Guidebook for Improved Electricity Access Statistics

International Energy Agency

INTERNATIONAL ENERGY AGENCY

The IEA examines the full spectrum of energy issues including oil, gas and coal supply and demand, renewable energy technologies, electricity markets, energy efficiency, access to energy, demand side management and much more. Through its work, the IEA advocates policies that will enhance the reliability, affordability and sustainability of energy in its 31 member countries, 11 association countries and beyond.

IEA member countries:

Australia
Austria
Belgium
Canada
Czech Republic
Denmark
Estonia
Finland
France
Germany
Greece
Hungary
Ireland
Italy
Japan
Korea
Lithuania
Luxembourg
Mexico
Netherlands
New Zealand
Norway
Poland
Portugal
Slovak Republic
Spain
Sweden
Switzerland
Republic of Türkiye
United Kingdom
United States

The European Commission also participates in the work of the IEA

IEA association countries:

Argentina
Brazil
China
Egypt
India
Indonesia
Morocco
Singapore
South Africa
Thailand
Ukraine

Source: IEA.
International Energy Agency
Website: www.iea.org

Abstract

The world remains far off track to reach universal access to electricity by 2030 – a key target canonised in UN Sustainable Development Goal 7 (SDG7). While many countries have set targets to advance universal electrification, it remains challenging to collect timely data on their progress, making it difficult to build or refine their electrification plans and policy strategies.

To provide governments with the tools to achieve their goals, the *Guidebook for Improved Electricity Access Statistics* (hereafter “Guidebook”) focuses on methodologies using readily available supply-side data from electric utilities, mini-grid operators, and off-grid system distributors to track access to electricity trends. This approach can be adopted at a low-cost and provides accurate estimates on access rates with a potential time lag on the order of months, instead of years. Supply-side data complements household surveys or censuses, which can give a more nuanced, detailed picture, but are typically only run every five to ten years due to their expense.

This report provides step-by-step guidelines on how to implement an improved supply-side data collection process, and produce more time-sensitive and comprehensive access to electricity indicators. In addition, the Guidebook proposes a standard method to capture mini-grids and stand-alone off-grid systems in calculating access to electricity, which are becoming increasingly more commonplace in strategies to reach universal access. Finally, the Guidebook suggests further steps practitioners can take to advance geospatial data collection and reporting, which is the best practice for electrification planning supporting both electric utilities and the vibrant off-grid industry.

Acknowledgements, contributors and credits

This study was prepared by the World Energy Outlook team in the Directorate of Sustainability, Technology and Outlooks in co-operation with other directorates and offices of the International Energy Agency (IEA). The study was designed and directed by **Laura Cozzi**, Chief Energy Modeller and Head of Division for Energy Demand Outlooks. The lead author and co-ordinator of this study was **Gianluca Tonolo**. **Daniel Wetzel** provided essential review and co-ordination support. **Pablo Freund**, **Jennye Greene**, and **Lisa Boone** from Bare Maximum LLC were among the principal external authors. Additional support was provided by **Roberta Quadrelli** and **Zakia Adams** from the IEA data centre, as well as by **Nouhoun Diarra** and **Darlain Edeme** from the Tracking Sustainable Transitions Unit. Diane Munro carried editorial responsibility.

This Guidebook is part of the IEA's series of statistical manuals and data tools, and builds on our work of enabling capacity building activities, including bilateral and regional training support to energy statisticians on an ongoing basis.

This report was made possible through support provided by the Power Africa initiative, led by the U.S. Government, and the U.S. Agency for International Development, under the terms of Award No. 720-674-20-IO-00001. The opinions expressed in this report are those of the authors and do not necessarily reflect the views of the U.S. Agency for International Development.

This report benefited greatly from comments and feedback by many government officials and international experts who reviewed preliminary drafts of the report. They include:

Hary Andriantavy	African Association for Rural Electrification (Club-ER)
Martin Asiimwe	Ministry of Infrastructure, Rwanda
Juliette Besnard	World Bank
Angelina Bombe	Ministry of Mineral Resources and Energy, Mozambique
Bryan Bonsuk Koo	World Bank
Elijah Chibwe	Ministry of Energy, Zambia
Salim Chitou	Economic Community of West African States (ECOWAS)
Arthur Contejean	International expert
Mamadou Diouf	Ministry of Petroleum and Energy, Senegal
Yiheyes Eshetu	Ministry of Water and Energy, Ethiopia

Francesco Fuso Nerini	KTH Royal Institute of Technology
Assane Gueye	Ministry of Petroleum and Energy, Senegal
Peace Kaliisa	Ministry of Infrastructure, Rwanda
Sjef Ketelaars	GOGLA
Shweta Kulkarn	Prayas Energy Group
Nicolina Lindblad	Sustainable Energy for All (SEforALL)
Mungai Kihara	Ministry of Energy, Kenya
Isaac Kiva	Ministry of Energy, Kenya
Félix Kouame Kouadio	Ministry of Mines and Energy, Côte d'Ivoire
Amadou Makhtar Sarr	Ministry of Petroleum and Energy, Senegal
Tom Maruti	Ministry of Energy, Kenya
Samson Masebinu	Power Africa
Emi Mizuno	Sustainable Energy for All (SEforALL)
Todd Moss	Energy for Growth
Steve Muamba	National Rural and Peri-urban Electrification and Energy Services Agency (ANSER), Democratic Republic of the Congo
Sreekumar Nhalur	Prayas Energy Group
Devotha Nshimiyimana,	Ministry of Infrastructure, Rwanda
Abdoulaye Oueddo	African Energy Commission (AFREC)
Jiyun Park	World Bank
Steven Payma	Rural and Renewable Energy Agency (RREA), Liberia
José Ignacio Perez-Arriaga	MIT Comillas
Jem Porcaro	Sustainable Energy for All (SEforALL)
Elisa Portale	World Bank
Oliver Reynolds	GOGLA
Dana Rysankova	World Bank
Andreas Sahlberg	KTH Royal Institute of Technology
Thomas Samuel	Moon
Jessica Stephens	Africa Minigrid Developers Association (AMDA)
Liliana Suchodolska	Paris21
Fatou Thiam Sow	Ministry of Petroleum and Energy, Senegal
Peter Thobora	Ministry of Energy, Kenya
Gbaty Tiadja Gbandey	African Energy Commission (AFREC)
Patrick K. Tonui	GOGLA
Emmanuel T. Ottamnan	African Association for Rural Electrification (Club-ER)
Susie Wheeldon	GOGLA

The individuals and organisations that contributed to this study are not responsible for any opinions or judgements it contains.

