management
- 5b. Resource management
- 5c. Ethical standards
- 5d. Continuous improvement
- 5e. Governance

In the NQAF, lines 1–19 in section 3, each has a set of indicators or "elements to be assured". See http://unstats.un.org/unsd/dnss/docs-nqaf/2012-13-NQAF-E.pdf and the NQAF Checklist.

Table 7.2:
European Statistics Code of Practice – 15 Principles

Institutional Environment: Institutional and organizational factors have a significant influence on the effectiveness and credibility of a statistical authority developing, producing and disseminating European statistics. The relevant issues are professional independence, mandate for data collection, adequacy of resources, quality commitment, statistical confidentiality, impartiality and objectivity.

1. *Professional independence:* Professional independence of statistical authorities from other policy, regulatory or administrative departments and bodies, as well as from private sector operators, ensures the credibility of European statistics.
2. *Mandate for data collection*: Statistical authorities have a clear legal mandate to collect information for European statistical purposes. Administrations, enterprises and households, and the public at large may be compelled by law to allow access to, or deliver data for, European statistical purposes at the request of statistical authorities.
3. *Adequacy of resources*: The resources available to statistical authorities are sufficient to meet European statistics requirements.
4. *Quality commitment*: Statistical authorities are committed to quality. They systematically and regularly identify strengths and weaknesses to continuously improve process and product quality.
5. *Statistical confidentiality*: The privacy of data providers (households, enterprises, administrations and other respondents), the confidentiality of the information they provide and its use only for statistical purposes are absolutely guaranteed.
6. *Impartiality and objectivity*: Statistical authorities develop, produce and disseminate European statistics, respecting scientific independence and in an objective, professional and transparent manner in which all users are treated equitably.

Statistical Processes: European and other international standards, guidelines and good practices are fully observed in the processes used by the statistical authorities to organize, collect, process and disseminate European statistics. The credibility of the statistics is enhanced by a reputation for good management and efficiency. The relevant aspects are sound methodology, appropriate statistical procedures, non-excessive burden on respondents and cost effectiveness.

7. *Sound methodology*: Sound methodology underpins quality statistics. This requires adequate tools, procedures and expertise.
8. *Appropriate statistical procedures*: Appropriate statistical procedures, implemented from data collection to data validation, underpin quality statistics.
9. *Non-excessive burden on respondents*: The reporting burden is proportionate to the needs of the users and is not excessive for respondents. The statistical authorities monitor the response burden and set targets for its reduction over time.
10. *Cost effectiveness*: Resources are used effectively.

Statistical Output: Available statistics meet users' needs. Statistics comply with European quality standards and serve the needs of European institutions, governments, research institutions, business concerns and the public generally. The important issues concern the extent to which the statistics are relevant, accurate and reliable, timely, coherent, comparable across regions and countries, and readily accessible by users.

11. *Relevance*: European statistics meet the needs of users.
12. *Accuracy and reliability*: European statistics accurately and reliably portray reality.
13. *Timeliness and punc*tuality: European statistics are released in a timely and punctual manner.
14. *Coherence and comparability*: European statistics are consistent internally over time and comparable between regions and countries; it is possible to combine and make joint use of related data from different sources.
15. *Accessibility and clarity*: European statistics are presented in a clear and understandable form, released in a suitable and convenient manner, available and accessible on an impartial basis with supporting metadata and guidance.

For each of the Code of Practice's 15 principles, there is a set of indicators of good practice.

See http://ec.europa.eu/eurostat/documents/3859598/5921861/KS-32-11-955-EN.PDF/5fa1ebc6-90bb-43fa-888f-dde032471e15 and http://ec.europa.eu/eurostat/documents/3859598/5923349/QAF_2012-EN.PDF/fcdf3c44-8ab8-41b8-9fd0-91bd1299e3ef?version=1.0.

Table 7.3:
The IMF Data Quality Assessment Framework (DQAF), dimensions and elements

0. Prerequisites of quality

0.1 *Legal and institutional environment* – The environment is supportive of statistics.

0.2 *Resources* – Resources are commensurate with needs of statistical programmes.

0.3 *Relevance* – Statistics cover relevant information on the subject field.

0.4 *Other quality management* – Quality is a cornerstone of statistical work.

1. Assurances of integrity: *The principle of objectivity in the collection, processing, and dissemination of statistics is firmly adhered to.*

1.1 *Institutional integrity* – Statistical policies and practices are guided by professional principles.

1.2 *Transparency* – Statistical policies and practices are transparent.

1.3 *Ethical standards* – Policies and practices are guided by ethical standards.

2. Methodological soundness: *The methodological basis for the statistics follows internationally accepted standards, guidelines, or good practices*.

2.1 *Concepts and definitions* – Concepts and definitions used are in accord with internationally accepted statistical frameworks.

2.2 *Scope* – The scope is in accord with internationally accepted standards, guidelines, or good practices.

2.3 *Classification/sectorization* – Classification and sectorization systems are in accord with internationally accepted standards, guidelines, or good practices.

2.4 *Basis for recording* – Flows and stocks are valued and recorded according to internationally accepted standards, guidelines, or good practices.

3. Accuracy and reliability: *Source data and statistical techniques are sound and statistical outputs sufficiently portray reality.*

3.1 *Source data* – Source data available provide an adequate basis to compile statistics.

3.2 *Assessment of source data* – Source data are regularly assessed.

3.3 *Statistical techniques* – Statistical techniques employed conform to sound statistical procedures.

3.4 *Assessment and validation of intermediate data and statistical outputs* – Intermediate results and statistical outputs are regularly assessed and validated.

3.5 *Revision studies* – Revisions, as a gauge of reliability, are tracked and mined for the information they may provide.

4. Serviceability: *Statistics, with adequate periodicity and timeliness, are consistent and follow a predictable revisions policy.*

4.1 *Periodicity and timeliness* – Periodicity and timeliness follow internationally accepted dissemination standards.

4.2 *Consistency* – Statistics are consistent within the dataset over time and with major datasets.

4.3 *Revision policy and practice* – Data revisions follow a regular and publicized procedure.

5. Accessibility: *Data and metadata are easily available and assistance to users is adequate.*

5.1 *Data accessibility* – Statistics are presented in a clear and understandable manner, forms of dissemination are adequate, and statistics are made available on an impartial basis.

5.2 *Metadata accessibility* – Up-to-date and pertinent metadata are made available.

5.3 *Assistance to users* – Prompt and knowledgeable support service is available.

For the prerequisites of quality and the five dimensions, the DQAF identifies several elements of good practice, and several relevant indicators for each element. See http://dsbb.imf.org/images/pdfs/dqrs_Genframework.pdf

Table 7.4:
Statistics Canada Quality Assurance Framework

Introduction
- Current circumstances and key issues driving need for quality management
- Benefits of a quality assurance framework
- Relationship to other statistical office policies, strategies and frameworks

Quality concepts and instruments
- Existing quality policies, models, objectives and procedures
- Role of a quality assurance framework: where it fits into the quality toolkit

Quality assurance procedures
- Managing user and stakeholder relationships: user satisfaction surveys, feedback mechanisms and councils
- Coordinating the national statistical system: protocols and standards
- Managing relevance: programme review, planning processes and data analysis
- Managing accuracy: design, accuracy assessment, quality control and revision policy
- Managing timeliness: advanced release dates, preliminary and final releases
- Managing accessibility: product definition, dissemination practices and search facilities
- Managing interpretability: concepts, sources, methods and informing users of quality
- Managing coherence: standards, harmonized concepts and methods
- Managing output quality trade-offs: especially relevance, accuracy and timeliness
- Managing provider relationships: response burden measurement and reduction, and response rate maintenance
- Managing statistical infrastructure: standards, registers and policies
- Managing institutional infrastructure: confidentiality, security, transparency, professional independence, impartiality and objectivity
- Managing metadata: relating to quality

Quality assessment
- Quality indicators: defining, collecting, analysing
- Quality targets: setting and monitoring
- Quality assessment programme: self-assessment, peer review, quality audit, certification

Quality assessment
- Quality indicators: defining, collecting, analysing
- Quality targets: setting and monitoring
- Quality assessment programme: self-assessment, peer review, quality audit, certification

Quality and performance management and improvement
- Performance management: planning, cost and efficiency, sharing good practices, change management and risk management
- Recruitment and training: resource planning, determining recruitment and training needs and developing and conducting training courses
- Continuous improvement programme: quality culture, ongoing within operating budgets
- Governance structure: for quality and performance trade-offs and re-engineering initiatives based on results of quality assessments

Based on http://unstats.un.org/unsd/statcom/doc10/2010-2-NQAF-E.pdf, Annex 1.
See also www.statcan.gc.ca/pub/12-586-x/12-586-x2002001-eng.pdf

7.10. Countries that intend to develop a quality assurance framework can find guidance on the content and structure from the examples above and in the additional examples in table 7.5. The NQAF website http://unstats.un.org/unsd/dnss/QualityNQAF/nqaf_country_search.aspx can also be consulted for a more comprehensive listing of national and international quality references.

Table 7.5:
Examples of Data Quality Frameworks

Australia
Australian Bureau of Statistics (2009). *The ABS Data Quality Framework*. Canberra. www.abs.gov.au/ausstats/abs@.nsf/mf/1520.0
Canada
Statistics Canada (2009). *Statistics Canada Quality Guidelines*, Fifth Edition. Ottawa. www.statcan.gc.ca/pub/12-539-x/12-539-x2009001-eng.pdf
Statistics Canada (2002). *Statistics Canada's Quality Assurance Framework*. Ottawa. www.statcan.gc.ca/pub/12-586-x/12-586-x2002001-eng.pdf
Colombia
National Administrative Department of Statistics (DANE) (2011). *National Code of Good Practices for Official Statistics*. www.dane.gov.co/files/noticias/BuenasPracticas_en.pdf
Organisation for Economic Co-operation and Development (OECD)
Organisation for Economic Co-operation and Development (2011). *Quality Framework and Guidelines for OECD Statistical Activities*. Paris. http://search.oecd.org/officialdocuments/displaydocumentpdf/?cote=std/qfs(2011)1&doclanguage=en
South Africa
Statistics South Africa (2010). *South African Statistical Quality Assessment Framework* (SASQAF), Second Edition. Pretoria. http://beta2.statssa.gov.za/standardisation/SASQAF_Edition_2.pdf.

C. Dimensions of quality

7.11. It is widely recognized that the concept of quality in relation to statistical information is multidimensional or multifaceted; there is no single measure of data quality and no longer is accuracy thought to be the one absolute measure or indicator of high-quality data. As mentioned above, data outputs are typically described in the various quality assurance frameworks in terms of dimensions or components of quality. The dimensions are assessed, measured, reported on and monitored over time to provide an indication of output quality to both the data users and data producers.

7.12. The following dimensions of quality reflect a broad perspective and have been incorporated in most of the existing frameworks: relevance; accuracy; reliability; timeliness; punctuality; accessibility; clarity; coherence; and comparability.[20] In the European Statistics Code of Practice, for example, they are presented as principles 11–15, and in the NQAF, they make up NQAF lines 14–19; all of them are grouped in a separate section on managing statistical outputs.

