

Effective Energy Data Management for Low-carbon Growth Planning: An Analytical Framework for Assessment

Bo Liu (lbesf.chn@gmail.com)¹, Meredydd Evans (mevans@pnnl.gov)^{1*}, Sha Yu (Sha.Yu@pnnl.gov)¹, Volha Roshchanka (volha@protonmail.com)², Srihari Dukkipati (srihari@prayaspune.org)³, Ashok Sreenivas (ashok@prayaspune.org)³

¹Joint Global Change Research Institute, Pacific Northwest National Laboratory, 5825 University Research Court, Suite 3500, College Park, Maryland 20740, USA

²Earth System Science Interdisciplinary Center, University of Maryland, 5825 University Research Court, Suite 3500, College Park, Maryland 20740, USA

³Prayas Energy Group, III A&B, Devgiri, Joshi Railway Museum Lane, Kothrud Industrial Area, Kothrud, Pune, Maharashtra 411038, India

*Corresponding author:

Tel: 1 (301) 314-6739

Fax: 1 (301) 314-6719

Abbreviations and Acronyms

AER	Alberta Energy Regulator
AGEB	Working Group on Energy Balances
APPA	American Public Power Association
BAFA	Federal Office of Economics and Export Control
BEE	Bureau of Energy Efficiency
BEIS	Department for Business, Energy & Industrial Strategy
BIS	Department for Business, Innovation and Skills
BMWi	Federal Ministry for Economic Affairs and Energy
CA	The Coal Authority
CBECS	Commercial Buildings Energy Consumption Survey
CCO	Coal Controller's Organisation
CEA	Central Electricity Authority
CGPL	Combustion, Gasification and Propulsion Laboratory
DECC	Department of Energy and Climate Change
Destatis	Federal Statistical Office
DGH	Directorate General of Hydrocarbons
DLHS	District Level Household and Facility Survey
DOC	Department of Commerce
DOE	Department of Energy
DOT	Department of Transportation
ECCC	Environment and Climate Change Canada
EIA	Energy Information Administration
ESCM	Energy Statistics Compilers Manual
FERC	Federal Energy Regulatory Commission

GSS	Government Statistical Service
IHDS	India Human Development Survey
InterEnerStat	Intersecretariat Working Group on Energy Statistics
IRES	International Recommendations for Energy Statistics
MNRE	Ministry of New and Renewable Energy
MoR	Ministry of Railways
MoRTH	Ministry of Road Transport and Highways
MoSPI	Ministry of Statistics and Programme Implementation
MoWR	Ministry of Water Resources
NEB	National Energy Board
NFHS	National Family Health Survey
NISE	National Institute of Solar Energy
NIWE	National Institute of Wind Energy
NRC	Nuclear Regulatory Commission
NRCan	Natural Resources Canada
NSSO	National Sample Survey Office
OEB	Ontario Energy Board
Ofgem	Office of Gas and Electricity Markets
OMB	Office of Management and Budget
ONS	Office for National Statistics
PPAC	Petroleum Planning and Analysis Cell
RECS	Residential Energy Consumption Survey
StatCan	Statistics Canada

Abstract

Readily available and reliable energy data is fundamental to effective analysis and policymaking for the energy sector. Energy statistics of high quality, systematically compiled and effectively disseminated, not only support governments to ensure national security and evaluate energy policies, but they also guide investment decisions in both the private and public sectors. Because of energy's close link to greenhouse gas emissions, energy data has a particularly important role in assessing emissions and strategies to reduce emissions. In this study, energy data management in four countries – Canada, Germany, the United Kingdom and the United States – are examined from both organizational and operational perspectives. With insights from these best practices, we present a framework for the evaluation of national energy data management systems. It can be used by national statistics compilers to assess their chosen model and to identify areas for improvement. We then use India as a test case for this framework to assess how India may adapt and evolve its energy data management systems. Its government is working to enhance India's energy data management to improve sustainable growth planning.

Keywords

Climate change, energy data management, low-carbon growth planning

1. Introduction

Climate change has been recognized as a common threat to both natural and human systems (IPCC, 2014). With years of negotiations on combating global climate change, the 21st Conference of the Parties (COP21) in Paris made history with an international agreement on transitioning to a low-carbon and climate-resilient future. As a major contributor to climate change, the energy sector accounts for approximately two-thirds of total anthropogenic greenhouse gas (GHG) emissions globally (IEA, 2015). Mitigating energy-related emissions plays a significant role in the future of sustainable development and low-carbon growth. However, accurately calculating GHG emissions requires activity-based data and reasonable accounting methods. Readily available and reliable energy data is fundamental to effective analysis and policymaking for the energy sector. Energy statistics of high quality, systematically compiled and effectively disseminated, ensure national security and evaluate energy policies and guide investment decisions in both the private and public sectors.

Quality assurance, for energy or other topics, requires a complex and multi-dimensional approach. Typically data quality includes assuring accuracy, completeness, timeliness, relevance and accessibility. Data quality also maintains objectivity and signals independence from political and other influence. That way public data users will believe the data, which in turn will increase the likelihood that a range of stakeholders make decisions based on the data. In addition, statistical agencies may be challenged by limited budgets and trade-offs between quantity and quality because of constraints on financial and human resources. Therefore assessing the quality of energy data requires an examination of the overall energy data management system. To maintain the effectiveness of data management systems, international organizations and governments have developed fundamental principles or codes of practices. There are three

leading statements of principles that largely overlap (Table 1). The U.N. code of practice is a broad statement of principles; the others are statements of guiding principles regarding practices within two specific geographic regions.

Harmonizing the processes for preparing data will ultimately improve data in all countries and make energy data more compatible worldwide. Beyond the general principles that countries may follow, the United Nations Statistical Commission created the Oslo Group and the Intersecretariat Working Group on Energy Statistics (InterEnerStat) in 2005 to provide general guidance on compiling national energy statistics. The Oslo Group aims at addressing methodological issues and international standards; InterEnerStat focuses on coordination among international organizations and harmonizing definitions of energy products and flows (UNSD, 2015). Two major accomplishments of the Oslo Group since its inauguration include the development of the International Recommendations for Energy Statistics (IRES) and the Energy Statistics Compilers Manual (ESCM). The IRES serves as a common process framework. Its recommendations include definitions and classifications, data sources, data compilation strategies, data quality assurance, dissemination policies, and other topics (UNSD, 2011). The IRES presents definitions and guidelines from a theoretical perspective. The ESCM, which still has to be finalized, complements IRES in a more practical way by providing illustrative examples from one country or another. IRES and ESCM provide an overview of common standard practices that maintain consistency and increase comparability across countries. They are not a collection of best practices as much as they are a set of voluntary guidance, which provide examples that compilers of statistics might use. A key outcome of this U.N. work is harmonizing how surveys are done. That makes data much more compatible and comparable than when data harmonization is merely added on at the end of the process. However, neither

IRES nor ESCM provide guidance for the evolution of effective energy data management in certain countries. There is still a need to explore how countries address energy data issues and how they overcome specific national challenges. That exploration is the organizational principle of our study.

Countries have adopted various models for national energy data management. These models depend on energy sector characteristics, economic structure, country size, a country's type of government, and other factors. No single energy data management system fits all countries, but a review of how several countries do it can suggest best practices for collecting, processing, and disseminating national energy data. With insights from these best practices, we present a framework for the evaluation of national energy data management systems. It can be used by national statistics compilers to assess their chosen model and to identify areas for improvement. To illustrate the framework's usefulness, we take India as a case study by investigating its national energy data management system.

2. Methodology and Data

To assess national data systems, we compiled and compared available information from official government websites and documents, including information on how national systems developed and improved over time. We then integrated this with information obtained directly from national statistics officials or government representatives in order to verify the information and to identify when ambiguity exists.