Table of contents

Introduction

The IEA maintains the world's most comprehensive database on energy and related statistics, including data on access to electricity. The IEA was the first international agency to start tracking global electricity access data in the early 2000s and since then also provides latest projections in the *World Energy Outlook*. The IEA is also one of the co-custodians of tracking progress on UN Sustainable Development Goal 7, or SDG7. As of 2022, The IEA estimates 775 million people worldwide are without access to electricity, of which 600 million live in Africa. Since 2014, Africa has made steady progress in reducing the number of people without access to electricity, but the Covid-19 pandemic and the 2022 energy crisis have now reversed this trend.

To accelerate and strengthen efforts, more sustainable foundations are needed for tracking access data. Timely data on progress towards universal electrification is essential for policy makers to plan subsequent steps in their electrification campaigns and to adjust as needed. Of particular importance is improving the disaggregation of technology, notably a split between connections by grid, mini-grid, or off-grid systems. Additionally, geographic information systems (GIS) are also becoming the de facto standard for access planning, emphasising the need to gather data with location specificity.

Today, countries that track access to electricity use either data on connections from electric utilities and other energy distributors (supply-side data) or household surveys and censuses (demand-side data), or a combination of the two. Both supply- and demand-side data have their benefits. Demand-side data collected through periodic household surveys or censuses provides essential user information but they are costly to conduct so are typically administered only every five to ten years. As a result, the long-time lag amid ongoing changes in policies, technologies and economic landscapes leaves decision makers with little current information to gauge progress. Supply-side data, by contrast, is more readily available within a relatively short time frame, and at a lower cost, advantages that are critically important for decision makers aiming to accelerate electrification plans. When carried out in parallel with surveys, supply-side data can provide timely estimates of progress towards universal access.

To capture the benefits from both approaches, the IEA recommends that governments or other administrative entities tracking access to electricity establish a data strategy that leverages the synergies between supply- and demand-side data sources to improve electrification planning.

The IEA's *Guidebook for Improved Electricity Access Statistics* lays the foundations to develop access indicators using supply-side data, covering all steps in the process including data collection, validation, processing and dissemination. It also details procedures to increase the quality, granularity, availability, comparability, and standardisation of access indicators.

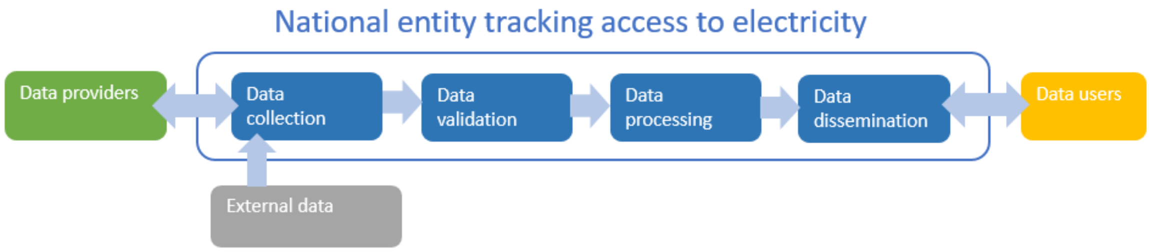
The Guidebook additionally focuses on the most significant barriers countries encounter when tracking access using supply data and proposes best practices to overcome these challenges. Notably, it details a methodology to estimate new access provided by stand-alone power systems, which are increasingly becoming a staple in many electrification programmes across Africa. By relying on sales/distribution data and adjusting for average product lifetimes, use and re-sale, appropriate estimates can be developed in a consistent, transparent manner, and calibrated against recent survey data.

Finally, the Guidebook highlights how to use supply data sources to start developing a geographical disaggregated access information system by jurisdiction or other regional designations, which is important for moving towards better informed GIS electrification planning.

The Guidebook, however, does not cover in detail how to track access of non-residential customers such as public and commercial buildings, or other key productive uses but similar approaches can be applied. This remains an important topic and the IEA plans to cover it in subsequent editions. New access solutions continue to emerge in the last decade many would not have envisioned the advances made in off-grid technologies. Accordingly, the IEA plans to also issue updates to the Guidebook as new access solutions emerge.

The guidelines have been released alongside the IEA's template for collecting relevant supply-side data. The template is used by the IEA to collect access to electricity data from governments and includes step-by-step instructions for how to report the data. It also provides embedded standard formulas used to process raw data into relevant access indicators. The template can be adapted for use as the foundation for countries in their own collection efforts, including data by jurisdiction then aggregated for national numbers.

Access to electricity data workflow



IEA. CC BY 4.0

The template on its own, however, cannot capture some of the nuanced challenges faced when validating and processing data, including making country-specific determinations on how to manage issues of double-counting, adjusting to country definitions, and how to fill data gaps. The guidelines outlined in this report focus on these important details, as well as provide clear definitions and explanations of the key principles of a supply-side data approach.

The Guidebook is divided into three chapters:

1. **Key concepts** covers all relevant definitions related to electricity access, from how to determine a standard level of minimum access as benchmark for reaching universal targets to defining connection technologies and power systems, such as mini-grids and stand-alone systems, and how to report them in calculating access to electricity. The chapter also discusses data sources used and the various methodologies for counting access to electricity, including their advantages and drawbacks.

2. **Overview of electricity access tracking** provides an outline on how to set up an initial process, and a blueprint of the data workflow, from the collection to the dissemination of these indicators. Recommendations for implementing a sustainable access to electricity data strategy are also included, from the implementation of administrative steps to stakeholder engagement.

3. **Walking through the process: from data collection to dissemination** goes through each stage involved in producing access to electricity indicators and provides recommendations on how to manage difficult issues. From developing more robust statistics methods, including greater granularity and the use of metadata (i.e. data about data), and step-by-step guidelines for collecting, validating, processing and disseminating access to electricity indicators.

1. Key concepts

Improving the tracking of access to electricity must first establish common definitions for data sources, collection methods, and assumptions needed to estimate access rates as well as other related indicators. This chapter outlines the main definitions for classifying access to electricity, connection types and the primary methodologies in use today, as well as their respective advantages and challenges. These common definitions and methodologies will serve as a foundation for discussing best practices and recommendations to track access to electricity with the use of data coming from electricity distributors, hereafter referred to as supply data.