7.13. The dimensions of quality are overlapping and interrelated and, therefore, the adequate management of each of them is essential if the information produced is to be fit for use. They should be taken into account when describing, measuring and reporting on the quality of statistics in general, and energy statistics in particular.

(a) *Relevance*: The relevance of statistical information reflects the degree to which the information meets or satisfies the current and/or emerging needs of key users. Relevance therefore refers to whether the statistics needed are produced and whether those produced are in fact needed and shed light on the issues of most importance to users; to know this requires the identification of user groups and knowledge about their various data needs and expectations. Relevance also covers methodological soundness, particularly the extent to which the concepts, definitions and classifications correspond to those users

[20] Some frameworks also include other dimensions, for example, interpretability (which is similar to clarity), credibility, integrity, serviceability, etc., but this chapter focuses only on relevance; accuracy; reliability; timeliness; punctuality; accessibility; clarity; coherence; and comparability.

require. Relevance can be seen as having the following three components: completeness; user needs; and user satisfaction.

One of the challenges of an energy statistics programme is weighing and balancing the conflicting needs of its current and potential users in order to produce energy statistics that satisfy the most important needs of key users in terms of the data's content, coverage, timeliness, and other factors, within resource constraints. To ensure or manage relevance, data compilers must engage with their users and data providers before and during the production process and after the outputs have been released. Strategies for measuring the relevance of an energy programme's outputs include consulting directly with key users about their needs, priorities and views on any deficiencies in the programme, tracking requests from users and evaluating the ability of the programme to respond, and analysing the results of user satisfaction surveys. Also, since needs evolve over time, ongoing statistical programmes should be regularly reviewed to ensure their continued relevance.

(b) *Accuracy and reliability*: The accuracy of statistical information reflects the degree to which the information correctly estimates or describes the phenomena it was designed to measure, that is, the degree of closeness of estimates to true values. It has many facets and there is no single overall measure of accuracy. It is usually characterized in terms of the errors in statistical estimates and is traditionally decomposed into bias (systematic error) and variance (random error) components. In the case of energy estimates based on data from sample surveys, accuracy can be measured using the following indicators: coverage rates, sampling errors, non-response errors, response errors, processing errors, and measurement and model assumption errors.

> As discussed in chapter VI on commodity and energy balances and in IRES, one measure of relevance for energy statistics would be to determine whether sufficient data have been collected to construct an energy balance, with the most relevant product categories and transformation and consumption broken down sufficiently. The ability to construct a detailed energy balance that complies with IRES principles should be considered automatically relevant for all countries.

Reliability is an aspect of accuracy and concerns whether the statistics consistently measure over time the reality they are designed to represent. The regular monitoring of the nature and extent of revisions to energy statistics is considered a gauge of reliability.

(c) *Timeliness and punctuality*: Timeliness of information refers to the length of time between the end of the reference period to which the information relates and its availability to users. Timeliness targets are derived from considerations of relevance, in particular the period for which the information remains useful for its main purposes. This varies with the rate of change of the phenomena being measured, the frequency of measurement, and the immedi-

> As data on oil products are often collected on both monthly and annual bases, each temporal data set can be used to assess the accuracy of the other. In addition, the size of the statistical difference can be considered an indicator of accuracy.

acy of user response to the latest data. Planned timeliness is a design decision and is often based on a trade-off between accuracy and cost. Thus, improved timeliness is not an unconditional objective. However, timeliness should be monitored over time to provide a warning of deterioration, especially as users' expectations of timeliness are likely to heighten as they experience faster and faster service delivery, thanks to the impact of technology. Punctuality refers to whether data are delivered on the dates promised, advertised or announced (e.g. in an official release calendar).

Mechanisms for managing timeliness and punctuality include announcing release dates well in advance, implementing follow-up procedures with data providers if they have not responded by the specified deadlines, releasing preliminary data followed by revised and/or final figures, making the best use of modern technology and adhering to the pre-announced release schedules (and if necessary, informing users of any divergences from the advance release calendar and the reasons for the delays). Paying attention to timeliness and punctuality and announcing schedules and release dates in advance help users plan, provide internal discipline and guarantee equal access to all by undermining any potential effort by interested parties to influence or delay any particular release for their own benefit.

(d) *Coherence and comparability*: The coherence of energy statistics reflects the degree to which the data are logically connected and mutually consistent, that is, the degree to which they can be successfully brought together with other statistical information within a broad analytic framework and over time. Comparability is a measurement of the impact of differences in applied statistical concepts, measurement tools and procedures, when statistics are compared between geographical areas or over time. The use of standard concepts, definitions, classifications and target populations promotes coherence and comparability, as does the use of a common methodology across surveys. The coherence and comparability concepts can be broken down into coherence within a dataset (internal coherence), coherence across datasets, and comparability over time and across countries.

Mechanisms for managing the coherence and comparability of energy statistics include adherence to the methodological basis of IRES recommendations when data items are compiled, and the promotion of cooperation and the exchange of knowledge between individual statistical programmes. Automated processes and methods, such as coding tools, can be used to identify issues and promote coherence and consistency within a dataset. The use of common concepts, definitions, classifications and methodology will result in coherence across datasets (e.g. between energy and other statistics, such as economic and environmental), and comparability over time and across countries. Divergences from the recommendations and common concepts, definitions, classifications and methodology, as well as breaks in series resulting from changes in the concepts, definitions, etc. should be explained.

Care should be taken when combining energy statistics with national accounts data such as GDP, because of the notable difference between energy statistics collected on the territory principle and national accounts compiled on the residency principle. However, in practice, this is not a huge difference and thus not a major concern for most countries.

(e) *Accessibility and clarity*: Accessibility of information refers to the ease with which users can learn of its existence, and locate and import it into their own working environment. It includes the suitability of the form or medium through which the information can be accessed and its cost. An advance release calendar or timetable to inform users in advance about when and where the data will be available and how to access them promotes accessibility and also enables equal access to information for all groups of users. A provision for allowing access to microdata for research purposes, in accordance with an established policy that ensures statistical confidentiality, also promotes accessibility.

Clarity refers to the extent to which easily comprehensible metadata are available, that is, where the metadata are necessary to give a full understanding of the statistical data. It is sometimes referred to as "interpretability". The clarity dimension is fulfilled by the existence of user support services and the provision of supplementary explanatory information and metadata, which are necessary for the proper understanding of the statistics and the appropriate uses to which they can be put. This information should normally cover the underlying concepts and definitions, origins of the data, the variables and classifications used, the methodology of data collection and processing, and indications of the quality of the statistical information. User feedback is the best way to assess the clarity of data from the user's perspective, for example, through questions regarding their understanding and interpretation in user satisfaction surveys.

Three items of metadata promoting clarity for an energy balance are: (a) the calorific values used to convert data from physical units; (b) whether the balance was compiled using gross or net calorific values; and (c) if the physical energy content method or partial substitution method was used.

7.14. The dimensions of quality described above are interconnected and as such are involved in a complex relationship. The accuracy/timeliness trade-off is probably the most frequently occurring and most important of the trade-offs. For example, striving for improvements in timeliness by reducing collection and processing time may reduce accuracy. A similar situation requiring consideration by energy statistics programmes, for example, would be the trade-off between aiming for the most accurate estimation of the total annual energy production or consumption by all potential producers and consumers, and providing this information in a timely manner when it is still of interest to users. IRES therefore recommends that if, while compiling a particular energy statistics dataset, countries are not in a position to meet the accuracy and timeliness requirements simultaneously, they must produce provisional estimates, which are to be made available soon after the end of the reference period but which will be based on less comprehensive data content. These estimates are to be supplemented at a later date with information based on more comprehensive data. In such cases, the tracking of the size and direction of revisions can serve to assess the appropriateness of the chosen timeliness-accuracy trade-off. Additional trade-offs, such as those between relevance and comparability over time, may need to be dealt with when changes in classifications to improve relevance lead to reductions in comparability over time due to breaks in series.

7.15. The trade-offs described above relate to those between two dimensions of output quality, but other conflicting situations may emerge requiring difficult trade-offs, such as those between one of the dimensions and such quality considerations as respondent burden, confidentiality, transparency, security or costs. For example, ensuring the efficiency or cost-effectiveness of a statistical programme may create challenges for ensuring relevance by limiting the programme's flexibility to address important gaps and deficiencies. A careful examination of all relevant factors and priorities will be required to make the necessary decisions relating to these types of difficult trade-offs.

> Monthly or quarterly oil and gas data, focusing on the principal components of supply and not attempting significant consumption or transformation breakdowns, can also be considered a practical application of the trade-off between timeliness and relevance, assuming that complete data are independently collected annually.

D. Measuring and reporting on the quality of statistical outputs

1. Quality measures and indicators

7.16. There are essentially two ways to measure quality: using quality measures and quality indicators. Quantitative and qualitative quality measures and indicators developed around dimensions, such as those described above (i.e. relevance, accuracy, reliability, timeliness, punctuality, accessibility, clarity, coherence and comparability), enable data compilers to describe, measure, assess, monitor and report output quality to assist users in determining whether the outputs meet their needs and are fit for use or for the purposes of their intended use. The measures and indicators can also be used by data compilers to monitor data quality with a view to continuous improvement.

7.17. *Quality measures*: Quality measures are defined as those items that directly measure a particular aspect of quality. For example, the time lag between the reference date and the day of publication of particular energy statistics, measured in the number of days, weeks or months, is a direct quality measure of timeliness. In practice, many other quality measures can be difficult and costly to calculate. In these cases, quality indicators can be used to supplement or act as substitutes for the desired quality measures.

7.18. *Quality indicators:* Quality indicators usually consist of information that is a by-product of the statistical process. They do not measure quality directly but can provide enough information to give an insight into quality. For example, in the case of accuracy, it is very challenging to measure non-response bias since the characteristics of those who do not respond can be difficult and costly to ascertain. In this instance, response rates are often utilized as a proxy to provide a quality indicator of the possible extent of non-response bias. Other data sources can also serve as a quality indicator to validate or confront the data. For example, in energy balances, energy consumption data can be compared with production figures to flag potential problem areas.

Examples and selection of quality measures and indicators

7.19. Examples of quality indicators and measures that have already been defined around specific dimensions and are in use by statistical organizations are numerous. Most are presented as descriptive statements or assertions (e.g. the majority of the indicators of good practice relating to the European Statistics Code of Practice's principles; the "elements to be assured" in the NQAF Guidelines and NQAF Checklist, and those relating to the DQAF's "elements of good practice"). Others may be quantitative statements or quantified measures calculated according to specific formulas (e.g. the European Statistical System (ESS) Standard Quality and Performance Indicators). The various quality indicators and measures are intended to make the description of a product by quality dimensions more informative and increase transparency. IRES encourages countries to develop or identify a set of quality measures and indicators that can be used to describe, measure, assess, document and monitor over time the quality of their energy statistics outputs and make them available to users. Table 7.6 presents a limited and domain-specific sample of some indicators and measures that countries' energy statistics programmes can consider using to indicate the quality of their energy statistics.