This study examines models of energy data management from four countries – the United States (U.S.), Canada, the United Kingdom (UK) and Germany. We selected these countries to demonstrate the diversity and effectiveness across various models. These countries are all

members of the International Energy Agency and all of them have energy data systems that are at a minimum several decades old. Thus, we describe the organization and operations of energy data management in these countries, especially how energy data are collected, processed, analyzed, and disseminated, and how these practices have changed as the countries have developed. We also discuss tools and practices used to overcome any challenges in their chosen models. This study's analytical framework is based on the processes and practices of energy data management in these four countries. However, the basic principles can be useful to a range of countries, including those with more recent histories of developing comprehensive energy statistics and energy balances.

India recognizes the importance of energy data in sustainable development and low-carbon growth planning and has been collaborating with the U.S. government to improve its energy data management. That led to the Sustainable Growth Working Group under the U.S.-India Energy Dialogue, in which energy data management is one of three focus areas. We use this analytical framework to assess the current energy data management in India and demonstrate the effectiveness of the framework in managerial assessment of national energy data management systems. At the same time, we feel this approach, which we have applied to India, might be helpful for a range of developing countries seeking to improve their energy data management systems.

3. Results and Discussion

3.1 Models of energy data management: an organizational perspective

Countries share a common understanding of effective data management, but have adopted various models to satisfy their needs (Figure 1). A country's political system, economic structure,

size, and type of government can have a large influence over how its national energy data management system operates. Data management models can be centralized or decentralized, depending on how much responsibility for official statistics falls on a central agency or on specialized government agencies. Decentralization can take place in many ways, so that functions are distributed, for example, by region (Germany, UK) or agency specialization (Canada, Germany). The national policy context determines the design of a statistical system; in turn, the design can have an impact on the process of policymaking.

3.1.1 The United States

The United States has a distinct agency for energy statistics: the Energy Information Administration (EIA). EIA, along with 12 other principal statistical agencies and some 114 statistical programs, constitute the Federal Statistical System (Office of Management and Budget, 2015). The Office of Management and Budget (OMB) oversees the system and is in charge of ensuring efficiency and compliance with federal regulations. As the principle source of U.S. energy statistics, EIA is responsible for collecting, analyzing and disseminating energy information. Although it is a part of the U.S. Department of Energy (DOE), EIA has been given total independence with respect to data collection, analysis of information, and preparation of reports (Department of Energy Organization Act of 1977). EIA has about 370 federal employees and in Fiscal Year 2016 had a budget of \$122 million (EIA, n.d.-a, c).

To avoid duplication, EIA coordinates with other government agencies and private organizations, such as the Federal Energy Regulatory Commission (FERC), the Department of Commerce, the Nuclear Regulatory Commission (NRC), the American Public Power Association (APPA), and others. These agencies and organizations typically collect energy information that

is necessary for, or is a result of, their organizational function. For example, the Division of Natural Gas Regulation at DOE's Office of Fossil Energy authorizes natural gas imports and exports. The Division collects information from all natural gas exporting and importing companies and shares that with EIA for data dissemination. Also, FERC separately collects electricity data which allows the agency to regulate the electricity market. EIA produces the majority of official energy data and analysis to support policymaking, business decisions, and public understanding of the energy sector. But for enforcement and other regulatory activities, EIA and others still rely on other government entities.

3.1.2 Canada

Canada's central statistical office, Statistics Canada (StatCan), is responsible for the collection, compilation, analysis, and publication of statistical information at both the national and provincial/territorial levels (Statistics Act of 1985). In 2014-2015, StatCan employed 4,758 full-time equivalent (FTE) staff members and had an expenditure of \$467 million (Statistics Canada, 2016a). StatCan's Energy Statistics Program (ESP), the main source of energy data in Canada, is managed by the Industry Statistics Division. The ESP had 432 FTEs with an annual spending of 47 million in 2014-15 (Statistics Canada, 2016a). While ESP is the primary source of production, transformation, and consumption data, other parts of StatCan also collect energy data. These include data on the transport, export, and import of energy products; on transport sector energy consumption; and on employment and investment in the energy sector (Statistics Canada, 2016c). Other major energy data sources, including the National Energy Board (NEB), Natural Resources Canada (NRCan), and Environment and Climate Change Canada, produce more specialized energy information and analysis related to their activities. As a regulatory

agency for the oil, gas, and electricity industries, the NEB compiles data from company reports and produces statistics related to the energy market. NRCan maintains the National Energy Use Database, which includes data on energy consumption and efficiency at the end-use level. Environment and Climate Change Canada develops the Canada's GHG emissions inventory, which covers emissions and activities associated with the energy sector. These agencies all comply with statistical standards developed by StatCan.

To address regional energy issues, many provinces and territories in Canada also have energy regulatory agencies. These include the Alberta Energy Regulator (AER), the British Columbia Oil and Gas Commission, and the Ontario Energy Board (OEB). Such agencies are key providers of energy data in their respective jurisdictions. With specific mandates, StatCan has established data-sharing agreements with provincial and territorial agencies to improve efficiency, to reduce costs, and to alleviate the burden to the respondents.

3.1.3 The United Kingdom

The UK Statistics Authority oversees the production and publication of official statistics in the UK and establishes the required code of practice and standards for agencies. Because of its political structure, the UK maintains a statistical system that has elements of regional decentralization. To achieve coherence in national statistics, the UK entered into an agreement with the three devolved administrations of Scotland, Wales, and Northern Ireland. The resulting Inter Administration Working Agreement on Statistics helps these separate government entities coordinate their responsibilities for producing official statistics. The executive office of the UK Statistics Authority, the Office for National Statistics (ONS), is the largest producer of official statistics. Other governmental departments produce statistics relevant to their national policy

areas, and devolved administrations collect statistics important to regional government policy priorities. The various government departments, ONS, and the administrations of Scotland and Wales form the Government Statistical Service (GSS), a virtual community for cooperation. The Northern Ireland administration is not part of the GSS, although they work very closely with the Service. GSS has about 7,000 civil servants who collect, produce, and disseminate official statistics.

The UK's Department of Energy and Climate Change (DECC)¹ is the lead body for energy policy. It produces statistics covering energy, climate change, fuel poverty, and related areas for both the whole country and for the devolved administrations. DECC employs around 1,600 staff members and in 2014-2015 had a gross expenditure of £5.7 billion (DECC, 2015). DECC statistics staff report to the DECC Head of Profession for Statistics, who is accountable to the UK Statistics Authority. DECC's statistical division has around 60 employees, half of whom are statisticians; the other half are administrative specialists and energy experts. Staff members work in teams of four to five people for each fuel or subject area (MacLeay, Iain, personal communication, September 19, 2014). Besides the statistical division, other DECC branches as well as other government agencies also support energy data production. Examples of such agencies include the UK Coal Authority, the UK Atomic Energy Authority, and the Office of Gas and Electricity Markets (Ofgem).

3.1.4 Germany

¹ In July 2016, Prime Minister Theresa May merged DECC with the formerly Department for Business, Innovation and Skills (BIS) and formed the new Department for Business, Energy & Industrial Strategy (BEIS).

Germany's statistical system is the most decentralized of the four countries in this study. Data collection and processing are typically conducted by the separate statistical offices of Germany's 16 federal states. The Federal Statistical Office (Statistisches Bundesamt, or Destatis) is the central authority responsible for the methodological and technical aspects of surveys, coordination with state statistical offices on surveys, compilation and publication of national statistics, and further development of the federal statistical system (Statistisches Bundesamt, n.d.). Destatis had a total staff of 2,282 and a budget of over 162 million euros in 2015 (Statistisches Bundesamt, 2016). The statistical offices of the federal states are not subordinate to Destatis; they are independent from the federal government in terms of administration.