1.1. General definition of access to electricity

A household is considered as having access to electricity if benefitting from an active connection to the grid or to off-grid electricity sources large enough to provide a defined minimum level of energy services. The definition of the minimum level of services is key to identify which type of electricity sources should be included or excluded when counting households with access. This requires periodic revisions on the definition of systems that can provide this minimum level as energy efficiency evolves.

The **access to electricity rate** for a certain country or area, which represents **the share of the population benefitting from access to electricity,** is defined as the total number of people living in households having access to electricity divided by the total population.

Access to electricity rate = $\frac{Population_{with\ access}}{Population_{total}}$

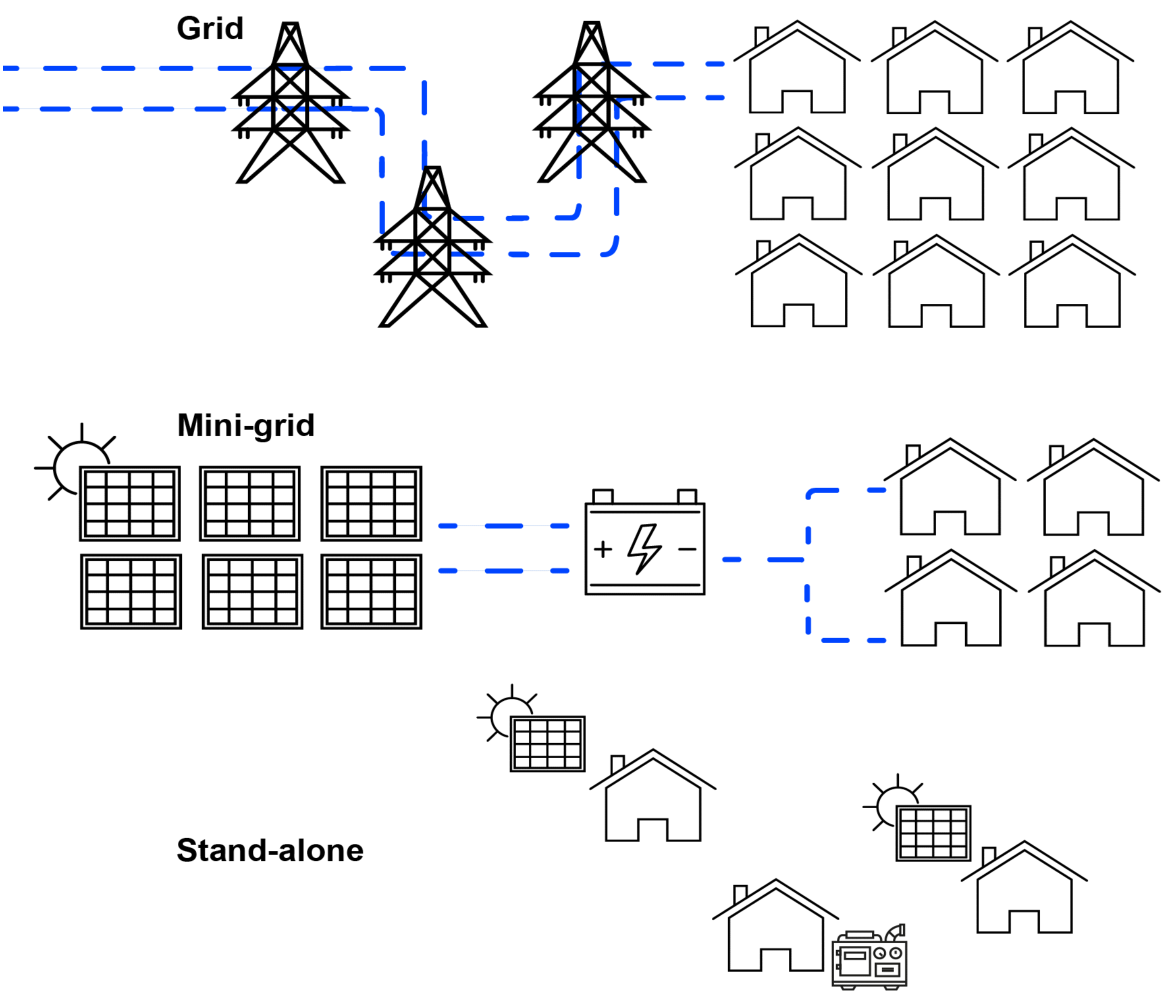
Although this definition does not capture all the nuances that help qualify access (e.g. how electricity is consumed, its contribution to quality of life, reliability, affordability), it represents a reasonable and practical way to track access to electricity through the use of supply data. Focused surveys can be used to track these nuances but existing supply data can also help in estimating them (see Section 3.4, “Other data and indicators”).

The share of the population is the main indicator for tracking access to electricity but the share of households with access is also used. Tracking and reporting using both metrics is welcome, with units clearly indicated and appropriate metadata included. Since access rates need to be defined against a minimum level of access definition, for harmonisation purposes, the IEA recommends using our basic bundle as a threshold for access accounting, but the IEA defines three different levels for access: basic, essential, and extended bundle of energy services (see Section 1.3, "Defining access to electricity"). Countries can report access rates against each of these definitions, or use other threshold definitions, but any reporting should specify the definition used in the metadata reported.

1.2. Defining connection technology types

Households can gain access to electricity through different sources or technologies, including a connection to the main grid, a mini-grid or the installation of a stand-alone system. Mini-grids and stand-alone systems are often referred to as off-grid solutions, however both can be operated in conjunction with the grid as backup or secondary power sources.

Connections by type of technology



Technology connection definitions:

Grid: A network of electrical stations and substations plus transmission and distribution lines that connect electricity generators (centralised and decentralised) to final consumers.

Mini-grids: Small electric grid systems comprised of a generation unit(s) and distribution lines, often not connected to main electricity networks, that link to households and/or other consumers. Mini-grids typically have power capacities ranging from 10 kW to 1 MW and differ depending on the supply technology used, including solar PV and batteries, hybrid solar and diesel, hydro, wind, biogas, among other sources. Mini-grids can be eventually connected to a main grid and in this case the reporting of the connection must be done by the distribution company selling to the final customer, which can be either the main grid utility or the mini-grid operator.