7.20. The objective of quality measurement is to have a practical set (limited number) of quality measures and indicators that can be used to describe and monitor over time the quality of the data produced by the responsible agencies and to ensure that users are provided with a useful summary of overall quality, while not overburdening respondents with demands for unrealistic amounts of metadata. As such, it is not intended that all quality measures and indicators be addressed for all data. Instead, as mentioned in IRES, countries are encouraged to select practical sets of measures and indicators to compile those that are most relevant to their specific outputs and can be used to describe and monitor the quality of the data over time. They should also ensure that the selected measures and indicators cover each of the quality dimensions that describe their outputs, have well-established methodologies for their compilation, and are easy to interpret by both internal and external users.

7.21. Data compilers should decide how frequently the measures or indicators for different key outputs are to be produced. Certain types of quality measures and indicators can be produced for each data item in line with the frequency of production or publication of the data; for example, response rates for total energy production can be calculated and disseminated with each new estimate. However, other measures could be produced once for longer periods and only be produced again for newly released data if there were major changes.

Table 7.6:
Selected indicators for measuring the quality of energy statistics[21]

Quality dimension	*Quality measure/indicator*
Relevance	• Procedures are in place to identify the users of energy data and consult with them about their needs. • Unmet user needs – Gaps between key user (e.g. energy policymakers, emissions inventory compilers, environmental statisticians) needs and compiled energy statistics in terms of concepts, coverage and detail are identified and addressed. • Requests for energy information are monitored and the capacity to respond is evaluated. • User satisfaction surveys on the agency's energy statistics outputs are regularly conducted and the results are analysed and acted upon.

[21] The indicators listed represent only a sample of possible indicators that can be used for measuring quality.

Accuracy and reliability	• Energy source data are systematically assessed and validated. • Sampling errors of estimates, e.g. standard errors, are measured, evaluated and systematically documented. • Non-sampling errors, e.g. item non-response rates and unit non-response rates, are measured, evaluated and systematically documented. • Coverage – The proportion of the population covered by the energy data collected is assessed (here the "population" refers to all entities either producing, trading, transforming or consuming energy, and is not necessarily the same as the common understanding of the term "population"). • The imputation rate is reported. • Information on the size and direction of revisions to energy data is provided and made known publicly.
Timeliness and punctuality	• A published release calendar announces in advance the dates that (key) energy statistics are to be released. • The time lag between the end of the reference period and the date of the first release (or the release of final results) of energy data is monitored and reported. • The possibility and usefulness of releasing preliminary data are regularly considered, while at the same time taking into account the data's accuracy. • The time lag between the date of the release or publication of the data and the date on which they were announced or promised to be released is monitored and reported. • Any divergences from pre-announced release times for the energy data are published in advance; a new release time is then announced with explanations on the reasons for the delays.
Coherence and comparability	• Comparison and joint use of related energy data from different sources is made. • Energy statistics are comparable over a reasonable period of time. • Divergences from the relevant international statistical standards in concepts and measurement procedures used in the collection/compilation of energy statistics are monitored and explained. • Energy statistics are internally coherent and consistent.
Accessibility and clarity	• Energy statistics and the corresponding metadata are presented in a form that facilitates proper interpretation and meaningful comparisons, and are archived. • Modern information and communication technology (ICT) is mainly used for disseminating energy statistics; traditional hard copy and other services are provided, when required, to ensure that users have appropriate access to the statistics they need. • An information or user support service, call centre or hotline, is available for handling requests for energy data and for providing answers to questions about statistical results, metadata, etc. • Access to energy microdata is allowed for research purposes, subject to specific rules and protocols on statistical confidentiality. • The regular production of up-to-date quality reports and methodological documents (on energy concepts, definitions, scope, classifications, basis of recording, data sources (including the use of administrative data), compilation methods, statistical techniques, etc.) is part of the work programme, and the reports and documents are made known publicly.

2. Quality reporting

7.22. In order for the users of energy statistics to be able to make informed use of the statistical information provided, they need to know whether the data are of sufficient quality. For some dimensions of quality, such as timeliness, users are able to easily assess the quality for themselves, while others, such as coherence and even relevance, may not be as obvious. The dimension of accuracy, in particular, is one which users may often have no way of assessing and must rely on the statistical agency for guidance. A quality report, or similar documentation, is meant to provide this guidance.

7.23. National practices for reporting on the quality of outputs vary. The quality documentation provided by data producers can range from short and concise to very detailed, depending upon the users for whom the information is meant. General users will most likely only be interested in a certain level of detail necessary for knowing whether they can rely on the data, while producers will want more detailed information in order to be able to evaluate whether the output quality meets the quality requirements and to identify strengths and any areas that need further improvement.

7.24. Quality information is often structured in a template format to promote comparability and consistency across statistical domains. Sometimes it is issued in a quality report that is separate from other metadata, not as a replacement for, but rather as a complement to them. Other times, it may be included as part of other metadata (for example, along with explanatory and technical notes and other more detailed documentation) provided by the compiling agency. Some compilers refer to the quality documentation as a quality statement or quality declaration. Typically, though, the quality reports or quality documentation examine and describe quality according to the dimensions the agency used to define its products' fitness for purpose, such as those focused upon in this chapter: relevance, accuracy, reliability, timeliness, punctuality, coherence, comparability, accessibility and clarity.

7.25. Two types of quality reports can be distinguished – the shorter "user-oriented" one and the more detailed "producer-oriented" one. The focus of the user-oriented reports is on output quality, so they are often limited to brief descriptions of the output dimensions, and generally will include just a few of the indicators listed in the previous section on measuring quality. On the other hand, the longer, "producer-oriented" quality reports, such as the comprehensive type ESS members are recommended to produce periodically (every five years or so or after major changes), go into greater detail on the dimensions, especially on errors and other aspects affecting accuracy, and provide additional information on the processes and other issues, such as confidentiality, costs and the response burden. For users, such details may be confusing and unnecessary for their purposes, but for the producers, the comprehensive reports serve as an internal self-assessment. Quality reporting therefore underpins quality assessment, which in turn is the starting point for quality improvements in statistical programmes.

7.26. An extract of a short, user-oriented quality report for energy statistics surveys in Slovenia is presented in box 7.1.

Box 7.1:
Country example, Slovenia: Quality Report on Annual Energy Statistics[22]

Brief description of the survey

The aim of the statistical survey – Annual energy statistics (E3-TOP/L) – is to gather data on the production (by kind of fuel), purchase and sale of heat, the consumption of fuels to produce heat and the use of electricity in heat plants.

Timeliness of the first results for the annual survey

Reference period	31.12.2009	31.12.2010	31.12.2011	31.12.2012	31.12.2013
Date of publication	4.10.2010	4.10.2011	2.10.2012	8.10.2013	9.10.2014
Time lag (days)	T+277	T+277	T+276	T+281	T+282

22 Source: Statistical Office of the Republic of Slovenia, www.stat.si/doc/metodologija/kakovost/24_LPK_E3TOPL_E3TOPM_2013_ang.pdf.

Box 7.1 (continued)

Burden of the reporting units[a]

	Annual survey E3-TOP/L	Monthly survey E3-TOP/M
Number of reporting units that submitted data	41	9
Annual number of questionnaires per unit	1	12
Time spent to fill in a questionnaire (hours)	0.50	0.33
Total time spent (hours)	20.5	36

[a] The burden of the reporting units was estimated by multiplying the time spent to fill in a questionnaire by the number of reporting units.

Survey costs at the Statistical Officeb

	Annual survey E3-TOP/L	Monthly survey E3-TOP/M
Number of working hours spent	46	8
Number of reporting units that had to fill in questionnaires	43	9
Survey period	annual	monthly
Number of questionnaires per year (total)	43	108

[b] We took the number of working hours spent as the measure for determining the costs of implementing the E3-TOP surveys.

Comparability over time

The questionnaire for the annual survey was first introduced in the survey for the 2000 annual data (some comparable data were available already from 1995). These results are only part of the data required to elaborate the heat balance but form the majority. There is a complete time series for 2000 onwards, but, as already mentioned, some comparable data on production were available (from 1995). A comparable time series for the monthly survey is available from 2004 onwards.

7.27. For ESS countries, a standard reporting structure has been developed for presenting quality information in a comprehensive quality report, that is a full scale, producer-oriented report with quantitative and qualitative information dealing in detail with all important aspects of output and process quality. Box 7.2 presents two examples of what to include in quality reports on two specific elements of survey design relating to accuracy and reliability, namely sampling errors and non-response errors.

Box 7.2:
Selected suggested content to be covered in ESS quality reports (for selected elements under the accuracy and reliability dimension)

What should be included in sampling errors

Always applicable:

- Where sampling is used, there should be a section on sampling errors.
- As far as possible sampling error should be presented for estimates of change in addition to estimates of level. If necessary, reasonable assumptions can be used.

Box 7.2 (continued)

If probability sampling is used:

- There should be a presentation of sampling errors calculated according to formulas that should also be made available. If the estimators include adjustments for non-sampling errors, for example non-response, this should be explained and included also in the accuracy assessment.
- The most appropriate presentational device should be chosen, normally coefficients of variation (CVs,) ranges of CVs, or confidence intervals.
- If outliers have received special treatment in estimation, this must be clearly described.

If non-probability sampling is used:

- For sampling with cut-off, an assessment of the accuracy due to the cut-off procedure should be included, in addition to the presentation of sampling error for the sampled portion of the population.
- For other forms of non-probability, a sampling model can be invoked for the estimation of sampling error. A motivation for the chosen model and a discussion of sampling bias should be included.

What should be included in non-response errors

- Non-response rates according to the most relevant definitions for the whole survey and for important sub-domains.
- Item non-response rates for key variables.
- A breakdown of non-respondents according to cause for non-response.
- A qualitative statement on the bias risks associated with non-response.
- Measures to reduce non-response.
- Technical treatment of non-response at the estimation stage.

Source: ESS Handbook for Quality Reports 2013, www.gov.uk/government/uploads/system/uploads/attachment_data/file/239316/decc_statistics_user_survey_2012.pdf.

7.28. While the guidelines of the *ESS Handbook* presented above were developed specifically for ESS Member States for the preparation of a comprehensive, full-scale, producer-oriented quality report, energy statistics programmes in non-ESS countries could draw upon the extensive work done in the ESS to set up their own quality reporting processes. They may find it useful to apply the ESS guidelines as a starting point for developing guidelines for their own energy programmes' quality reports. Another helpful ESS resource for both quality reporting, as well as for providing comprehensive metadata to users, is the Single Integrated Metadata Structure (SIMS), an inventory of statistical concepts used for quality and metadata reporting in the ESS (See box 7.6 in the metadata section below).

7.29. The preparation and updating of quality reports depend on the survey frequency and the stability of the quality characteristics. A balance should be sought between the need for the most recent information and the report compiling burden. If necessary, the quality report should be updated as frequently as the survey is carried out. However, if the characteristics are stable, the inclusion of the quality indicators in the latest survey results could be enough to update the report. Another option is to provide a detailed quality report less frequently, and a shorter one after each survey, covering only the updated characteristics, such as some of the accuracy-related indicators. In IRES, countries are encouraged to regularly issue quality reports as part of their metadata.

E. Quality reviews

7.30. Quality reviews can be done in the form of self-assessments, audits or peer reviews. They can be undertaken by internal or external experts and the time frame can vary from days to months, depending on the scope of the review. However, the results are more or less identical – the identification of improvement actions/opportunities in processes and products. In IRES, it is recommended that some form of quality review of energy statistics programmes be undertaken periodically, for example, every four to five years, or more frequently if significant methodological or other changes in the data sources occur.