National energy data in Germany are managed by Destatis and other government agencies. Destatis mainly compiles statistics on energy supply, energy consumption, domestic energy prices, and emissions. Other government agencies publish data relevant to their administrative functions. As the leading government body on energy policy, the Federal Ministry for Economic Affairs and Energy (Bundesministerium für Wirtschaft und Energie, or BMWi) publishes energy data and forecasts. It also leads the Working Group on Renewable Energy Statistics that engages all energy-related government entities for the integration and publication of renewable energy data. The Working Group on Energy Balances (Arbeitsgemeinschaft Energiebilanzen, or AGEB) cooperates with working groups at the level of the federal states in order to produce annual reports on national energy balances. In addition, the Federal Office of Economics and Export Control (Bundesamt für Wirtschaft und Ausfuhrkontrolle, or BAFA) collects and maintains statistics on fossil fuels, including data on imports and cross-border prices.

As these four country models demonstrate, energy data can be managed by systems with various institutional structures. The main roles and responsibilities for energy statistics may follow one of four main systems: a specialized statistical agency, such as the EIA in the United States; a general statistical agency in coordination with specialized administrative agencies, as in Canada's StatCan, NEB, and NRCan; a statistical branch within a specialized administrative agency that coordinates with counterparts at devolved administrations, such as the UK's DECC; or regional statistical authorities in coordination with a centralized general statistical agency, other administrative agencies, or working groups, as in the case of Germany's Destatis, BMWi, and AGEB. The functions of these energy data management systems are always constrained by their budgets and personnel resources, but energy statistics professionals usually find ways to overcome those challenges. With these four countries providing examples of energy data management, and in search of best practices, we next explore the effectiveness of these systems through the lens of operations.

3.2 Best practices on energy data management: an operational perspective

Besides the organizational aspect, operations play a significant role in effective energy data management. In this section, we examine the operations of these four country-based statistical systems to understand what practices and tools are in place. We look for fundamental principles, and for means of data collection and coverage, quality assurance and validation, modeling and projections, and data dissemination.

3.2.1 Fundamental principles for operations

Countries typically maintain a set of fundamental principles for the operations of their national energy data management systems. Employing these principles in actual operations reflects some good practices. We describe some key characteristics in this aspect in the following categories: (a) legal status and independence; (b) relevance to policymakers and users; (c) coordination; (d) implementation and adequacy of resources; (e) quality of data and methodologies.

All four statistical systems we examine here have a legal framework that defines the roles and responsibilities of stakeholders; provides a mandate for data collection, which safeguards independence from political and other undue influences; ensures confidentiality of data providers; and encourages data sharing. For example, the U.S. Congress established both DOE and EIA in 1977 through the Department of Energy Organization Act. These built upon systems and organizations first established by the Federal Energy Administration Act of 1974, which gave EIA the authority to gather data from energy producers and consumers. In Canada, a single law, the Statistical Act, lays out responsibilities and establishes a very broad mandate for Statistics Canada to enter into data sharing agreements with other agencies and provinces. In the UK, the Statistics and Registration Service Act and in Germany the Law on Statistics for Federal Purposes both secure the protection of personal information and ensure data confidentiality.

Official energy statistics advance public understanding of the energy sector and contribute to sound energy policymaking. So staying relevant and accessible are essential attributes of an effective statistical system. Typically agencies stay relevant to data users by satisfaction surveys, conferences, and other forms of consultations. In the United States, for instance, EIA ensures the utility of information through customer surveys, formal solicitations of comments, interviews and focus group studies, user conferences, and other outreach programs (EIA, n.d.-b). When he

was chief statistician at StatCan (1985-2008), Ivan Fellegi was determined to overcome the challenge of being irrelevant. He set up regular meetings for senior representatives from interested ministries to provide suggestions for improvements and for new products at StatCan (Scheuren et al., 2012).

Effective statistical systems also require coordination among agencies, adequate resources, sound methodologies, and quality control of data. Coordination is vital for mobilizing budgetary and staffing resources. That makes it easier to prioritize tasks, exploit potential synergies, improve efficiency (such as using standardized methods or tools), and ensure coherence of outputs (Fellegi, 1996). Coordination is more challenging for decentralized systems like Germany. But two or three times a year Destatis meets with representatives of statistical offices in the federal states to discuss methodological and technical details of surveys. Such meetings are more frequent if necessary for producing nationally uniform results (Kopsch, 2002). In the UK, the recent creation of BEIS by combining responsibilities of two formerly separated agencies, DECC and BIS, is seen as a strategy to enhance coordination and closely link business, industry, and science with energy and climate change.

Statistical systems need to ensure the adequacy of resources for implementation. That means having built-in procedures for evaluating and acquiring necessary funding, staffing, and other resources. In the United States, EIA submits an annual budget request with detailed justifications and its performance measures. In Canada, StatCan retains staff and advances their skills through continuous learning and selective rotations. Such practices help secure funding for operational needs, build up staff expertise and momentum, and increase creativity. An essential aspect of maintaining credibility is relying on sound methodologies and producing quality data.

Each of the four statistical systems we studied has established methodological guidelines and quality assurance measures to ensure and enhance the quality level achieved.

3.2.2 Data collection and coverage

A country's statistical model and its institutional setup determine how energy data is collected, compiled, and disseminated, as well as how energy statistics are covered. The four countries studied in this analysis have different mechanisms for collecting and maintaining energy data. In particular, energy data in these countries differ in type, frequency, coverage of geographic areas, and time series, as well as in the methodologies used to collect data. In the United States and Canada, EIA and StatCan both collect and disseminate energy information and data, including energy production, stocks, demand, imports, exports, and prices. In the UK, DECC publishes statistics relating to energy, climate change, energy efficiency, fuel poverty, radioactive incidents, and coal health. In Germany, Destatis only provides energy data on electricity and gas production; on distribution, heat generation, and energy consumption in industry; and on hard coal imports.

Across all countries, energy data are typically collected by means of administrative data, census, statistical surveys and, to some extent, by measurement and modeling (United Nations, 2013). Data collection methods are typically specific to energy data categories and data types (Table 2). Energy data in the United States² and in Canada are collected at the individual level, directly from mines, wells, refineries, traders, manufacturers and end users. But in the UK most

² In the United States, EIA collects from individual companies and cross checks that data with surveys of smaller producers, state data, and data from private information services—all on the production side. Data from the demand side is collected directly from companies by means of surveys. EIA does not directly collect data from individual wells, but gets that data from a private data service.

energy data are collected at a largely aggregate level – for example, data from mining companies instead of individual mines. The frequency of data collection also varies across countries. EIA has the most frequent collection of energy supply data and the widest data coverage among all four countries.

3.2.3 Statistical surveys

Production and consumption surveys play an important role in collecting energy data. Production surveys, directed to suppliers of specific energy sources, in general measure the quantities of specific fuels produced for and/or supplied to the market. Consumption surveys collect information on the types of energy used by consumer groups and on the consumer characteristics associated with energy use. In other words, production surveys may provide information on sales to different groups of end users, which helps in compiling data on consumption. But adding consumption surveys may provide much more detail and may confirm the accuracy of data on consumption patterns. For example, EIA collects U.S. household energy characteristics, usage patterns, and other demographic information through the Residential Energy Consumption Survey (RECS). It combines the findings with data from energy suppliers to these households. That way it is possible to estimate energy use and costs for more disaggregated end uses such as heating, cooling, and appliances (EIA, n.d.-d).