Stand-alone systems (SASs): Electricity generation devices supplying single households or small businesses. They can vary in size from several watts to a few kW and include:

- **Solar-home systems (SHSs):** Small-scale photovoltaic and battery stand-alone systems with capacity equal to or higher than 10 watt peak (Wp). They are used off-grid but also as backup where grid supply is not reliable. The IEA counts solar home systems as access to electricity but it is strongly recommended to track them separately according to a specific size range per the taxonomy proposed in this report, which is based on GOGLA classifications. A growing number of large SHSs, ranging from several hundred Wp to a few kWp, are being sold and often referred to as solar generators.
- **Solar multi-light systems (SMLS):** Solar PV-based lighting systems of 3 Wp to 9 Wp that provide more than one lighting point and phone charging. The IEA excludes these systems from access to electricity accounting, but strongly recommends to track and report them separately.
- **Solar lanterns:** Solar PV-based single-light systems of power below 3 Wp, which provides limited lighting (and sometimes phone charging). Although they represent a strong improvement in life quality, they do not provide a minimum level of energy service that most organisations define as access to electricity. For that reason, the IEA excludes them from access to electricity accounting, but recommends tracking them separately given the high volume of use. Almost two-thirds of off-grid solar systems sold by GOGLA members in 2021 and early 2022 were solar lanterns.
- **Off-grid solar system (OGS):** Includes SHSs, SMLS and solar lanterns.
- **Fossil fuel generators:** Electricity generation devices based on the use of internal combustion engines from gasoline or diesel and a generator.

- **Other types of stand-alone systems:** Although available but less common, other types of SASs include micro-hydro, wind, rechargeable batteries and hybrid systems.

1.3. Defining access to electricity

Organisations such as the IEA and the World Bank, as well as some individual countries, have adopted different classifications for defining access to electricity, with the benchmarks used for minimum levels of service, in particular, diverging. As a result, global access statistics may differ due to incompatible definitions.

The IEA defines access to electricity as a household receiving enough electricity to power at least a minimum level of services capable of growing over time. The IEA minimum level of services is defined as the “basic bundle”, which includes more than one light point providing task lighting, phone charging and a radio, or broadly equivalent to a range of around 50-75 kWh per household per year, depending on efficiency levels. Households connected to national grids and mini-grids are considered having access since these systems can satisfy this criterion.

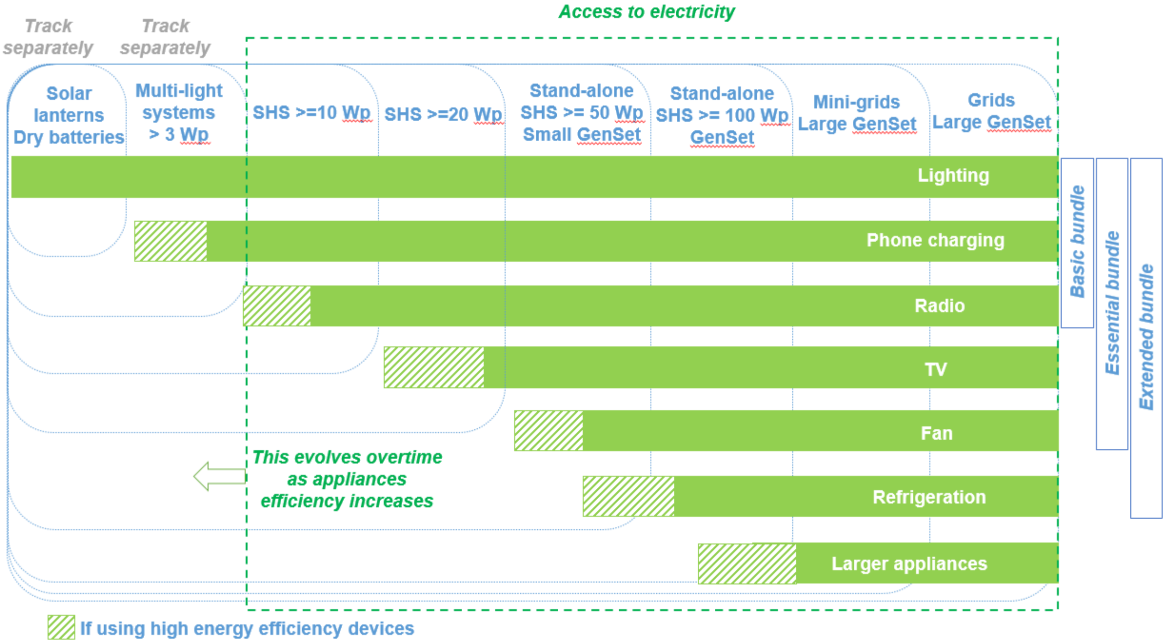
In the case of OGS, the IEA considers solar home systems (SHS) of 10 Wp and above as providing access to electricity. Over time, the increasing efficiency of electrical appliances and devices will bring down the minimum size of the systems able to provide the basic bundle. As an example, before the use of efficient LED lighting, the basic bundle would have required higher capacities. Although SMLS and solar lanterns are excluded from access tracking, the IEA strongly recommends to keep track of their deployment as they represent an important first step towards gaining access.

Many governments include small off-grid solar systems and some SMLS in their electrification programmes since these both can help households make the first step on the energy ladder and represent the initial stage in plans for a gradually growing energy service.[1]

The IEA’s “essential bundle” and “extended bundle” also allow for the need to expand electrification plans in the medium term. The essential bundle includes four light bulbs for four hours per day, a fan for three hours per day, and a television for two hours per day, which equates to roughly 500 kWh per household per year. This can be provided by a SHS of 50 Wp and above or a small fossil generator. The extended bundle implies a refrigerator, four hours for lighting, four hours for TV, and six hours for a fan, which equates to roughly 1250 kWh per household per year, and can be supplied by a fossil generator or a large SHS of more than a 100 Wp.

[1] This includes the Democratic Republic of Congo, Ethiopia, Kenya, Madagascar, Malawi, Mozambique, Nigeria, Rwanda and Uganda.