7.31. *Self-assessments*: Self-assessments are comprehensive, systematic and regular reviews of an organization's activities and results are referenced against a model or framework. They are "do it yourself" evaluations. Typically, self-assessment checklists or questionnaires are developed to be used for systematic assessment of the quality of the statistical production processes.[23]

7.32. A *quality audit*: A quality audit is a systematic, independent and documented process for obtaining quality evidence concerning the quality of a statistical process and evaluating it objectively to determine the extent to which policies, procedures and requirements on quality are fulfilled. In contrast to self-assessments, audits are always carried out by a third party (internal or external to the organization). Internal audits are conducted with the purpose of reviewing the quality system in place (policies, standards, procedures and methods) and the internal objectives. They are led by a team of internal quality auditors who are not in charge of the process or product under review. External audits are conducted either by stakeholders or other parties that have an interest in the organization, by an independent external auditing organization, or by a suitably qualified expert.

7.33. *Peer reviews*: Peer reviews are a type of external audit which aim to assess a statistical process at a higher level, and not to check conformity with requirements item by item from a detailed checklist. They are therefore often more informal and less structured than external audits. Normally, peer reviews do not address specific aspects of data quality but focus rather on broader organizational and strategic questions. They are typically systematic examinations and assessments of the performance of one organization by another, with the ultimate goal of helping the organization under review to comply with established standards and principles, improve its policymaking and adopt best practices. The assessments are conducted on a non-adversarial basis, and rely heavily on mutual trust between the organization and assessors involved, as well as their shared confidence in the process.

F. Data quality and validation checks for international energy statistics at IEA

7.34. The International Energy Agency (IEA) has implemented a variety of data quality and validation checks for annual energy statistics and balances which tie in with the coherence, comparability and accuracy dimensions discussed above, and could be relevant for national administrations wishing to validate their own commodity statistics and balances. The checks are grouped around three main categories: internal consistency, external consistency and plausibility.

[23] For more information, see for example, the European Statistical System's Development of a Self-Assessment Programme (DESAP) and Data Quality Assessment and Tools, http://ec.europa.eu/eurostat/documents/64157/4373903/07-Checklist-for-Survey-Managers_DESAP-EN.pdf/ec76e3a3-46b5-409e-a7c3-52305d05bd42).

Box 7.3:
Examples of data quality/validation checks from IEA

Internal consistency

For individual energy commodities:

- Time series checks, automated as well as visually supported, help identify outliers or missing data and assess whether trends and growth rates look reasonable.
- Revisions over time should be justified by changes in methodologies or availability of more accurate information; the explanations should be documented and provided to users.
- Numbers should not be negative, except when logically possible (e.g. stock changes, transfers, statistical differences).
- Subtotals should add up to totals, both for products (e.g. individual oil products to the total oil products category) and for flows (e.g. industrial subsectors to total industry, energy and non-energy use to total final consumption, etc.).
- Production should not generally be lower than exports, except in very specific situations (e.g. large stock withdrawals due to large storage, high levels of imports, etc.).
- Total imports and exports by origin and destination (if shown) should add up to total imports and exports. If the detailed information is available, checks could be made whether the trading partners reported consistent quantities.
- Opening stocks should equal closing stocks of the previous time period and stock changes should equal the difference between opening and closing stock levels.
- Calorific values per fuel type should be reported wherever relevant, fall within given ranges, be consistent with data reported in physical and energy units, and vary over time in a reasonable way. (For information on typical net calorific value ranges for energy products, see table 4.1 in IRES.)

Across different energy commodities

- All inputs to electricity and heat generation shown for each of the various commodity statistics (e.g. oil, gas, renewables and coal) should match the values for such inputs shown in the electricity and heat statistics, fuel by fuel.
- All biofuels (e.g. biogasoline, biodiesel or biogas) reported as blended with conventional fuels in the renewables statistics should match the receipts from renewable sources in the respective commodity statistics (e.g. oil or gas for the examples above).
- For all transformations between commodities (e.g. coal to liquids, gas to liquids, etc.), data should be consistent across the various commodity statistics.
- Specific sectoral consumption flows could be checked across all fuel types, and unusual trends (e.g. fuel switching) should be justified.

Consistency with external data sources

- Data could be checked against other available national sources (publications of different ministries, statistical offices, energy suppliers, etc.).
- Data could be checked against publications of international organizations (e.g. IEA, United Nations Statistics Division, Eurostat, IAEA for nuclear, the Food and Agriculture Organization of the United Nations (FAO) for biofuel information, etc.).
- Data could be checked against publications by private sector organizations, including those with a sectoral energy industry focus (e.g. for natural gas statistics: Cedigaz, GIIGNL, Eurogas, etc.; and for oil, the *Oil and Gas Journal*).
- Annual data could be compared with monthly data, where available.

Box 7.3 (continued)

- Trends in energy consumption data could be checked against trends in relevant activity data (e.g. data on physical production of cement, pig iron, steel, electricity, etc.) to assess whether they look reasonable. Some national and international industrial associations could hold relevant data (e.g. the World Steel Association).

Plausibility

- The size of the statistical difference should be reasonable compared to supply, both in physical and in energy units.
- For all transformation processes, efficiencies, defined as output/input in energy units, should be within reasonable ranges depending on the technologies used, but, in any case, less than 100 per cent; variations in time should be justified by technical factors.[24] Transformation processes include electricity and heat generation (electricity only, combined heat and power (CHP) and heat only), and production in refineries, coke ovens, blast furnaces, charcoal plants, coal liquefaction plants, etc.
- For transformation processes with multiple outputs, such as coke ovens and refineries, the yields by fuel type and their variation in time would need to be within expected ranges, given the technologies used in the processes.
- For energy industries, such as electricity generation, refineries, etc., production should be consistent with a reasonable utilization rate of the existing capacity; own use and losses should be reasonable percentages of production.
- Based on data in energy units, a number of indicators, such as total energy production, total primary energy supply, energy self-sufficiency, final consumption, energy intensity (TES/GDP), as well as CO_2 emissions estimates, could be monitored over time to assess whether trends look reasonable and how values compare with those of other countries.

[24] Typical efficiency ranges in transformation processes are presented in chapter VI, tables 6.10 and 6.11.

G. Country practices – quality in energy statistics programmes

1. Country feedback about experience assuring output quality

7.35. As part of the preparation of this *Manual*, information on country practices was collected through the country practice template developed by the Oslo Group on Energy Statistics and the United Nations Statistics Division. Questions about the dimensions of quality, as described in IRES and in this *Manual*, were included in section 5 of the template. Countries were requested to share information on the quality dimensions, specifically their country practices on relevance and satisfying user needs; accuracy and calculating and handling errors; timeliness and punctuality, and managing timing of and deadlines for data releases; accessibility and clarity, and providing access to data and metadata; and ensuring and explaining the degrees of coherence, comparability and consistency. Below, in box 7.4, is a selection of excerpts and practices from country replies. The complete set of replies can be accessed at https://unstats.un.org/unsd/energy/template/responses_t.htm.

Box 7.4:
Examples of country practices for assuring quality in energy statistics: quality dimensions

Relevance

"Opinions of users on data quality and for which data the statistics should be conducted are learned in workshops and their suggestions are taken into account in practical work. Monitoring is conducted continuously on the website of the State Statistical Committee in order to learn users' opinion." (Azerbaijan)

"The statistics produced are highly relevant and are the subject of interest to all levels of government, industry and the general public. The information collected is of use to analysts and policymakers in both the public and private sectors for informed decision-making." (Australia)

Accuracy

"Every effort was made to minimise reporting error, by the careful design of questionnaires, intensive training of survey analysts, and efficient data processing procedures. Main non-sampling issues for EWES: Some businesses' utility payments, such as electricity, gas and water, are included in lease agreements and not able to be separately reported. Instructions on the form advise how to report in these situations. This issue results in overestimates in the rental, hiring and real estate services industry, balanced by underestimates across the rest of the economy. Selected data item's relative standard errors are shown on the ABS website." (Australia)

"Measurement errors can be considered negligible for most of the items covered (<5%) since the respondents keep good records of their dealings. Processing errors can also be considered to be minimal (<10%). Some of the errors are mainly due to allocations of energy items to end users for certain energy items. For instance, the use of fuel oil by the different manufacturing industries might not be exact as the data suppliers (petroleum companies) do not keep detailed databases by end users. Other sources of error: There might be some model assumption errors, such as while estimating the use of biomass and while allocating energy use by end users. These errors are minimized by cross checking results with different sources of information relevant to the item (benchmarking)." (Mauritius)

"Uncertainty is estimated using confidence intervals and variation coefficient." (Finland)

"Non-response errors are corrected by sending warning letters or charging penalty fees." (Hungary)

"Measurement errors may occur in the data due to misunderstandings or lack of knowledge. Most respondents are not energy experts and may find it difficult to answer all the questions. Thus, we try to guide them with user instructions and advise them if they contact us. We also use a great deal of resources on making optimal questionnaires." (Norway)

Timeliness and punctuality

"Preliminary results: on 100th day – information to the Bank of Latvia on main energy resource consumption by branch; on 120th day – to the Ministry of Economics and Ministry of Agriculture (preliminary heat balance and fuel consumed by transformation sector). Final results: on 210th day (informative leaflet "Energy Balance of Latvia"); on 270th day (Internet database of the Central Statistical Bureau); on 330th day (data transmission to Statistical Office of the European Union (Eurostat), International Energy Agency (IEA), United Nations (UN))." (Latvia)

"A specific day (15th day of every month)." (Japan)

Box 7.4 (continued)

"The reference period for the monthly electricity generator surveys is month minus 3, the frequency monthly. The reference period for the TSO survey is month minus 1, the frequency monthly. This ensures that all twelve months are received in time for a provisional Energy Balance which is normally published in March. The finalized Energy Balance is published in October." (Ireland)

Accessibility and clarity

"An advance release calendar is published on CYSTAT's website and also sent to the subscribers to the alert function of the website on the Friday prior to the week of publication, specifying the exact release date. On publication day, an email is again sent to the subscribers informing them of the day's releases and it is also published on the Home Page. Access to the data as well as to metadata is easy as there is a clickable link (both in the email and on the home page) which directs the reader to the exact webpage where the new data are published." (Cyprus)

"The results are published on the Internet through the predefined tables and graphs, as well as through interactive query in a data bank. In addition, the results are given in a compact disc. It also provides various 15 print products. In addition, the information is available at the National Institute of Statistics and Geography (INEGI) through its consultation centers and marketing, established in the various points of the national territory, or electronically via the Institute's web page." (Mexico)

"SORS has a publications calendar, metadata are easily available, all groups have access to the published statistics." (Serbia)

Coherence and comparability

"Coherence with statistics on renewable electricity have been [sic] guaranteed by specific agreement." (Italy)

"There is ensured comparability of the data presented in the Energy Balance with the similar data presented in the other statistical areas. For example, the production of the electricity, heating energy and other products are comparable with the data of the industrial production statistics, or data on import of energy products presented in the Energy Balance are comparable with the similar data on external trade." (Moldova)

"Comparability over other domains: Data are cross-checked with other sources, namely, External Merchandise Trade, Annual Survey on Economic Activities, etc." (Macao)

2. Country experience in a decentralized statistical system

7.36. Ensuring data quality is an important function of any statistical organization, whether it be centralized or decentralized. Below is a brief description of Sweden's decentralized statistical system and quality of official statistics.