Surveys on energy supply are easier to carry out and less expensive than consumption surveys. The number of respondents in production surveys is often small and the competence of the respondents is comparatively high. In consumption surveys the population of potential respondents regarding energy usage patterns at residential and commercial buildings, for example, is huge and that population's knowledge of the topics is relatively low. Therefore,

consumption surveys are often less frequent and more expensive, and need more resources to carry out. For example, EIA keeps most U.S. production surveys on a monthly basis, while consumption surveys, such as the Commercial Buildings Energy Consumption Survey (CBECS) and RECS, are only conducted once in a few years. The capacity of interviewers is critical to energy surveys, especially consumption surveys. Consumption surveys are often conducted by contractors. To ensure that surveys are conducted correctly and that the results are reliable, it is important for statistical agencies to provide enough training for survey interviewers and to conduct quality checks. EIA is also considering recommendations by the National Academy of Sciences to conduct smaller, continuous surveys – in part to maintain the capacity of interviewers and to have less staff turnover (Marton and Eddy, 2012).

Surveys are expensive and complex, and there is clearly a trade-off among cost and quality, timeliness, and precision. The United Nations uses nine steps to describe developing and implementing energy surveys: specify needs, design, develop, collect, process, analyze, disseminate, archive, and evaluate (United Nations, 2013). Many factors contribute to the quality of survey results, including sample size, sampling methods, response rate, and the capacity of interviewers. The frequency of data collection relates to the way in which surveys are carried out. Most countries only collect monthly data from large energy producers and suppliers; then they extrapolate data for small producers and suppliers. Sample size and sampling methodologies also depend on the frequency of data collection. For example, if the same data are collected monthly and annually, the corresponding monthly survey usually has a smaller sample size.

How comprehensive energy surveys are depends on data needs and available budget and resources. In the United States, EIA distributes 64 active surveys to collect energy data (EIA, 2016). Other agencies, such as FERC and the Department of Interior, also distribute energy

surveys for their internal administrative purposes. The frequency of data collection is hourly and daily for Balancing Authority operations, but in most cases data is collected annually. Reporting from respondents is usually mandatory for production surveys.

EIA also conducts customer surveys to evaluate and improve its data services. StatCan uses 29 surveys to collect energy data (Statistics Canada, 2016b) on a monthly, quarterly, or on an annual basis. Reporting from respondents is also mandatory. In the UK, DECC circulates 29 statistical surveys to collect energy data (DECC, 2011). These surveys have sample coverage ranging from 25 percent to 100 percent; response rates range from 80 percent to 100 percent. Most data are collected on a monthly, quarterly, or annual basis. According to Destatis, in Germany there are monthly and annual data collection surveys. In each country, various methods of data submission apply. Generally, respondents can submit survey data by mail or by fax, or increasingly through the Internet.

3.2.4 Quality assurance and data validation

Ensuring data quality is a core challenge for agencies that collect and disseminate energy data. Some frameworks have been developed to improve and assure the quality of data. One such framework was developed by the United Nations Economic Commission for Europe (UNECE), Eurostat, and the Organisation for Economic Co-operation and Development (OECD) in order to manage data quality and metadata: the Generic Statistical Business Process Model (GSBPM). This model has been adopted by EIA, StatCan, and by many European statistical agencies to varying degrees (Lalor, n.d.). The GSBPM encourages a standard view of the statistical business process, but it is also a matrix with many possible combinations of steps and substeps.

Quality assurance needs to be conducted at every stage of energy data management, from survey design to data dissemination and archiving. For example, all four countries require data validation after energy data are collected or reported. Each of the four countries follows a similar trajectory for data validation (Figure 2). Energy data collected from respondents go through a number of plausibility checks before they are used to produce energy statistics. At a minimum, these checks cover the following issues: whether the data are complete in terms of quantity and content; whether the data contradict historical data; whether the data significantly differ from estimates of energy models; and whether the energy flows balance out correctly. These plausibility checks are usually based on formal checklists, logical conclusions, and experience (Statistical Offices of the Federation and the Länder, 2005). When respondents are not able to respond with correct data immediately, the statistical agencies make an estimate and then later correct it as appropriate.

3.2.5 Modeling and projection

Energy modeling is widely used by statistical agencies to collect and analyze data. Modeling can potentially reduce the frequency and complexity of data collection and validate data collected from surveys. Analyzing and interpreting data are also useful functions of energy models. For example, building energy consumption by fuel and service is critical information for demand-side management. However, fuel shares in various services are often difficult to obtain from household surveys because most households do not have submetering; in that case, modeling can provide such information.

Energy modeling also provides an indication of how the energy sector might be in the future. In the United States, EIA produces the Annual Energy Outlook (AEO), a publication

based on results from the National Energy Modeling System. It focuses on factors that shape the U.S. energy system over the long term. EIA produces the publication annually to discuss selected energy issues with updated assumptions about socioeconomic development, specific technologies, and policies. AEO also provides the basis for strategic planning on energy and it contributes to the evolution of national energy policy making.

3.2.6 Data dissemination

To remain relevant to stakeholders, statistical agencies typically disseminate energy data in a consistent and user-friendly way. Effective and timely data dissemination will benefit policymaking, private and public investment decisions, and strategic planning at both the national and regional levels.

All four countries we assessed disseminate a wide range of historical energy data. The United States has the most consistent time series. Nearly all annual energy data reported by EIA start from 1949; most monthly data start from the 1970s or the 1980s. In all four countries, projections of future energy supply and demand are also publicly accessible. Germany's energy forecasts are less frequent compared to the other three countries, where such analysis is updated annually. But all four countries tend to provide analysis on potential trends for the next 20 to 35 years. The institutional setup for data dissemination can either be centralized (the U.S. and the UK) or decentralized (Canada and Germany) (Figure 1). In the United States, EIA aggregates and disseminates energy data at various levels, including state, census region, and national. EIA also provides detailed state energy statistics profiles, which are beneficial for state and local policymaking as well as for interstate comparisons. In Canada, StatCan and other agencies usually report data at the provincial and national levels. In the UK, DECC publishes energy data

at the national level, with some subnational details in the consumption statistics. In Germany, energy data are mostly available in a nationally aggregated form, while some data are organized at the federal state level.

All four countries make data available for viewing and downloading for free. Energy data in the United States are fully available to the public. The availability of energy data is limited to some extent in the other countries. Canada does not disseminate detailed renewable energy and nuclear/uranium data in the public domain, but these data are available on request. According to the UK Energy Research Centre, nuclear statistics and related information in the UK are managed by the Dalton Nuclear Institute, but they are not publicly available. In terms of format, energy data can be disseminated both electronically and in official statistical publications.

Dissemination formats should meet the needs of stakeholders. For example, press releases on energy statistics need to be organized in a way that is easy to re-disseminate by mass media. Comprehensive energy data, meanwhile, can be posted online in multiple electronic formats (PDF, XLS, CSV or interactive graphs/tables, etc.). The variety of data products matters in effectively communicating information. Most countries develop annual energy reviews or balances in order to summarize information on energy supply, demand, transformation, and efficiency. Countries such as the United States also provide monthly data on the energy market and analyze changes in market trends, which is critical for business planning and investment in the near term. There are also press releases and news updates (EIA's *Today in Energy*, for example) that discuss specific energy issues and communicate information in a timely manner. In addition to data, statistical releases are also accompanied by other materials, including survey or product names, survey objectives, time frame, concepts and definitions, target population, collection method, survey forms, sample size, sampling methodologies and response rates, error

detection, missing data, revisions, analytical methods used, provisions regarding confidentiality, and disclosure control and response methods.