IEA access to electricity by technology

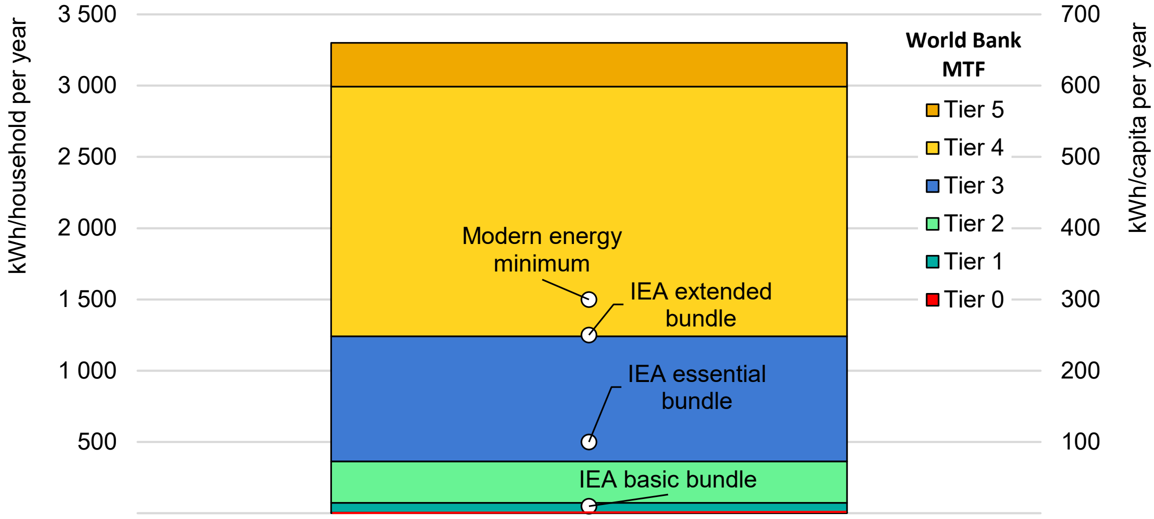


Notes: The figure shows the potential of powering the energy services which evolves with the efficiency of appliances. GenSet refers to fossil fuel generators. Larger appliances include cooling, cooking, cleaning and other electric devices.
Source: This figure is adapted from a version originally featured in the IEA's Energy Access Outlook 2017.

The World Bank is charged with the tracking of the proportion of the population with access to electricity under UN Sustainable Development Goals (SDG7.1.1) and it does this through the use household surveys and censuses. The World Bank considers electricity access down to small solar multi-light systems of 3 Wp, but also proposes a scaled approach to measure energy access known as the "Multi-tier Framework" (MTF) developed by its Energy Sector Management Assistance Program (ESMAP). The MTF establishes six tiers of energy access characterised by a combination of the amount of energy consumption, reliability and affordability. To collect the data under the MTF ad hoc surveys need to be run in a country, and from 2015 to 2019 there have been 17 MTF surveys run in countries across Latin America, Africa, and Asia. The IEA's "basic bundle" is equivalent to MTF Tier 1 while the "essential bundle" to Tier 3 (see figure below).

Several organisations are advocating for greater ambition in access provision to support a minimum level of energy consumption reflective of average global needs for energy worldwide. For example, the Energy for Growth Hub proposes a new Modern Energy Minimum (MEM) to support quality of life and economic growth in all countries. The minimum is 1 000 kWh per person per year, of which 300 kWh is for residential use. Although this level represents a good long-term target, the proposed MEM does not mirror the reality of the large population gaining access for the first time who will need to start at much lower levels, which is more accurately reflected in the IEA basic bundle or even below. The values in the basic bundle might seem to start relatively low but the IEA access vision implies households will climb the energy ladder with the support of national policies and further electrification planning. Indeed, the MEM is very similar to the IEA "extended bundle" of 250 kWh per person per year.

Comparing minimum levels of energy services defined by different organisations



Notes: The tiers represented in the figure refer to the World Bank's Multi-Tier Framework definitions. The IEA minimum level of energy services, the basic bundle, falls into Tier 1. The Modern energy minimum corresponds to the residential share of the minimum proposed by the Energy for Growth Hub. For converting from household to per capita consumption the regional average household size of five people is used.

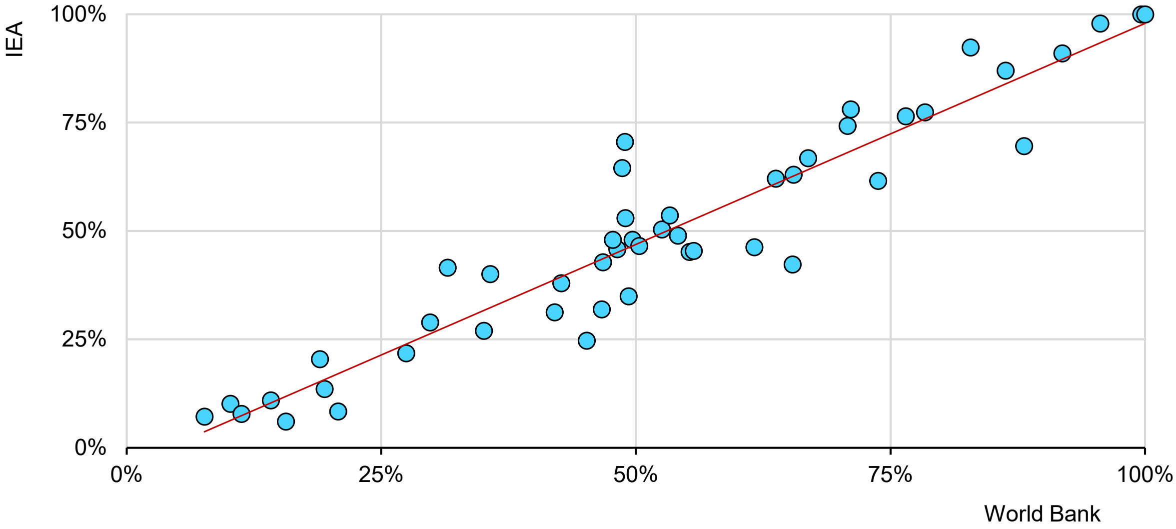
The IEA maintains a slightly higher threshold for the minimum level of service in its access definition than the World Bank, but, similar to the multi-tier framework, sets different thresholds to reflect the varying contexts. This is underpinned by the objective that access planning must aim at supporting rising levels of household energy use in the future, while balancing the consumers' ability to pay.

Harmonising access to electricity definitions used by organisations in the field and by countries is an important but elusive task, raising a challenge for standardising how countries report their data. Some countries, such as Zambia or Rwanda, have chosen to define electricity access at a specific MTF tier (Tier 2 and Tier 1, respectively). The use of metadata to clarify definitions in administrative data is a critical component of enabling transparency in order to improve comparability of access indicators (see Section 3.3, "Coherence, comparability, and transparency of released data: Metadata").