Box 7.5:
Example of data quality in a decentralized system[25]

The Swedish official statistical system

The Swedish official statistical system has been decentralized since statistics reform was implemented on 1 July 1994. Statistics Sweden is responsible for cross-sectoral statistics, such as economic statistics and national accounts, while 26 sector-specific agencies are responsible for official statistics in each of their respective areas. The Swedish Energy Agency (SEA) is responsible for official energy statistics.

Quality of official statistics

The review of Sweden's official statistical system, accomplished by an inquiry – through contact with users and foreign agencies – indicated that current official statistics are of good quality. This does not mean that there are no problems, or that quality cannot be improved in certain respects. But the measures that may be required to improve quality essentially involve making a basically well-functioning system even better.

The most important quality requirements for Sweden's official statistics are stated in the *Official Statistics Act*. The review of Sweden's official statistical system proposed that the wording of the quality requirements in the *Act* should be modelled on the quality criteria in the European Union's statistics regulation. Much of the content of the European Statistics Code of Practice has already been met in Swedish law, although additional principles contained in the Code may also need to be regulated by law. The review proposed that the principle of non-excessive burden on respondents be introduced in Sweden's Official Statistics Ordinance. The suggestion is that most of the principles would be more suited as regulations by Statistics Sweden rather than general guidelines, as they are currently.

Statistical agencies, such as SEA, are working on quality issues in a professional manner, both individually and in joint forums, within the framework of the Council for Official Statistics in Sweden. Since good quality is crucial for the reliability and credibility of official statistics, it is essential that the agencies continue to work actively at ensuring data quality. The Council for Official Statistics has established certain guidelines and criteria to promote sufficient quality in official statistics. Based on them, statistical agencies can take an "official statistics pledge", which means that they promise to operate in accordance with the criteria. Currently, two agencies, the National Board of Health and Welfare and the Swedish Board of Agriculture, have taken the pledge. It would send an important signal to the users of statistics if more statistical agencies, and particularly Statistics Sweden, were to also take the pledge.

Unfortunately, errors occasionally occur in the statistics, but these do not seem to be due to fundamental system errors. They most likely arise because, in practice, it is impossible to completely avoid errors in such a complex system as the statistics system. When errors occur, it is important that the statistical agencies have procedures and routines in place to identify, correct and learn from them. It is also important that the agencies openly report the errors, so that commissioning organizations, users and others will be aware of the circumstances.

In Sweden, as well as in other countries, response rates for statistical surveys show a declining trend. This development is a real problem. The statistical agencies are aware of it and are trying to find different methods for dealing with it. The problem of declining response rates seems to be a fundamental trend due to structural causes. It is presently more difficult to reach people using traditional methods than it was in the past. Producers of statistics may need to develop better methods for managing the continuing decline and also find other ways for accessing the required information. If response rates continue to decline sharply, it may be necessary to consider the introduction of an obligation on the part of private individuals to provide information, as is the case in many other countries.

[25] The Swedish Official Statistical System, 28 January 2013.

Box 7.5 (continued)

According to some estimates, more than 95 per cent of official statistics are based on administrative data collected for purposes other than statistics. Problems with administrative data can arise, however, if the agency responsible for those data in question chooses to change or terminate their collection. This could be solved if the agencies responsible for registers were required to first consult with Statistics Sweden in such a situation. Statistics Sweden, in its role as coordinator of the statistics system, could be given the task of safeguarding the interests of statistical agencies. The Council for Official Statistics could probably be responsible for coordinating this. Furthermore, the obligation to consult should only concern those registers whose data are reasonably likely to be used for the production of statistics. However, the consultation obligation should not be introduced before the finalization of the European Union's statistics regulation.

H. Metadata on energy statistics

7.37. Types of statistical data include microdata, macrodata and metadata. Microdata are non-aggregated observations or measurements of characteristics of individual units, macrodata are data derived from microdata by grouping or aggregating them, and metadata are data that describe the microdata, macrodata or other metadata. This section of the chapter focuses on metadata.

7.38. Greater emphasis has been given over the years to the importance of ensuring that statistics published by national statistical offices, international organizations and other data producing agencies are accompanied by adequate metadata. Metadata, that is, "data about data" (and statistical metadata, or "data about statistical data") are a specific form of documentation that defines and describes data so that users can locate and understand them, make an informed assessment of their strengths and limitations and, therefore, their usefulness and relevance to their purposes, and use and share them. Without metadata, statistical data are just numbers.

7.39. Metadata, therefore, are important tools that support the production and final use of statistical information. The main types of metadata are structural metadata and reference metadata.

> *(a) Structural metadata*: Structural metadata are identifiers and descriptors of the data which are essential for discovering, organizing, retrieving and processing statistical datasets. They can be thought of as the "labels" that need to be associated with each data item in order for it to have any meaning at all, such as the names of the table columns, unit of measurement, time period, commodity code, etc. Structural metadata items are an integral part of the statistics database and should be extractable together with any given data item. If they are not associated with the data, it will be impossible to identify, retrieve and browse the data.
>
> (b) *Reference metadata*: Reference metadata describe the content and quality of the statistical data (e.g. conceptual metadata, describing the concepts used and their practical implementation; methodological metadata, describing methods used for the generation of the data; and quality metadata, describing the different quality dimensions of the resulting statistics, e.g. timeliness, accuracy, etc.). These reference metadata are often linked ("referenced") to the data but, unlike structural metadata, can be presented separately from the data on the Internet or in publications.

7.40. *Metadata items*. When disseminating comprehensive energy statistics, the compiling agency has the responsibility of making the corresponding metadata available and easily accessible to users. There are numerous metadata items that describe a statistical series, and many countries and organizations have developed metadata templates, lists or inventories for the presentation of the concepts, definitions, and descriptions of the methods used in the collection, compilation, transformation, revision, dissemination and evaluation, among other factors, of their statistics. One such comprehensive inventory is the Single Integrated Metadata Structure (SIMS) for metadata and quality reporting in the European Statistical System (ESS), whose methodological and quality metadata items are presented below in box 7.6. In practice, the amount of metadata detail that different countries disseminate along with their energy data varies, as does the way in which the metadata are presented. The basic purpose though is always the same – to help users understand the data and their strengths and limitations.

Box 7.6:
Metadata items for statistical releases (from SIMS)

Code	Survey/Product name
S.1	Contact (organization, contact person, address, email, phone, fax)
S.2	Introduction
S.3	Metadata (last certified, posted and updated)
S.4	Statistical presentation
S.4.1	Data description
S.4.2	Classification system
S.4.3	Sector coverage
S.4.4	Statistical concepts and definitions
S.4.5	Statistical unit
S.4.6	Statistical population
S.4.7	Reference area
S.4.8	Time coverage
S.4.9	Base period
S.5	Unit of measure
S.6	Reference period
S.7	Institutional mandate (legal acts and other agreements, data sharing)
S.8	Confidentiality (policy, data treatment)
S.9	Release policy (release calendar, calendar access, user access)
S.10	Frequency of dissemination
S.11	Dissemination format, accessibility and clarity (news release, publications, online database, microdata access, other),
S.12	Accessibility of documentation (documentation on methodology, quality documentation)
S.13	Quality management (quality assurance, quality assessment)
S.14	Relevance (user needs, user satisfaction, completeness)
S.15	Accuracy and reliability (overall accuracy, sampling error, non-sampling error (coverage, measurement, non-response, processing, model assumption errors))
S.16	Timeliness (time lag – final results) and punctuality (delivery and publication)

Box 7.6 (continued)

Code	Survey/Product name
S.17	Comparability (geographical, over time)
S.18	Coherence (cross domain, internal)
S.19	Cost and burden
S.20	Data revision (policy, practice)
S.21	Statistical processing
S.21.1	Source data
S.21.2	Frequency of data collection
S.21.3	Data collection
S.21.4	Data validation
S.21.5	Data compilation
S.21.6	Adjustments
S.21.61	Seasonal adjustment

Source: Technical Manual of the Single Integrated Metadata Structure, http://ec.europa.eu/eurostat/documents/64157/4373903/03-Single-Integrated-Metadata-Structure-and-its-Technical-Manual.pdf/6013a162-e8e2-4a8a-8219-83e3318cbb39.

Box 7.7

Country example, Indonesia: Metadata structure for energy statistics collected by BPS (Statistics Indonesia)

The Indonesian Central Bureau of Statistics (BPS) (Statistics Indonesia) is directly under the authority of the President of Indonesia to guarantee statistical independence. The scope of energy statistics collected by BPS covers electricity statistics, water supply statistics and city gas statistics. Data on fossil fuel production are collected by the Ministry of Energy.

Data compilation

Data source: Establishment surveys and administrative data.

Data conversion: The conversion of energy data from various measurement units to terajoules uses the standard conversion factors given in previous United Nations manuals.

Survey methodology

Sampling method: Complete enumeration (all companies are covered).

Population units: State enterprises producing energy.

Collection system: Data collection is carried out by direct interview, mail and fax.

Coverage: The survey covers all State enterprises in 33 provinces, in 645 regencies (municipalities).

Periodicity: The statistics are compiled both quarterly and annually.

Timeliness: From data collection to publication takes one year, hence an approximately 13-month publication lag.

Data characteristics collected: Name and address of companies, number of workers, main products produced, quantity and value of each product type; input value and output value (hence the value added derived); others.

Classification system: The classification system used is called Klasifikasi Baku Lapangan Usaha Indonesia, which is a national classification complying with ISIC, Revision 3.

Box 7.7: (continued)

Publication type: annual electricity statistics, annual city gas statistics, energy balance of Indonesia

Confidentiality: With the exception of company names and address, all individual company data are confidential. If a collated data point represents less than three companies it will be collated with a similar ISIC category.

Use of the data: Mostly for government purposes in designing, planning, monitoring, evaluating, and making policy, in addition to the needs of private establishments and international organizations.

(Based on http://unstats.un.org/unsd/energy/meetings/mexico2008/Presentations/Session%204%20-%20 Indonesia.pdf).

7.41. *Users and levels of metadata detail*. There are many types of users for any given set of data. The wide range of possible users, with their different needs and statistical expertise, means that a broad spectrum of metadata requirements has to be addressed. The responsible agencies, as data suppliers, must make sufficient metadata available to enable both the least and the most sophisticated users to interpret and readily assess the data and their quality. In IRES, it is recommended that different levels of metadata detail be made available to users to meet the requirements of the various user groups. An approach for presenting metadata is to organize them as if they were in layers, where the methodological information describing statistics becomes more detailed as one moves down from the narrower top layer (where the summary metadata are) to the deeper and wider layers. In this way, users will be able to dig as deeply as they want or need to get a more thorough understanding of the concepts and practices.

7.42. Box 7.8 provides links to examples of the main metadata items for energy statistics that several countries disseminate on their websites. As can be seen in these examples, the amount of metadata and their presentation vary among countries. Box 7.9 provides links to general metadata resources and guidelines.