To meet evolving stakeholder needs, countries are improving data dissemination by leveraging technology. Internally, statistical agencies are streamlining their reporting systems and improving data storage (for example, by moving towards cloud storage). Countries are also extending data formats, enhancing data accessibility, and improving visualization. For example, EIA has interactive data browsers that allow users to customize the time series to be viewed or plotted and that give users options for charts and maps. In addition, countries are diversifying data products to meet evolving needs. EIA is collecting hourly electric power data from all but two of the 67 respondent Balancing Authorities, for example. Such data would supplement the existing biweekly, monthly, and annual collections of electricity data.

3.3 Assessing national energy data management: an integrated framework

In reviewing and analyzing approaches in these four countries, we here present an analytical framework for the assessment of national energy data management (Figure 3). This framework covers critical elements of effective data management from both the institutional and operational perspectives. Since energy statistics are part of a country's national statistics, energy data management must comply with the guidelines and standards of national statistics more generally. Thus, the institutional model for energy data management is often heavily influenced by the national statistical system, which in turn depends on multiple factors such as economic structure, a country's size, and its type of government.

As demonstrated in the four countries, energy data management can either be situated within a centralized statistical agency or multiple agencies that concentrate on certain aspects of

energy. No single institutional structure is perfect, but these four countries share some commonalities in addressing challenges within their operations. A robust energy data management system has multiple dimensions: relevance to policies, the credibility and objectivity of data, accurate methodologies and data products, coherence in time series and data, accessibility, and user friendliness. End users of energy data in these four countries are also similar, and generally include policymakers, industry, nonprofits and international organizations, academia, media, and the public. The U.S. and the UK systems are very user friendly since they integrate all aspects of energy data on a single platform. But such systems typically require an adequate amount of dedicated resources upfront. The Canadian and German systems have their roles and responsibilities distributed across agencies and they require relatively more coordination to avoid overlapping.

Climate concerns are driving many countries to enhance their energy data in order to better plan for low-carbon growth. On the other hand, many developing countries lack the extensive energy data that are key to low-carbon growth planning. We developed this analytical framework based on the commonalities in the four energy data management systems we studied, yet challenges of energy data management remain in these four countries as well. Examples of some common issues include coordination, data quality and methodologies, and data confidentiality, as we discussed in the previous sections. However, how countries address specific challenges historically and at present is not the main focus of this study. Any country can use this framework to evaluate its institutional model for national energy data management and to take steps to make improvements. Each country may face different challenges under its own situation, and future research on this topic can focus on international best practices on specific issues within the context of a country that is being assessed. The next section of this

paper describes a case study of how India could use our framework to analyze and improve its existing statistics.

3.4 India's national energy data management: a case study

India has multiple agencies collecting energy data, and therefore coordination can be challenging across these agencies and jurisdictions. Managing energy data is often challenged by inconsistency in data compilation, the use of incompatible definitions and formats, and by the existence of data gaps. A detailed assessment of energy data management using the proposed framework can highlight options for improving the institutional structure and processes for managing India's energy sector and to identify areas of improvement in order to overcome the challenges within the system.

3.4.1 Organization of India's energy data management

Unlike the United States and the UK, India's energy data management is fairly decentralized (Figure 4). Institutions involved in the management of India's energy data fall into three groups. First, there are energy-related line ministries and their subordinate agencies or offices. These include the Ministry of Power (Central Electricity Authority and Bureau of Energy Efficiency), the Ministry of Coal (Coal Controller's Organisation), the Ministry of Petroleum and Natural Gas (Petroleum Planning and Analysis Cell, the Directorate General of Hydrocarbons and Petroleum Conservation Research Association), the Ministry of New and Renewable Energy, and line ministries dealing with major energy consuming sectors such as agriculture, industry and transport. Second, India has statistical agencies such as the Ministry of Statistics and Programme Implementation (MoSPI) and state-level Directorates of Economics and Statistics.

Third are India's planning institutions, such as the National Institution for Transforming India (NITI Aayog) and state-level planning departments. There is no one central institution solely responsible for collecting, analyzing, or disseminating all of the country's energy data. Instead, data are collected by individual line ministries for their own administrative purposes and disseminated as per policies and practices that have evolved over time within each ministry. In addition, some areas of the energy sector (electricity, for example) are also under state jurisdiction. Given the decentralized nature of India's energy data management, coordination is critical to ensuring consistency within the system.

3.4.2 Operations of India's energy data management

Fundamental principles. India has enacted various legal provisions over the years that empower government agencies to collect various kinds of data, including energy data with relatively weaker mandate for dissemination. Yet there is no robust mechanism for assessing the relevance of data or for incorporating feedback from data users in both the public and private sectors. MoSPI currently serves as the nodal agency for coordination, but it has inadequate authority and expertise for this role. It cannot gain access to many types of data, and it cannot require ministries to coordinate statistical survey methodologies (National Statistical Commission, 2011). In its report on data management, the National Statistical Commission also recognized that the lack of resources for sample surveys has made it challenging to expand existing surveys or to introduce new surveys (National Statistical Commission, 2011).

Data collection. Consistent with international practices, India's energy data are collected through administrative processes, regulatory proceedings, the census, and statistical surveys.

These data cover various fuel types as well as a number of processes: production, imports, transmission and distribution, and end use (Figure 4). Most end-use surveys are conducted by statistical branches within MoSPI. Some surveys conducted by the Ministry of Health and Family Welfare also contain energy-relevant information. However, none of these surveys focuses exclusively on energy. Instead, energy-related questions are embedded within surveys that focus on other issues, such as socioeconomic or health data. The majority of publicly available energy data is focused on the supply side; a limited amount of consumption data is available through the census and the National Sample Survey Office at MoSPI, although such consumption data are usually collected from energy producers and suppliers. In addition, energy data collected through surveys are spread over many questionnaires meant to answer specific questions but not to inform energy sector analysis in general. Moreover, such surveys may be conducted during different time periods, making it difficult to access coherent energy data. There are no systems in place to gather data from these disparate sources into a single, coherent whole.

Quality assurance. Individual government entities like PPAC have been increasing their efforts to address quality issues and to promote practices of evaluation and quality assurance. Yet there is no shared approach that spans the whole system. Given the current organizational structures, line ministries typically do not have statistical expertise and statistical organizations do not have topical expertise. This potentially creates a challenge for both data collection and quality checks during the compilation of energy statistics. There are also no known processes to actively seek feedback and inputs from data users such as researchers, investors, and others.

Data dissemination. Data handling within line ministries and regulatory agencies in India is often not geared towards public dissemination but rather towards administrative purposes. Data users may also find it difficult to access energy data through typical data-relevant publications such as annual reports and regulatory orders. The objective of such publications is not primarily to disseminate data. In most cases, dissemination of quality data is likely to involve additional budgets, personnel and training.

3.4.3 Recommendations for improvement

As in India, many countries are seeing rapid growth in their energy sectors in sync with population growth and urbanization. Many of these countries also do not have a long history with energy data management, or adequate experience. This makes effective policymaking a challenge in the energy sector. Our framework provides options for countries like India to strengthen their energy data management.

Consistent with its political and administrative structure, India's energy data management system is decentralized in nature. It has some weaknesses in data collection, reconciliation and dissemination as discussed in the previous section. Keeping in view the current institutional setup and processes in India, it is pragmatic to take an evolutionary approach to overcoming these challenges, beginning with near-term improvements to current operations leading to longer-term organizational reforms. For example, while an independent statutory agency for collecting, analysing and disseminating energy data, such as the EIA, can be effective in addressing these challenges, such an agency may not be appropriate or even desirable for India where data agencies within multiple ministries are already collecting and disseminating energy data. However, some form of institutionalised coordination mechanism between these data

agencies can greatly help to improve quality and consistency of energy data across multiple jurisdictions and frequencies. The following recommendations reflect this approach.