1.4. Supply-side vs. demand-side data sources

Electricity access data can be collected either by using supply-side or demand-side sources. Supply-side data is collected from the companies providing electricity services to end users such as utilities, distribution system operators (DSOs), mini-grid operators, SHS distributors, among others. Demand-side data, on the other hand, relies on direct reporting of access status by consumers via surveys or censuses. The IEA access to electricity database is mainly based on supply sources, however the resulting rates are in most cases comparable with World Bank data, which are mainly from demand-side sources (see figure below).

IEA vs. World Bank 2021 national access rates in selected African countries

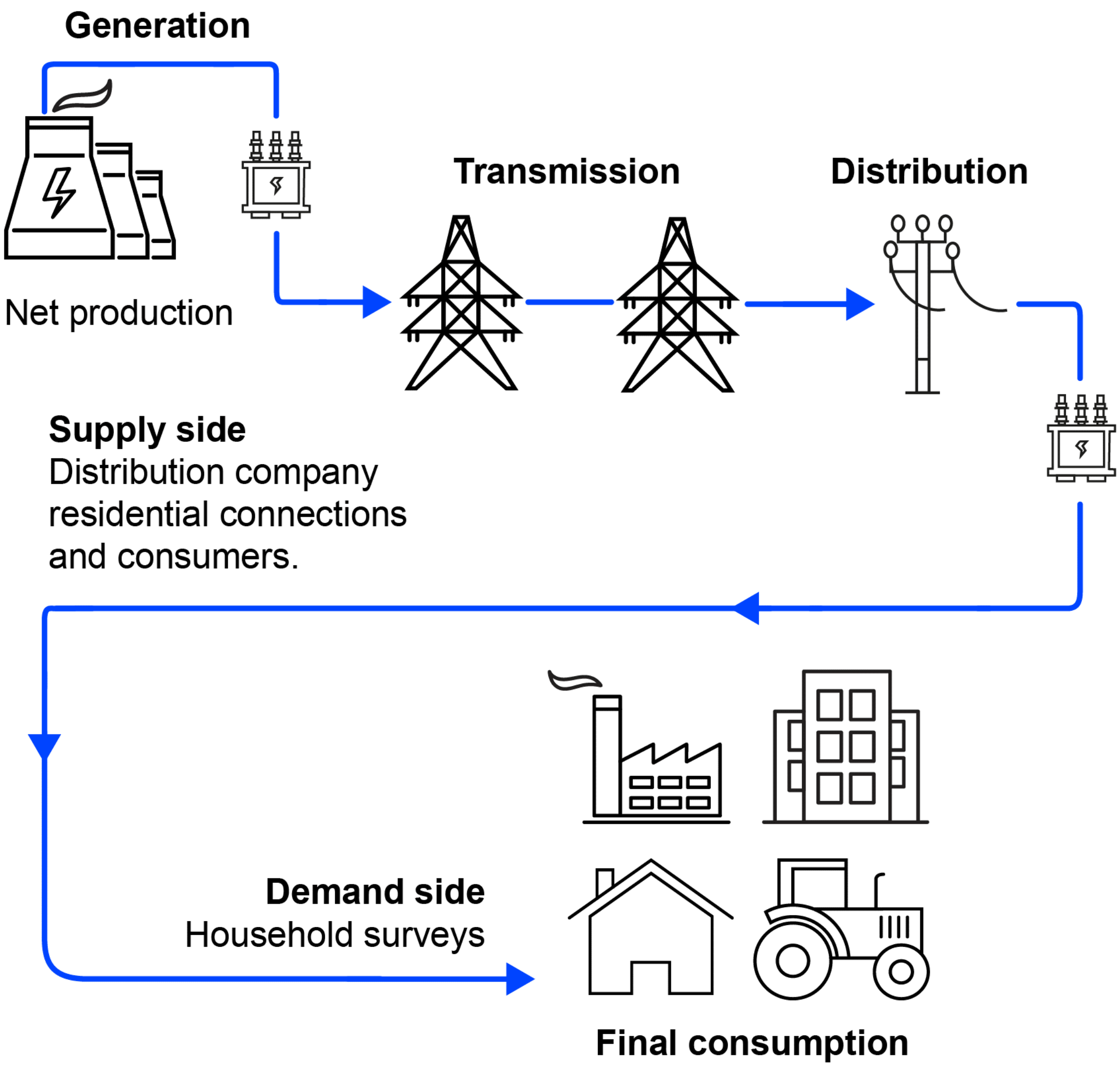


However, demand data is largely generated by household surveys or censuses, which are typically preformed every five to ten years. By contrast, using supply data to track access to electricity data is more readily available and timely, advantages that are key to tracking the most recent trends at a low cost and provide much-needed real-time feedback for policy makers in their planning strategies and for industry players weighing investment decisions. However, the increasing uptake of stand-alone technologies for access to electricity, such as the use of solar home systems, would lead to an increased disparity between supply-side and demand-side data if countries do not start collecting and processing supply data for stand-alone systems.

Both supply and demand side data have their strengths and weaknesses but they are also complementary and if used together provide many synergies for a sound data strategy.

Supply-side data typically comes from national utilities, licensed embedded concessionaires such as private electricity distributors, mini-grid operators, and retailers of stand-alone electricity-generating equipment like SHS distributors. Secondary sources include data aggregators and/or reporting agencies, such as ministries, regulatory and electrification agencies, industry associations (AMDA and GOGLA), programme donors and international organisations. Supply-side data is mostly based on the number of serviced customers (e.g. the number of connections) or sales data (e.g. the number of SHSs sold). Supply-side data is relatively cheap to gather and can be updated frequently. Smart meter deployment will make this data even easier to gather. However, supply-side data cannot directly account for informal connections and multiple systems in a single household such as solar home systems as a backup to a grid connection.

Supply-side and demand-side data sources – example for grid connections



IEA. CC BY 4.0.

The supply data IEA compiles to measure access to electricity comes mostly from national administrations such as ministries of energy and utilities, energy regulators and rural electrification agencies, among other official data reporting organisations that most often rely on supply data. Administrative entities using supply data for calculating access to electricity rates need to collect the number of residential customers of grid and mini-grid companies, the number of stand-alone systems sold or distributed, and the household sizes. Household sizes are used as a proxy to estimate how many people are benefitting from a connection and the IEA recommends using location-specific (e.g. jurisdiction or rural/urban) household sizes as the difference might be important. The supply-side data used by the IEA is collected on an annual basis, which provides the most up-to-date comprehensive electrification rate information currently available.