Box 7.8:
Country examples of metadata on energy statistics

Canada

Statistics Canada. Annual Industrial Consumption of Energy Survey (ICE). Record number 5047. www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=5047

Statistics Canada. Oil and Gas Extraction. Record number 2178. www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=2178

Finland

Statistics Finland (2013). Energy supply and consumption. Helsinki.

www.stat.fi/meta/til/ehk_en.html

The Netherlands

Statistics Netherlands (2015). "Methods. Renewable Energy". The Hague.

www.cbs.nl/en-GB/menu/methoden/dataverzameling/duurzame-energie-ob.htm

Slovenia

Statistical Office of the Republic of Slovenia (2013). Annual energy statistics, Methodological Explanations. Ljubljana, 2013.

www.stat.si/StatWeb/File/DocSysFile/8345

7.43. *Use of metadata to promote international comparability*. Metadata provide a mechanism for comparing national practices in the compilation of statistics. This may help and encourage countries to implement international standards and to adopt best practices in the collection of data in particular areas. The use of standard terminology and definitions and better harmonization of approaches adopted by different countries will improve the general quality and coverage of key statistical indicators.

Box 7.9:
Metadata resources and guidelines

Australia

Australian Bureau of Statistics (2012). What is Metadata? February 2012.

www.abs.gov.au/websitedbs/a3121120.nsf/home/statistical+language+-+what+is+metadata

Australian Bureau of Statistics, National Statistical Service. *10 Principles for Managing Metadata.*

https://statisticaldataintegration.abs.gov.au/topics/data-management/providing-metadata

Finland

Statistics Finland (2005). Metadata guidelines, June 2005.

www.stat.fi/meta/index_en.html

OECD

Organisation for Economic Co-operation and Development (2007). *Data and metadata reporting and presentation handbook*. Paris.

www.oecd.org/std/37671574.pdf

United Nations

United Nations (2012). Metis Wiki, March 2013.

www1.unece.org/stat/platform/display/metis/METIS-wiki

United Nations (2012). *United Nations Common Metadata Framework*, March 2013. www1.unece.org/stat/platform/display/metis/The+Common+Metadata+Framework

United Nations (2009). *Statistical Metadata in a Corporate Context*. Geneva.

https://unece.org/DAM/stats/publications/CMF_PartA.pdf

7.44. *Metadata must be high priority in all countries*. In IRES, it is recommended that countries accord high priority to the development of metadata, keeping them up-to-date, and considering the dissemination of metadata as an integral part of the dissemination of energy statistics.

Chapter VIII: Data dissemination

A. Introduction

8.1. The first fundamental principle of official statistics states, inter alia, that "official statistics that meet the test of practical utility are to be compiled and made available on an impartial basis by official statistical agencies to honour citizens' entitlement to public information".[26] Dissemination is an activity for fulfilling this responsibility and refers to the provision to the public of statistical outputs containing data and related metadata. Energy data are usually disseminated by national responsible agencies in the form of various statistical tables or by providing access to relevant databases; but country practices differ significantly in their effectiveness and methods.

[26] See A/RES/68/261.

8.2. This chapter introduces recommended basic principles of data dissemination in a normal statistical system, with energy-specific examples wherever possible; it also outlines data dissemination practices which statistical units/offices can advisably adopt. In this chapter, seven key dimensions of an effective data dissemination system will be discussed: reference period; dissemination schedule; confidentiality; revision policy; dissemination format and access; and meeting user needs. Finally, the chapter concludes with a brief discussion of energy indicators, which can be thought of as a dissemination tool.

B. Reference period

8.3. A reference period is the time period for which statistical results are collected or calculated and to which, as a result, the calculated values refer. The time period may be a calendar year (reference year), a fiscal year, a semester, a quarter, a month and even a day. Energy statistics may refer either to stock data, such as refinery capacity or electricity capacity measured in one specific point in time (e.g. a specific day, net installed electricity capacity at 31 December) or to flow data, such as production, imports or exports of energy products, which comprise measures over periods of time like a year or a month. The reference period should be distinguished from the publication time, the period or point in time at which the statistical data are published. The publication year of statistical results may be significantly later than the reference year for which they were collected.

8.4. It is recommended that statistical agencies produce their energy data on a calendar basis, according to the Gregorian calendar, and consistent with the recommendations set out in IRES. For international comparability and reporting, agencies using a fiscal year should make the necessary adjustments, and clearly indicate that they are using a fiscal year if it is not possible to make those adjustments. This is particularly important when measuring energy stock levels, as these can vary significantly depending on the time of year (largely due to climatic factors).

C. Dissemination schedule

8.5. For statistics of common interest, the statistical agency should issue press releases to the media. The schedule for issuing press releases on the regular series of statistics for each calendar year should be announced well in advance in the preceding year. The schedule should be circulated to the media and announced to the general public on the website of the particular statistical agency. Apart from exceptional circumstances, all releases should be made exactly on the scheduled dates.

8.6. Prior to the lifting of the embargo on their release, all statistics should be treated with strict confidentiality and access to embargoed statistics slightly in advance of the release date should be confined to a very limited number of people and be based on genuine need. The primary purpose of allowing designated parties to have advance access to statistics is to enable them to prepare their commentaries on the economic or social conditions as may be reflected by the statistics. The extent of advance accessibility should be limited to a reasonably short period, normally three to four working days before the release date. In order to promote neutrality and objectivity in releasing statistics, the commentaries provided by other government officials must be clearly distinguishable from the statistical descriptions released by the statistics authority.

8.7. IRES encourages countries to release monthly energy data on production, stocks and stock changes within two calendar months after the end of the reference month, quarterly energy data within three months after the end of the reference quarter, and annual energy data within 15 calendar months after the end of the reference year (see IRES, para. 10.19).

D. Confidentiality

8.8. The privacy of data providers (households, enterprises, administrations and other respondents), the confidentiality of the information they provide, and its use only for statistical purposes must be absolutely guaranteed. Protection of record-level data and information is vital in preventing unlawful disclosure of individual data and information. Statistical confidentiality is critical in gaining and maintaining the trust of those required to provide data, and must be differentiated from other forms of confidentiality under which information is not provided to the public due to other considerations, such as national security concerns.

8.9. There are two possible ways to identify statistical units by which individual information can be disclosed: direct identification and indirect identification. Direct identification is possible if data of only one statistical unit are reported in a cell, while indirect identification or residual disclosure may take place if individual data can be derived from disseminated data (e.g. because there are too few units in a cell, or because of the dominance of one or two units in a cell).

8.10. The issue of confidentiality is also of national importance, and is also becoming international in scope for the following reasons:

- The high degree of interest in cross-country comparisons (in particular relating to various measures of energy and greenhouse gas (GHG) intensity).
- The globalization of energy markets.
- The geopolitical effects of a country's energy resources (particularly for oil and gas reserves).

- The liberalization of markets, meaning that a number of energy providers might be active where previously there was only one.
- Increase in data dissemination over the Internet. As a result, there is presently a demand for countries' data at a very detailed level and in some cases, even a demand for countries' microdata.

8.11. A variety of measures are available to ensure sustained confidentiality of statistical information. These include but are not limited to the following:

- The rules for ensuring statistical confidentiality should be defined in the relevant acts governing statistical production in the country (e.g. Statistics Act).
- State civil servants and agents must be subject to legislative and regulatory provisions on professional confidentiality and the obligation of discretion. These rules should apply to all files and information that come to their knowledge during the course of their work. Like all civil servants, public service statisticians must be subject to these obligations.
- Many Statistics Acts provide for criminal prosecution for any breach of statistical confidentiality.
- Physical, technological and organizational provisions should be in place to protect the security and integrity of statistical databases.
- Strict protocols must apply to external users accessing statistical microdata for research purposes. External users must be individually approved and they must sign confidentiality agreements governing their access to microdata. For the most sensitive data, access should be allowed only through a computer terminal in a secure environment. The microdata cannot be removed from that environment.

Table 8.1:
Confidentiality practices in selected countries

Country	How confidential data are identified	How confidential data are handled
Australia	The likely identification of an individual organization or business.	Data suppression; cell collapsing; seek consent to publish.
Canada	The value in a tabulation cell is composed of a few respondents; the cell is dominated by a few companies.	Confidential data are suppressed. Data may also be suppressed in a secondary manner if knowledge about them would inadvertently render a confidential value elsewhere to be identifiable.
China, Macao SAR	Total number of observation unit is less than three.	Cell suppression, except with written consent of the respondent.
Finland	The statistics compiled from the data are public under the condition that no individual establishment's data can be identified from them.	Rough classification; dominance of significant energy users considered; consent to publish sought.
Ireland	Primary confidentiality: a category is confidential if any one of the following conditions applies: (i) There are less than three units (ii) One unit accounts for more than 80 per cent of the total (dominance rule 1) (iii) Two units account for more than 90 per cent of the total (dominance rule 2)	Confidential data are not published. Anthracite and patent fuels are aggregated in order to preserve confidentiality.
Japan	One or two establishment(s) is/are represented or dominate the cell value.	Suppression is a method of disclosure avoidance used to protect companies' confidentiality. To make sure the primary suppressions cannot be closely estimated by subtracting the other cells in the table from the marginal totals, additional cells are also suppressed.

New Zealand	All counts of businesses are randomly rounded to base three. All summary estimates must pass a percentage rule check: an estimate is at risk if the value for any contributor can be calculated to within a certain accuracy.	If a specific respondent's value can be estimated very closely, it should therefore be hidden – either by aggregating categories or through suppression (with secondary suppression).
Norway	Sales figures for petroleum products are confidential when one participant's marked share exceeds 90 per cent or two participant's share exceeds 95 per cent in one segment.	We merge purchaser groups or do not publish figures on county level; only at a national level.
Poland	There are less than three units in a group or one of the units has more than 75 per cent participation in a group.	Data are not available.
Russian Federation	Production is carried out only by one manufacturer.	Data are not published.
Slovenia	Less than three units contribute to the cell value or if one reporting unit represents more than 60 per cent of the total value.	Data are suppressed.

Box 8.1:
Privacy practices at Statistics Canada

Statistics Canada is committed to respecting the privacy of individuals – whether as respondents to a survey, clients purchasing a product or service, or users of the website.

All personal information created, held or collected by Statistics Canada is protected by the Privacy Act and by the Statistics Act in the case of respondents to surveys. This means that everyone will either be asked for permission to collect their information or advised of the authority for such collection.

Survey respondents

All information provided to Statistics Canada through surveys, censuses or any other source is confidential. Statistics Canada does not release any information that identifies an individual or group without prior consent. Similarly, no other government institution has the right to see the answers given in confidence to Statistics Canada without this consent.

Because of its unique mandate as the national statistical agency in collecting personal information solely for statistical and research purposes, Statistics Canada has prepared privacy impact assessments that address privacy issues associated with its survey activities.

Clients

Statistics Canada is fully committed to protecting the confidentiality of the personal information provided by its clients who purchase products or services. This information is used only to support the relationship with clients of Statistics Canada.

Third-party social media

Statistics Canada uses social media (Facebook, Twitter and YouTube) to provide access to relevant, accurate and timely statistical information and to foster engagement, cooperation and information-sharing among people who use statistical information.