1. **Establish common core principles.** This is especially important for developing standardized definitions, classification, and data quality standards. Examples from the principles and codes of practices from the U.N. Statistical Commission, the United States, and others can help countries like India strengthen their energy statistics by framing a common set of principles for all organizations involved in energy data management.
2. **Enhance coordination.** Agencies need to coordinate their activities in order to maximize the use of existing data and to improve data quality and relevance. In this regard, setting up a suitably empowered cell or agency to reconcile data from various sources and to exchange data management experiences would help. This would be somewhat similar to the roles played by DECC or Destatis which coordinate and reconcile data in a decentralized system.
3. **Prioritize data improvements.** Data on energy consumption and non-commercial energy sources is currently inadequate in many developing countries. Collecting such data needs to be prioritized, since it is critical to planning for inclusive and low-carbon development.
4. **Adopt technology and statistical methods.** To increase efficiency in data collection, countries should use common formats for similar data providers, where applicable, and should automate data collection to ensure accuracy and completeness. Recent technological advances make it easier to collect and disseminate data. Examples include automated metering, online data submission,

and interactive data dissemination. Adopting such technologies can improve efficiency and reduce manual processing errors. India could also consider adopting sophisticated statistical approaches that can help collect accurate data at lower cost through sampling. Most of the developed countries studied as part of this work use such approaches to reduce the costs of data collection.

5. **Introduce data quality assurance mechanisms.** These make it possible to implement systematic quality checks. Ensuring data quality is a core challenge for agencies that collect and disseminate energy data. Quality assurance needs to be conducted at every stage of energy data management. All four countries studied undertake data validation exercises after collection, and have other mechanisms to continually improve data quality through measures such as user conferences (EIA), regular meetings across various concerned agencies (StatCan and DeStatis) and so on.
6. **Institute continuous training and strategic staff planning.** This maximizes the effective use of staff at various levels. Adequate staffing and financial support also needs to be ensured for energy data management.
7. **Improve data relevance.** This is accomplished by working closely with stakeholders. Greater relevance of data to policymaking, analysis, and research can also help strengthen data-related budgets.
8. **Improve data dissemination.** Data dissemination in convenient formats needs to be prioritized and treated on par with data collection for administrative purposes. All four countries studied freely disseminate the collected data in timely, easy-to-

access formats. This requires better data integration, uniform dissemination standards, customer-oriented data formats, and better use of modern technology.

4. Conclusions and Policy Implications

Driven by climate concerns, many countries have started to enhance their energy data in recent years in order to better plan for low-carbon growth. Developing countries have a stronger demand for better energy data management, but developed countries also have room for improvement. Such improvements may include expanding coverage, doing additional quality checks, and more.

This study reviews energy data management in five countries – the United States, Canada, the UK, Germany, and India. All five countries have adopted different organizational models for energy data management. Despite differences in these models, countries with effective energy data management systems follow some common principles. These include making data relevant to policymakers and users; establishing the legal status of data and its independence from outside influences; data coordination; the implementation and adequacy of data management resources; and assuring the quality of data and methodologies.

In addition, statistical agencies require a variety of expertise, including statisticians, modelers, and topical experts in different energy-related fields. That demand for expertise needs to be considered in managing personnel and it must be maintained with continuous training and capacity building. Energy data management systems usually parallel a government's overall approach on the spectrum of being either more centralized or more decentralized. In countries with decentralized data systems, extra coordination and communication are essential to maintaining data quality.

This study also assesses energy data coverage in these countries from both the perspective of data collection and data dissemination. Administrative and census data, statistical surveys, modeling, and on-site measurement are used to collect a wide range of energy data. This includes data on energy production, transformation, consumption, energy efficiency, and other energy-related indicators. A new system design needs to consider some key aspects of data coverage. These include categories of data, frequency of data collection and dissemination, geographic areas covered, and the time series of data. The United States has the most extensive and comprehensive energy data, with the longest time series, the most subnational details, and the most frequent data dissemination.

Learning from the experiences of these countries, governments may wish to emphasize five key areas in developing a robust energy data management system: improving consumption data; strengthening coordination; prioritizing energy data; ensuring data quality; and enhancing access to data.

Although consumption data are critical for designing energy and social policies, in many countries they are more dispersed than production data. To improve consumption data, statistical agencies should enhance the coverage and frequency of consumption surveys based on data needs, while at the same time coordinating with existing surveys and information collected by other agencies. Statistical systems require coordination among institutions to achieve greater efficiency and coherence. Coordination is more challenging in the case of decentralized energy data management systems, although centralized systems also require coordination. Government agencies are constrained by budget and resources. Meanwhile, it is critical to prioritize data collection and reporting based on stakeholder needs, major trends in the energy sector, and gaps in existing data.

There is a need to continually evaluate the quality of data being collected, processed, and disseminated in order to make sure it is relevant to data users. This is particularly true of energy data because the energy sector undergoes rapid transformation technologically and structurally and because such data is increasingly augmented by renewable energy systems, which are often decentralized. All of these data quality improvements, of course, are best made in close collaboration with data users.

And what is data without making it known? Data dissemination is as important as data collection. Effective and timely data dissemination ensures the relevance of the statistical agency and it contributes to policy development and decision-making. All four developed countries analyzed in this study provide energy data free of charge. They also constantly improve data dissemination in order to diversify data products and formats and to provide maximum access to energy data.

While this study considered India as a case study to verify the applicability of the broad framework, the underlying principles of the framework should be equally applicable to all developing countries though it would need to be customized to the context of each country.

Acknowledgements

The research was supported by the United States Agency for International Development and the National Renewable Energy Laboratory. The Pacific Northwest National Laboratory is operated for the United States Department of Energy by Battelle Memorial Institute under contract DE-AC05-76RL01830. Prayas (Energy Group) thanks NITI Aayog for the opportunity to be part of the Sustainable Growth Working Group.

References

United States Department of Energy Organization Act of 1977, Pub.L. 95-91, 42 U.S.C. 7135.

Canada Statistics Act of 1985, Pub.L. R.S.C., c. S-19.

DECC, 2011. DECC Statistics Surveys. Retrieved from:

<https://www.gov.uk/government/publications/statistical-surveys> (accessed: May 18, 2016).

DECC, 2015. Department of Energy & Climate Change Annual Report and Accounts 2014-15.

Retrieved from:

https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/447144/6.656_DECC_JG_Annual_Report_2014-15_AW_WEB.pdf (accessed: April 22, 2016).

EIA, 2016. Survey Forms. Retrieved from: <https://www.eia.gov/survey/> (accessed: May 18, 2016).

EIA, n.d.-a. Budget and Performance. Retrieved from:

https://www.eia.gov/about/budget_performance.cfm (accessed: March 9, 2016).

EIA, n.d.-b. Information Quality Guidelines. Retrieved from:

http://www.eia.gov/about/information_quality_guidelines.cfm (accessed: September 10, 2014).

EIA, n.d.-c. Mission and Overview. Retrieved from:

https://www.eia.gov/about/mission_overview.cfm (accessed: March 9, 2016).

EIA, n.d.-d. Residential Energy Consumption Survey (RECS). Retrieved from:

<http://www.eia.gov/consumption/residential/about.cfm> (accessed: March 9, 2016).

Fellegi, I.P., 1996. Characteristics of an effective statistical system. Canadian Public Administration 39, 5-34.

IEA, 2015. World Energy Outlook 2015 Special Report on Energy and Climate Change.

International Energy Agency, Paris, France. Retrieved from:

<https://www.iea.org/publications/freepublications/publication/WEO2015SpecialReportonEnergyandClimateChange.pdf> (accessed: January 18, 2016).