History of the IEA's work on electricity access rates

The IEA was the first international organisation to track and produce a global database with information on access to electricity, starting in the early 2000s. The IEA has continued to update these figures annually, providing timely estimates for global progress toward universal electrification. Our latest data by country can be found in our online database, along with other key metrics related to the UN's SDG7, which calls for access to affordable, reliable, sustainable and modern energy for all by 2030. The IEA is one of the official co-custodians for SDG7, and contributes annually to official reporting on global progress toward the targets.

The IEA also produces projections for electricity access *for different scenarios* in its annual *World Energy Outlook*, assessing expected progress under current policies and trends and under official national targets as well as proposing a least-cost, feasible pathway to universal access by 2030.

Demand-side data typically relies on household surveys or census results and the main data providers are national or central statistics offices. Surveys represent a fundamental tool for data collection even though they are less timely and more expensive than supply data. Surveys gather information directly from the source and answer the key question of who is or who is not getting access, avoiding double counting issues. However, because of its delayed results and low frequency, surveys cannot gauge how different policies and investments are helping progress toward universal access. Demand-side data often rely on year-on-year estimates based on previous trends and can miss sudden changes of directions.

Beyond nationally developed and implemented surveys, specific surveys covering access to electricity include Demographic and Health Surveys (DHS), the Living Standards Measurement Study (LSMS), the Multiple Indicator Cluster Surveys (MICS), and the World Health Survey Plus (WHS+). Surveys can capture households using stand-alone systems, or if they are connected informally to the grid through a collective metering point. Larger ad hoc surveys can gather information about the level of energy use, end uses, reliability and affordability.

However, surveying also carries challenges in designing a representative sample, especially of remote areas, and can introduce subjectivity in responses. For example, asking “Does your household have access to electricity” may be answered yes even if a household just has a solar lantern, which is not considered access in most international access frameworks. In some cases, subjectivity in survey responses can be efficiently reduced by training interviewers and adopting clear definitions and questions but this requires more resources.

The World Bank access tracking relies mostly on demand-side data from relatively expensive surveys. For example, its MTF survey requires around 3 500 household interviews equally divided between rural and urban regions. The estimated cost is approximately USD 300 000 for the data collection of one MTF survey, excluding data processing and analytical activities related to the reporting. ESMAP and others have developed an abbreviated alternative to the MTF survey that can be deployed in conjunction with existing household surveys, such as ones on the labour force. ESMAP has also conducted a study on the possible application of satellite and other geospatial data to complement surveys and to address the lack of subnational information.

Interoperability between supply-side and demand-side data

Combining supply-side and demand-side access to electricity data can be useful to help calibrate industry data on stand-alone and mini-grids and help estimate the number of informal grid connections, as well as provide other critical qualitative information. This would require sharing and exchange of data as well as coordination and co-operation among all supply and demand sources.

In addition, access to electricity indicators produced from supply data could be checked, completed, and improved thanks to information gathered through household surveys and vice versa. This would harness the synergies of the two approaches and aid in identifying and closing any major data and information gaps that exist.

2. Overview of electricity access tracking

This chapter gives a broad outline of the overall process of producing access statistics, including an overview of how to implement a new supply-side data collection process, from administrative steps to stakeholder engagement.

2.1. Data work flow overview

Tracking access to electricity involves collecting, validating, processing, and disseminating data (see figure below). This should be coordinated by a single entity, referred throughout this Guidebook as the **national entity tracking access to electricity, or "the Entity".** An overview of each step is given below.

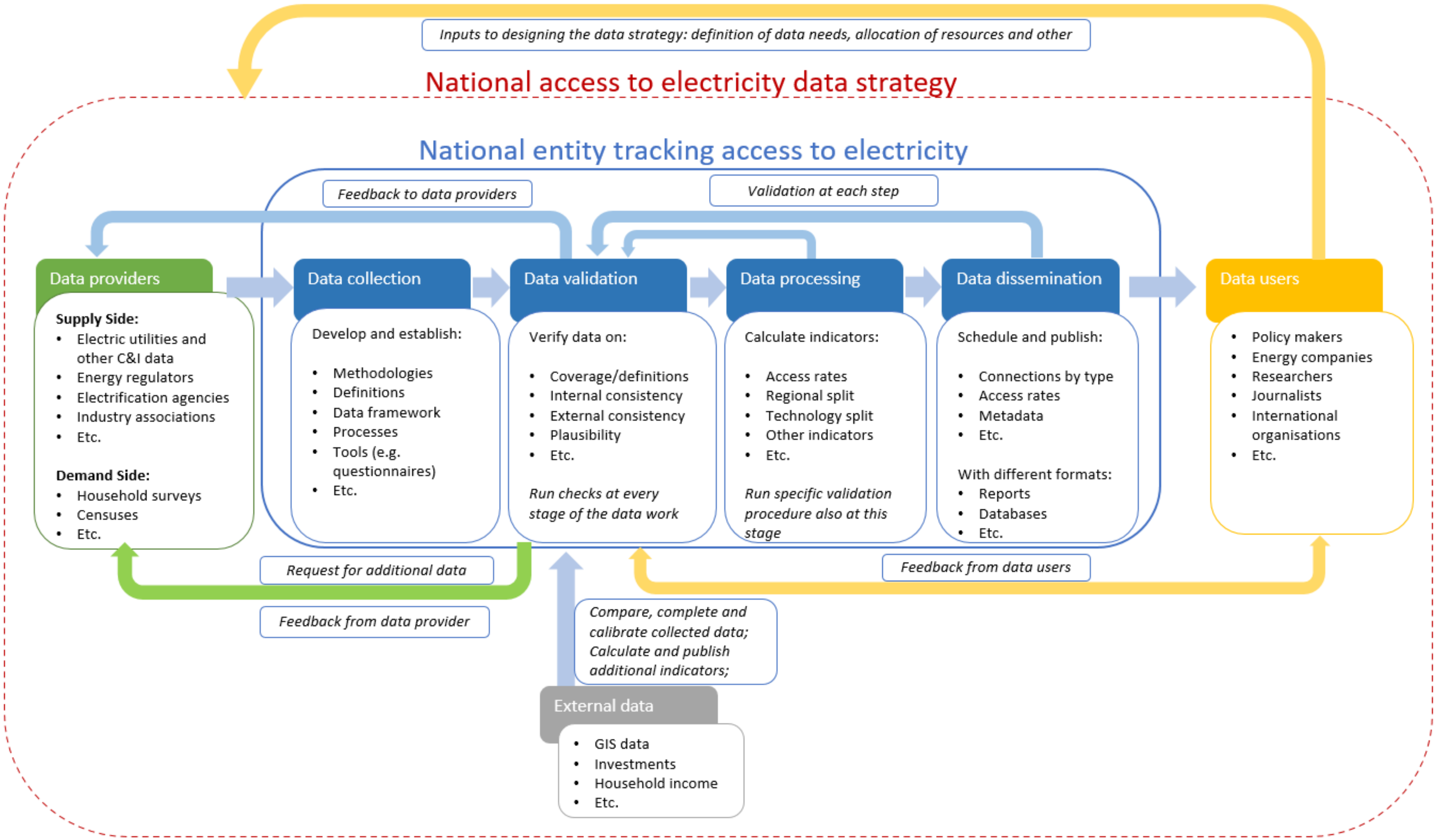
Data collection: The Entity first works with data providers on the collection and submission process for the agreed data. Data providers include utilities, energy distributors and national entities such as electrification agencies or energy regulators. The Entity should work with providers and data users to agree on what data is required, which may already be available, and then agree on a standardised format and timing for data providers to submit this data.