Personal information provided to the Government of Canada via social media accounts is subject to the provisions of the Access to Information Act and the Privacy Act. This information is collected to capture conversations (e.g. questions and answers, comments, "likes", retweets) between users and Statistics Canada. It may be used to respond to inquiries, or for statistical, evaluation and reporting purposes.

Box 8.1 (continued)

Visitors to the Statistics Canada website

The Government of Canada and Statistics Canada are committed to providing visitors with websites that respect their privacy.

The Statistics Canada website does not automatically gather any specific personal information, such as names, phone numbers or email addresses, unless they are provided by users by sending an email or registering in a secure area of the site. The information that is received and stored is safeguarded to prevent unauthorized disclosure.

Information on individual website visitors is used by Statistics Canada employees who need to know the information in order to respond to their requests or to ensure the security of the system. The information is only shared with another government department if an inquiry relates to that particular department. The information is not used to create individually identifiable profiles, nor is it disclosed to anyone outside the federal government.

Adapted from: www.statcan.gc.ca/reference/privacy-privee-eng.htm, October 2014.

E. Revision policy

8.12. Revisions are an important part of the compilation of energy statistics. Revisions result from a reassessment of the past values of statistical variables, and are determined by a variety of factors which include, but are not limited to:

- Conceptual changes (e.g. changes in nomenclatures and definitions).
- Improvement of methodological procedures (e.g. changes in detail and stratification of sample survey data).
- A change in statistical data sources (e.g. replacement of survey data with data from administrative sources).
- Provisional data or estimates based on partial data being updated with “firm” data.
- Inclusion of additional observations, which, in the case of statistics obtained from econometric time series methods, determines the revision of estimated coefficients and possibly the model used (e.g. seasonally and/or calendar adjusted variables).
- The correction of errors.

8.13. There are two main types of revisions which statistical bodies commonly consider:

- ***Routine (normal/concurrent) revisions***: These are revisions which are part of the regular statistical production process and whose aim is to incorporate new or updated data, or to correct data or compilation errors. In general, these revisions should be carried out up to the moment there is sufficient information to publish a reasonably accurate figure. These revisions may occur for both annual and monthly and quarterly statistics, though more frequent statistics are more prone to revisions given the time spans involved. Routine revisions can also reflect the results of structural statistical operations, such as population censuses or a household budget survey. These revisions thus occur at a relatively regular periodicity, thereby reflecting the frequency of the structural operations. In certain cases, revisions may be extensive, aimed at building back series that ensure inter-temporal comparability. These general routine revisions should be used to the greatest extent possible as a way to introduce new statistical sources, changes to the conceptual framework and improvements in methodological algorithms. They tend to occur on a less frequent basis in the case of annual (or less frequent) statistics. With regards to infra-annual statistics, general routine revisions can be in the form of an annual revision, with a view to integrating more complete data made available for a whole year (sometimes referred to as benchmarking).

- ***Major or special revisions***: These are not part of the regular revision schedule and are conducted in order to incorporate major changes in concepts, definitions and classifications, and changes in data sources.

8.14. In order to effectively manage revisions, it is highly important for statistical agencies to develop a revision policy that will guide revisions to the statistics they produce. The development of such a policy should be aimed at providing users with the information necessary for coping with revisions in a systematic manner. The absence of coordination and planning of revisions is considered a quality problem by users. Features of a well-established revision policy could be a predetermined revision and release schedule; reasonable stability from year to year; advance notice of reasons and effects and easy access to sufficiently long time series of revised data; and adequate documentation of revisions included in statistical publications and databases. A sound revision policy is recognized as an important aspect of good governance in statistics, as it will not only help national users of the data but will also promote international consistency. Further, IRES recommends that any energy statistics revision policy be well coordinated with those of other areas of statistics.

Box 8.2:
General EIA weekly natural gas storage report revisions policy

The unscheduled release of revisions to weekly estimates of working gas held in underground storage shall occur when the cumulative effect of data changes or corrections is at least 10 Bcf for the current or prior week. Revisions shall be disseminated on a federal workday between 2 p.m. and 2.10 p.m. (Eastern Time), following notice of the pending release to the public between 1 p.m. and 1.10 p.m. (Eastern Time). If a revision is made, changes to all affected regions are to be recorded in the 2–2.10 p.m. release. Public notification can occur in a number of ways, including a website notice of the impending release of revised data to replace the current weekly natural gas storage report, email notification to selected media, and an email notice sent to all users of the weekly natural gas storage report data who have signed onto a free distribution service. There are two special circumstances related to handling certain data changes in the weekly natural gas storage report. First, this unscheduled release policy does not apply to data changes resulting from changes in the estimation methodology or parameters because those changes are announced in advance. Second, reclassification of gas (between working gas and base gas inventories) is to be reported only in regularly scheduled releases of the weekly natural gas storage report. The policy for unscheduled releases of revisions was announced in April 2005, as a Federal Register notice (70 FR 21406), a copy of which is available at www.eia.gov/survey/frn/naturalgas/WNGSR-Unscheduled-Release-Policy-Final-April2005.pdf.

Adapted from http://ir.eia.gov/ngs/revisions.html, October 2014.

F. Dissemination format and access

8.15. A flow of comprehensive, reliable, accessible and timely statistics is indispensable to informed policymaking, and the wide availability of statistics to the general public and media helps support an informed public debate. These roles for statistics lead to certain basic principles for the official statistics function within a country, as described in the Fundamental Principles of Official Statistics, adopted by the General Assembly in 2014 (resolution 68/261). To meet the divergent needs of data users, energy statistics should be disseminated through:

- The issuing of press releases containing statistics of common interest to the general public;
- The release of statistical publications providing statistical data and methodologies in greater detail;
- CD-ROM products for distribution to key users and subscribers;
- Online interactive systems which enable data users to build statistical tables to their own specific requirements through a simple and user friendly step-by-step process; and
- The provision of data enquiry services for users requiring specific data. Statistical agencies are encouraged to make releases in as many formats as possible.

8.16. However, since electronic formats greatly increase the number of parties and geographic locations to which simultaneous access can be provided, it is recommended that energy statistics be made available electronically. With regards to international reporting, statistical agencies are encouraged to disseminate their energy statistics, as soon as they become available, to national users. In order to ensure a speedy and accurate data transfer to international and regional organizations, it is recommended that countries use the Statistical Data and Metadata Exchange (SDMX) format, if possible, for the exchange and sharing of their data.[27]

[27] At the time of writing, this had not yet been developed for energy.

Box 8.3:
Energy statistics online databases in France

The French Ministry of Ecology, Sustainable Development and Energy has two online databases on energy statistics.

The database **Pégase** (French acronym for oil, electricity, gas and other energy statistics) contains energy statistics compiled by the Department of Observation and Statistics (SOeS). Users can create and download customized tables, both of monthly and annual data.

Monthly data refer to production, consumption, imports and exports; a simplified monthly energy balance is created with the data available at the end of each month. Price statistics contain examples of prices for the various energy sources, for household consumption or for consumption of firms with prices, exclusive of value added tax (VAT). Pégase also includes data on the foreign trade of energy products and international prices.

The series are revised each month. For example, the estimated value for wind generation in April is released in early June and may be significantly revised in early July.

Annual energy statistics refer to the supply and consumption of different energy products. Pégase contains the annual series in original units (tons for coal or oil, kWh for gas or electricity) but also in tons of oil equivalent to allow comparisons between products. Annual statistics on prices, imports, exports and international prices are also available, either deduced from monthly statistics or based on data collected through annual surveys.

Box 8.3 (continued)

The following figure provides an example of a table available on Pégase.

Imports, production and consumption of electricity in France (GWh), online database Pégase
(Ministère de l'écologie, du développement durable et de l'énergie, France)

Rapports | Tableau | Graphique — Liens utiles

Pégase – Importations, production, consommation d'électricité en France (en GWh) (1970-2013)

AUTRE : Units - donnees

Periode / Libelle	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013
Production brute d'électricité primaire	514 437	509 498	514 333	507 886	514 015	480 224	506 516	504 530	508 505	520 954
Production brute d'électricité primaire renouvelable	65 196	57 969	64 142	68 150	74 568	70 488	78 237	64 447	83 098	97 269
Production brute d'électricité nucléaire	448 241	451 529	450 191	439 730	439 447	409 736	428 279	442 082	425 406	423 685
Importations d'électricité	5 571	8 062	8 522	10 796	10 748	19 213	19 475	9 501	11 984	11 687
Exportations d'électricité	-68 477	-58 390	-71 863	-67 525	-58 736	-44 913	-50 206	-65 914	-56 933	-60 148
Total disponibilités d'électricité	452 531	449 170	450 992	451 147	466 027	454 524	475 785	450 117	463 555	472 493
Consommation d'électricité de la branche énergie	-59 841	-56 672	-60 226	-62 013	-60 113	-58 849	-62 815	-56 245	-55 954	-54 080
Consommation finale d'électricité; corrigée du climat	423 492	423 175	429 437	431 011	439 862	425 276	443 033	438 532	438 833	441 440
Consommation d'électricité de la sidérurgie	11 657	10 978	11 948	11 588	11 830	8 782	10 464	11 155	10 584	10 268
Consommation d'électricité de l'industrie	127 710	126 273	123 958	123 785	120 851	108 125	110 486	108 019	107 807	106 958
Consommation d'électricité du résidentiel-tertiaire; corrigée du climat	262 064	256 654	273 932	276 535	287 870	288 623	302 243	291 241	299 546	303 012
Consommation d'électricité de l'agriculture	6 899	7 326	7 485	6 907	6 606	7 460	7 616	8 004	8 450	8 650
Consommation d'électricité des transports	12 162	11 944	12 114	12 196	12 705	12 286	12 224	12 113	12 446	12 552
Consommation totale d'électricité primaire; corrigée du climat	452 114	448 711	452 505	455 955	467 920	455 588	469 669	458 377	464 767	469 561
Correction climatique pour la consommation d'électricité	-417	-459	1 513	4 808	1 893	1 064	-6 116	8 268	1 212	-2 932

The database **Eider** contains some of the key statistical information on environment, housing, transport and energy in the regions and departments of France. Data are available in three formats (detailed tables, regional portraits and long series) and are regularly updated.

Source: www.statistiques.developpement-durable.gouv.fr/donnees-ligne/th/energies-climat-1.html, October 2014.

8.17. Countries should also provide adequate metadata to define and describe energy statistics to enable users to make an informed assessment of the usefulness and relevance of the statistics and to address their needs.