IPCC, 2014. Summary for Policymakers, in: Field, C.B., Barros, V.R., Dokken, D.J., Mach, K.J., Mastrandrea, M.D., Bilir, T.E., Chatterjee, M., Ebi, K.L., Estrada, Y.O., Genova, R.C., Girma, B., Kissel, E.S., Levy, A.N., MacCracken, S., Mastrandrea, P.R., White, L.L. (Eds.), Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom, and New York, NY, USA, pp. 1-32.

Kopsch, G., 2002. The role and function of regional and local statistical offices; interactions with regional and local authorities. United Nations Statistics Division, Yangon, Myanmar.

Lalor, T., n.d. National Implementations of the GSBPM – A Summary Based on METIS Case Studies.

Marton, K., Eddy, W.F., 2012. Effective Tracking of Building Energy Use: Improving the Commercial Buildings and Residential Energy Consumption Surveys. National Academies Press, Washington, DC.

National Statistical Commission, 2011. Report of the Committee on Data Management. New Delhi, India. Retrieved from:

http://mospi.nic.in/Mospi_New/upload/finalreportonData%20management01082011.pdf

(accessed: June 19, 2016).

Scheuren, F., Willig, B., Katz, L., Rae, M., Maki, B., Lee, D., Hayes, W., Lee, C., Dehaan, R., Mushtaq, A., 2012. Profiles in Statistical Courage. George Washington University, Washington, DC.

Statistical Offices of the Federation and the Länder, 2005. Quality Standards in German Official Statistics. Retrieved from: <http://unstats.un.org/unsd/dnss/docs-nqaf/Germany-BrochureQualityStandards.pdf> (accessed: September 10, 2014).

Statistics Canada, 2016a. 2014-15 Departmental Performance Report for Statistics Canada. Retrieved from: <http://www.statcan.gc.ca/eng/about/dpr/2014-2015/index> (accessed: April 5, 2016).

Statistics Canada, 2016b. Complete List of Surveys and Statistical Programs on Energy. Retrieved from: http://www23.statcan.gc.ca/imdb/pIX.pl?Function=getThemeSV&PItem_Id=97413&PCE_Id=384&PCE_Start=01010001&CItem_Id=97413&CCE_Id=384&CCE_Start=01010001&lang=en (accessed: May 18, 2016).

Statistics Canada, 2016c. A Statistical Framework for Energy in Canada. Retrieved from: <http://www.statcan.gc.ca/pub/57-602-g/57-602-g2015001-eng.htm> (accessed: April 5, 2016).

Statistisches Bundesamt, n.d. Our Mission: Statistics. Retrieved from: <https://www.destatis.de/EN/AboutUs/OurMission/Statistics/Statistics.html> (accessed: April 26, 2016).

United Nations, 2013. Energy Statistics Compilers Manual. Retrieved from: <http://oslogroup.org/index.asp?page=escmmainpage.html> (accessed: February 16, 2016).

UNSD, 2011. International Recommendations for Energy Statistics. United Nations, New York, USA. Retrieved from: <http://unstats.un.org/unsd/statcom/doc11/BG-IRES.pdf> (accessed: January 19, 2015).

UNSD, 2015. Energy Statistics Section. Retrieved from: <http://unstats.un.org/unsd/energy/default.htm> (accessed: January 19, 2016).

Table 1: Principles of Official Statistics

	Principles	UN	EU	US
a.	Proximity and relevance to policy issues and topics	✓	✓	✓
b.	Independence from political and other undue influences	✓	✓	✓
c.	Trust among data providers and confidentiality	✓	✓	✓
d.	Public perception and credibility	✓	✓	✓
e.	Timeliness and punctuality	✓	✓	✓
f.	Impartiality and objectivity	✓	✓	✓
g.	Cost effectiveness	✓	✓	✓
h.	Non-excessive burden on respondents	✓	✓	✓
i.	Availability, accessibility, and clarity to the public	✓	✓	✓
j.	Quality of data, products, methodologies, and procedures	✓	✓	✓
k.	Coherence and comparability		✓	✓
l.	Adequacy of resources		✓	✓
m.	Mandate for data collection		✓	✓
n.	Coordination among various agencies/branches for consistency and efficiency	✓		
o.	Use of international concepts, classifications, and methods	✓		
p.	Bilateral and multilateral cooperation	✓		

Table 2: Data coverage and collection methods.

Data category	Collection methods	Examples	Characteristics
Energy supply	Administrative data, census ³ , surveys, modeling	Imports/export, stock, energy production, transformation output, deliveries to the market	Data sources are a few large producers / industries / agencies. Fossil fuels, primary heat and renewables often have different methodologies.
Energy transformation	Census, surveys	Transformation input/losses, transportation output	The number of data owners is relatively small.
Energy consumption	Surveys, business data from energy industries, modeling	Final energy consumption by end users, transformation	Consumption surveys have a large population and are

³ Census here means that the data was collected by sending questionnaires to 100% of the organizations in a given category.

		inputs	expensive and time-consuming.
Energy prices	Census, surveys	Expenditures, costs, and taxes	Price data can be obtained from suppliers/traders or consumers. Price data for renewables are limited. Taxes and incentives data are often not included in the energy data system.
Energy efficiency and indicators	Modeling, surveys, metering/measurement	Energy consumption by service, per capita or per unit of floor space/output	Energy indicators are often estimated by assembling different datasets, which presents challenges in terms of data compatibility.
Energy poverty	Administrative data, surveys, modeling	Energy expenditure relative to income	Data are often obtained from consumption surveys and case studies.