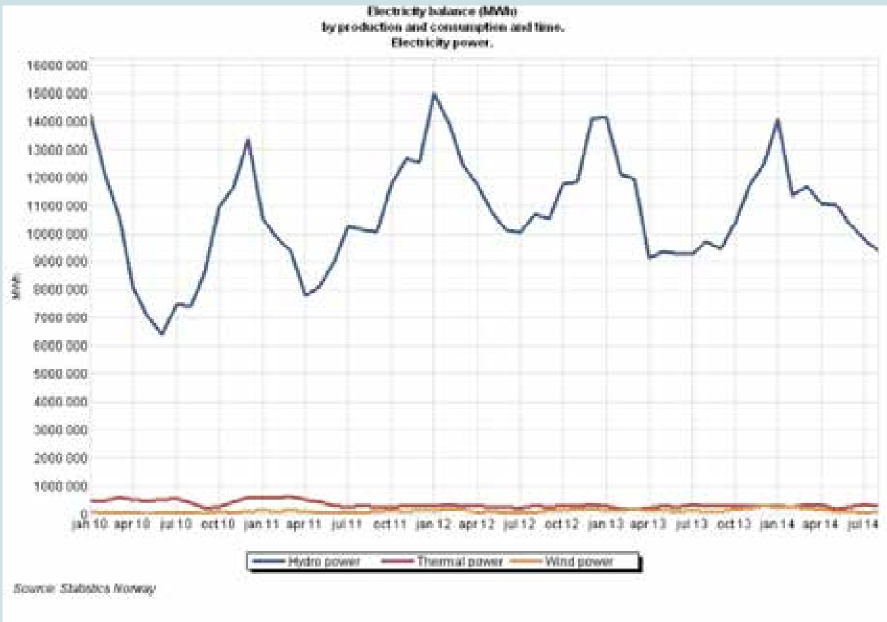
Box 8.4: Energy statistics metadata on Bulgaria's NSO website

The National Statistical Institute of Bulgaria publishes on its website a number of tables related to monthly statistics on production and deliveries of energy products (solid fuels, oil and petroleum products, natural gas, electricity), on electricity and natural gas prices, on combined heat and power (CHP) units, and on energy balances. For each topic, a methodology summary (containing descriptions of variables and commodities, definitions and used units of measure) and a quality report are available. Quality reports comprise general information on data collections, such as: name of the data collection; national body responsible; objective and history; reference period, frequency and transmission deadline. Additional detailed information included in the report refers to: legal framework; variables surveyed and derived; type of data collection (business/household survey, administrative source, etc.); population frame and reporting unit; sample size and unit non-response rate; classifications; compilation of the final data set, models and statistical estimation techniques used; overall accuracy; data revision policy; methodology; and quality documentation.

Adapted from www.nsi.bg/en/content/4948/energy, October 2014.

Box 8.5:
Norway: Preliminary energy statistics

Norway compiles and presents preliminary commodity balances and a preliminary energy balance about four months after the reference year, based on short-term (monthly) statistics. These are more aggregated and unreliable than the detailed balances published towards the end of the year. Still, much interest is generated for the preliminary figures because of the short time lag between the reference period and the release date, making their impact more relevant.



G. Meeting user needs

8.18. There are many types of users of energy data, such as politicians, government agencies, public and private sector companies, civil society organizations, academia, the media, the general public and international agencies. They need disparate statistics for various purposes, and they vary in capacity and sophistication in their use of statistics. Some needs may be suppressed because of the lack of available statistics, and potential demand should be considered, as well as current demand.

Box 8.6:
Statistics South Africa's experience with social media

A new method of disseminating data is the use of social media. Social media platforms provide organizations with access to a wide spectrum of users on a daily basis. It is also a platform for getting instant feedback from external stakeholders.

Statistics South Africa has been using mainly Twitter and Facebook very successfully, reaching new and current users instantly.

The fact that social media platforms are immediate has both positive and negative implications. The immediacy of the platforms allows an organization to disseminate information

Box 8.6 (continued)

quickly and interact with users and followers on a timely basis. This also means that the accounts have to be monitored on a consistent basis as delays in responding on these platforms jeopardizes the organization's credibility.

Statistics South Africa uses these platforms to inform followers about the organization and the services and products provided, and sends out live messages from big events, as well as information about daily releases. These posts usually contain a link to its website for easy access to the releases, as well as a graphic with the highlights of the release. This drives more traffic to the website, one of the main dissemination platforms.

Challenges to the use of social media platforms are the monitoring of these platforms and fast interaction with users; and the creation of a constant stream of new and innovative content.

8.19. User needs cannot be properly met unless these have been properly identified, synthesized, understood and prioritized. It is important to understand that users invariably have a long list of statistical needs, and every effort should be made to guide users to identify their priorities. Also, user needs and priorities are always changing and tracking these changes requires ongoing consultation and dialogue with users. This can be done through various approaches to assess user needs.

Box 8.7:
Statistics South Africa's Dissemination Programme

Statistics South Africa's (Stats SA) mandate to meet user needs must take into account the fact that different users have differing requirements in terms of volume and detail of information required. Once these users have been identified, it is necessary to determine their unique data needs. Stats SA makes use of a range of dissemination channels for the release of data to the users identified, the primary dissemination tool being StatsOnline (Stats SA's webpage). The following tools are also used:

- Statistical publications (available in hard copy or PDF online).
- Electronic media (CD-ROM).
- Statistical databases (Nesstar).
- Microdata.
- Information services, provided by advisers from individual statistical fields.
- User information services enquiries.
- Training and workshops to stakeholders.

The combination of dissemination tools is based on the nature of the release. All outputs disseminated by Stats SA (both statistical and non-statistical) must be released in the form of one or more of the dissemination methods listed above.

South Africa is faced with the particular issue of Internet access not being widely available to all those wanting to access data, and those that have access may not have adequate speed to use the online data dissemination methods, as a result of challenges due to infrastructure. This is where the use of electronic media (CD-ROM) is critical in disseminating data. An example is the Census 2011 data that had to be made available on a CD-ROM pack and provided to data users on request. This allows users to upload the data onto their computers and to use them as they see fit without the need for Internet connection. A CD-ROM pack can be picked up from Stats SA's head office or any one of the regional offices.

8.20. Consultations and discussions with users should aim to establish, among other things, how they use statistics in their operations, the availability of required statistics and how they may have been constrained by lack of adequate statistics, their assessment of the adequacy of existing statistics in terms of relevance, accuracy, consistency, completeness, timeliness, level of disaggregation (geographic, product type, etc.) and accessibility, their current and perceived future statistical needs and priorities and where there are gaps, and how they think their needs can best be met.

8.21. One positive approach is to identify those users interested in particular data sets or groups of data and arrange contact with them. Users can be identified or traced via user registration forms which can be made available as a prerequisite for the download of a statistical publication or prompted to reply to a survey through a dissemination website. The mailing list used for disseminating statistics and identifying those frequently in contact with the statistical agency may be a good starting point in determining the main data users. Selected institutions from each of the main user groups (academia, government, industry, etc.) should be included in the consultations and discussions held with them, either individually or in small groups, whilst others might be invited to contribute in writing. The process should ensure that policy and decision makers, as well as technical staff in user institutions are consulted.

Box 8.8:
United Kingdom's Department of Energy and Climate Change User Survey

Meeting users' needs is at the heart of the United Kingdom Statistics Authority's (UKSA) Code of Practice. Capturing users' views with the aim of delivering transparent and quality statistics is an important part of the authority's assessment process. The Department of Energy and Climate Change (DECC)'s commitment to the Code of Practice is reflected in all of its professionally produced statistics which have been assessed as national statistics by the UKSA.

DECC values the needs of its data users and aims to maintain its delivery of reliable and quality data to satisfy their requirements. By way of refining and improving its data delivery, DECC carried out a survey, in December 2012, of all its users, including those in government and local authorities, who signed up to receive statistics alerts via the DECC website. In 2010, DECC undertook a similar survey of users of its statistics.

The survey's key findings, especially of current and future user needs together with satisfaction with DECC's interaction with users, can be found at the following link:

Adapted from www.gov.uk/government/uploads/system/uploads/attachment_data/file/239316/decc_statistics_user_survey_2012.pdf, October 2014.

H. Energy indicators as a dissemination tool

8.22. In order to reach wider audiences, the use of summarized data in the form of internationally accepted energy indicators is a way to disseminate information that may otherwise be too complex to raise the interest of the non-expert. The calculated indicators often better explain the phenomenon to be analysed than raw energy data (either in original or balance form). While intuitive, energy indicators also shed light on what driving forces lie beneath the growth in energy use. This is useful information for policymakers who want to develop evidence-based policy initiatives and also to evaluate the effect of past and future initiatives. As such, they are a key tool in effective energy statistics dissemination. This section does not provide recommendations on energy indicators, or even give a comprehensive description of all possible indicators. Instead, some common energy indicators, using data from energy balances, are explored and described here.

8.23. The most basic energy indicators can be obtained from an energy balance, as described in chapter VI. Values for total production, total energy supply and final consumption of energy are three of the most basic derived energy statistics. Even only using these three numbers, meaningful insights to a territory's energy situation can be obtained:

- Dividing production by the total energy supply can show how reliant a country is on imports to satisfy its energy needs, thus measuring energy security. This is often called an energy self-sufficiency indicator. (A similar indicator, energy dependency, is often calculated by dividing total imports by total energy supply).
- Dividing final consumption by total energy supply can show how much energy is being lost in transformation (together with losses and energy industries own use). This figure is valuable in showing efficiency of a country's transformation processes, and is thus a useful energy efficiency indicator (albeit rather a crude one, if the specifics of its power generation make-up are not considered).

8.24. The next layer in the indicator pyramid combines these balance data with other socioeconomic statistics to make cross-country and inter-temporal comparison meaningful. The two figures most commonly used for this purpose are population and GDP level, and thus two commonly composed indicators are:

- Total energy supply per capita, indicating how much energy is used per person in the territory, and
- Total energy supply per unit of GDP, indicating how much energy is required to create one unit of economic output, sometimes called the energy intensity of the economy.

8.25. At this stage, other indicators can also be of use. Instead of using total energy supply, sometimes more interest lies in electricity supply, a proxy for electricity access, and so electricity supply per capita is calculated. This both highlights the utility and limitation of certain energy indicators; while this indicator can give an idea of how much electricity is consumed per person, in countries where only a few users use large quantities of electricity the indicator will obviously fail to reflect electricity access.

8.26. On the environmental level, energy supply can not only be combined with environmental data but used also to calculate environmental data itself. One of the principal uses of energy statistics is to calculate greenhouse gas emissions, in particular CO_2 emissions from fuel combustion. The two environmental indicators relating to energy are thus:

- *Emissions of GHG (or CO_2) per capita (or unit of GDP).* This is a much-used measure in the climate change debate over how much GHGs a country produces over a given time period. While it is an easy figure to analyse and compare across countries and time periods, it is important to understand that countries with large manufacturing, energy-producing or general industrial bases will naturally have high values. Similarly, large, cold (or very hot) countries would be expected to use more energy on transport and climate control, thus a larger value for them is understandable.

- *Emissions of GHG (or CO_2) per unit of energy supply*. This can be considered the carbon intensity of the territory, and countries with similar energy supply levels may have very different values depending on the source of their energy, particularly their power generation.

8.27. Energy efficiency indicators are energy indicators that measure improvements (or deteriorations) in energy efficiency across countries, industries and technologies. While energy statistics, as defined by IRES, are used in energy efficiency indicators, in general, further, more detailed data are required, such as floor space, steel production, passenger kilometre or motorization rate. It is beyond the scope of this *Manual* to describe these indicators in detail, as IRES does not provide recommendations on them. For further reading, the IEA manual, *Energy Efficiency Indicators: Fundamentals on Statistics*,[28] can be consulted.

[28] See www.iea.org/reports/energy-efficiency-indicators-fundamentals-on-statistics.