Figure 1: Energy data management models

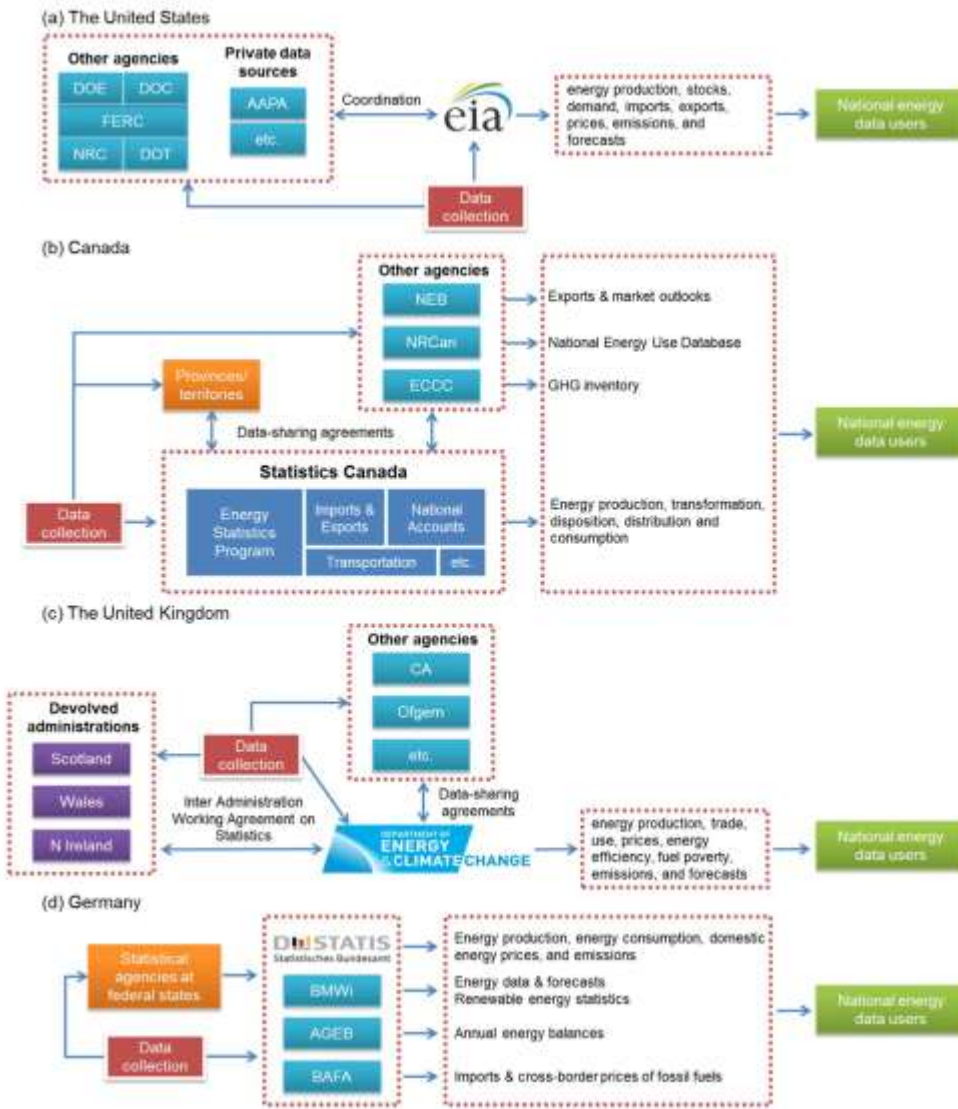


Figure 2: Common processes for data validation

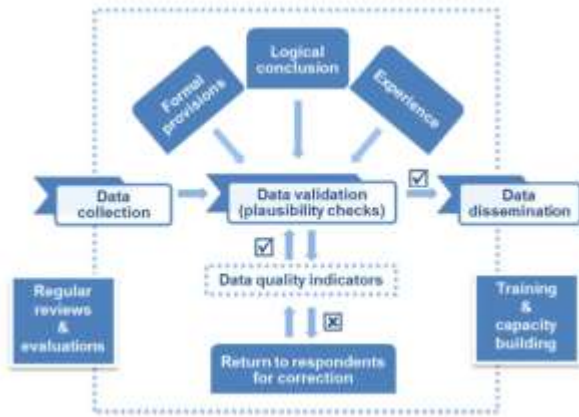


Figure 3: The analytical framework for the assessment of national energy data management

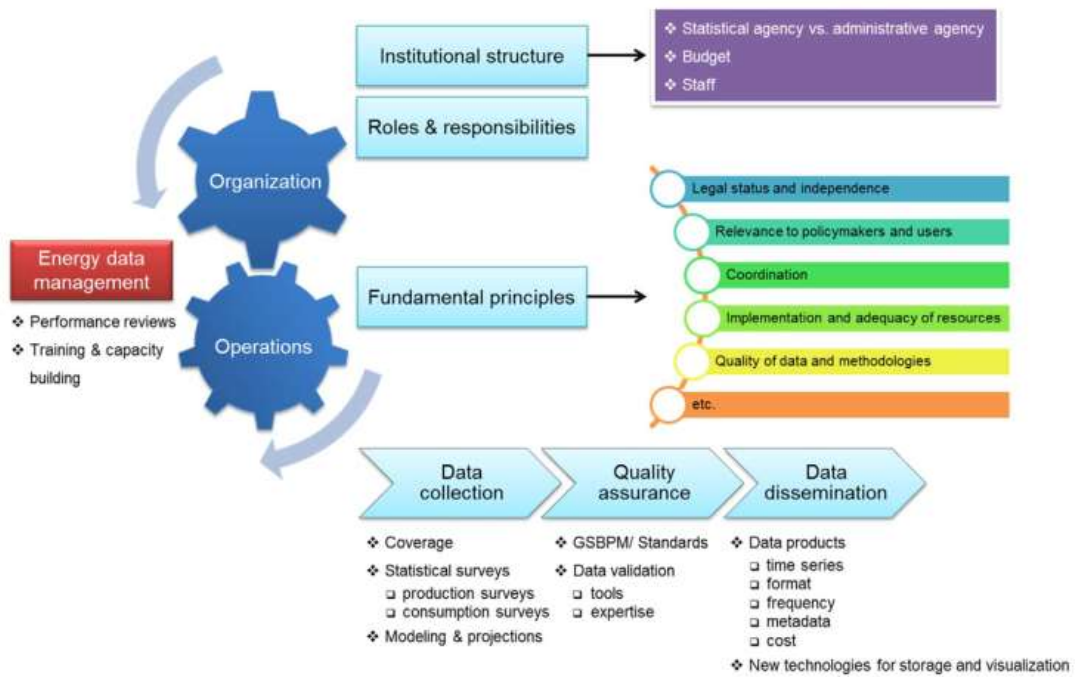
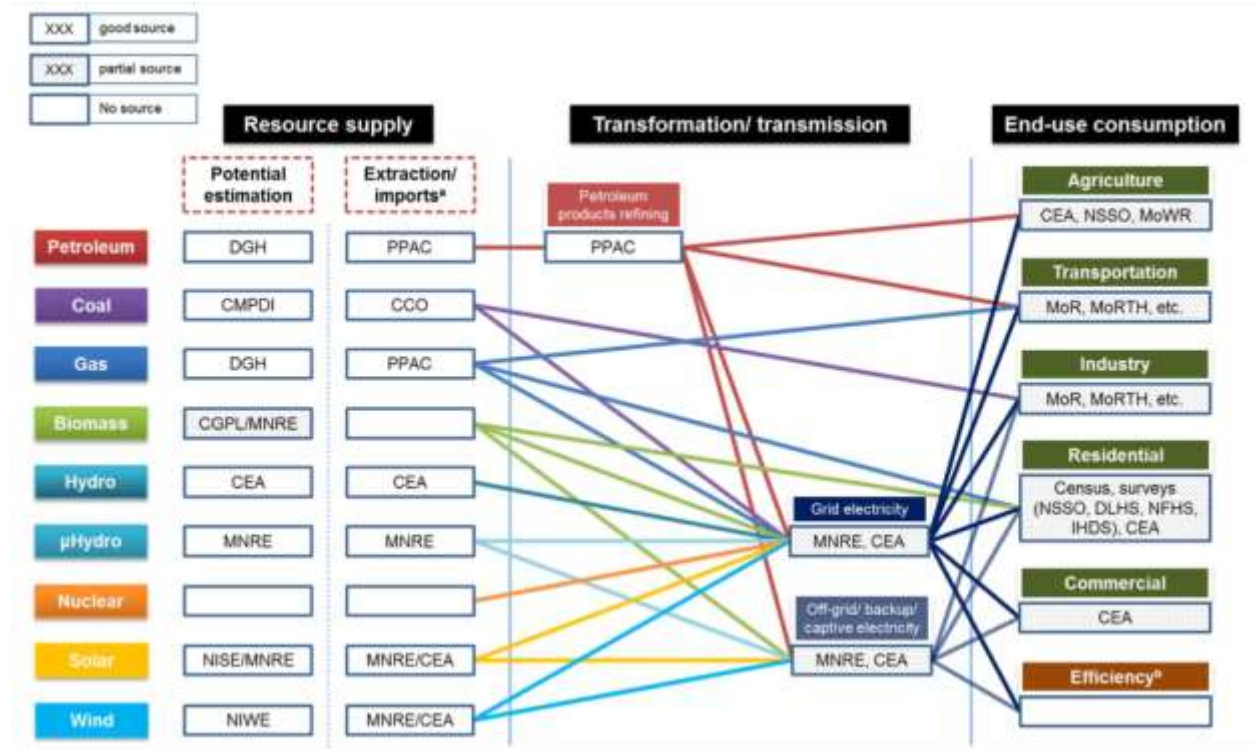


Figure 4: Government agencies involved in India's national energy data management.



Notes:

- Imports include direct resource imports and the import of final energy sources
- Efficiency includes appliance and agricultural pump use, building performance, transport fleet efficiency etc.