Data validation: Data should be checked to ensure its quality and accuracy. As problems are identified, the Entity should interface with data providers to address any issues or inconsistency. For example, a sudden variation in connection numbers between two periods is a red flag and should be investigated. This might necessitate a second data submission. The Entity should keep track of identified data issues and record the corrections to ensure internal transparency. Data should also be checked against related data from national sources or other comparable agencies to ensure internal coherence.

Data processing: Raw data is used to calculate the final indicators, such as access to electricity rates. At this step, it is recommended to use automatised calculations. Data validation also occurs in this step, with the processed data checked for accuracy and consistency. If underlying problems are uncovered, statisticians need to go back to the original provider to confirm the data.

Data dissemination: After final data and indicators are completed they are published in different forms such as databases and reports, and need to be easily accessible and understandable, especially to key data users. Data users may identify errors or have additional questions, which provide important feedback for statisticians and help with data validation and long-term planning.

Access to electricity tracking workflow



Note: The workflow described above is for reference only.

2.2. Implementing an access data strategy

Launching and sustainably establishing an energy information system requires more work and effort than the actual running of the process to produce the statistics. Therefore, this section gives high-level guidance for those implementing a new or willing to improve an existing access to electricity information system, although the exact steps vary across different contexts. In the initial set-up, the IEA recommends an annual collection of access data, however, with the use of supply-data, this can be done more frequently if countries want to produce more timely estimates on progress from ongoing access provision campaigns, as is the case in Rwanda.

- **Identify the key data providers and users and establish an access to electricity data working group.** These can be private and public entities at the national, sub-national and international level, and could include distribution companies, regulators, electrification agencies, national statistics offices (NSOs), energy ministries and financial or technology programme donors, among others. The working group should hold a kick-off meeting to align on what data are available and what data are needed to support policies and programmes for electricity access. This group should be consulted throughout the data strategy development and meet periodically to provide guidance and can build on existing energy data working groups. Tracking access to electricity should be embedded in the main energy data working group's strategic mandate, where possible.
- **Develop solid legal foundations.** A clear legal framework is needed to obligate data providers to supply the needed data on time, in the correct format and at the level of detail needed for the statistics. Legal frameworks can charge the Entity with the responsibility to ensure confidentiality of specific datasets. It is recommended that the Entity in charge is provided with a legal mandate to enforce data collection from data suppliers and to maintain a national Energy Information System (EIS).
- **Define a clear plan for the data work**. All stakeholders involved should agree on a realistic timeline and scope for the data work. Data collection tools such as questionnaires or online tools and their structure should be discussed with data providers to align, where possible, the reporting format to how those entities collect and store their data. Documenting procedures to perform this task at each phase of the data production process and keeping track of specific actions and communications with other stakeholders is key to ensure a sustainable and more efficient execution of the process.
- **Allocate financial, human and IT resources.** Long-term financing for tracking access to electricity needs to be secured in advance to ensure the sustainability of the data work. Raising the importance of data in any political and investment decision at higher levels is fundamental to ensure allocation

of funds. Part of these resources should be allocated to training the needed staff. Recording trainings can help lighten this load for new staff in the future. Allocations also need to be made to set up the appropriate IT systems, including computers, software, internet connections, and servers to store the data. Having one person on staff who works on maintaining and upgrading these data systems (not only for access data but all data), is essential.

- **Provide continuous training to data providers, users and statisticians.** Make annual training sessions available to all data providers, which covers changes in data reporting, uses and verification. Providing training to the team in charge of tracking helps them keep pace with a changing sector. For example, tracking off-grid connections is a new concern and staff need to be trained on new methodologies and definitions. IEA delivers many training workshops, which can be made available to countries.
- **Make improving data coverage and quality a priority.** Allocate resources to make improvements, such as extending the coverage to new indicators, revising historical data in light of corrections, and integrating new data sources. The top new priorities should be agreed in consultation with the access to electricity tracking working group as well as other national entities and energy sector stakeholders.

3. Walking the process: from data collection to dissemination

This section provides detailed step-by-step procedures to develop and implement a comprehensive and granular access to electricity data system. It also highlights the most common issues faced when tabulating this data and recommends best practices to address concerns, including the suggested notes to include in metadata when reporting final data. Finally, it highlights opportunities to collect and collate additional data that can help with access planning.

3.1. Data collection

At the heart of electricity access statistics is counting the number of households that have a connection able to provide a defined minimum level of energy services. It is important to keep track of the number of connections disaggregated by type, such as grid, mini-grid and stand-alone systems, and by geographic location, including by jurisdiction or by rural and urban areas.

Below are guidelines for counting connections by different types. We also detail which supply data can be used and what are the most common sources. It is recommended to collect data for the most recent year (or period) and revisions in counting connections of past reporting.

Counting grid connections

The number of grid-connected households can be derived from electric utility or distribution company data. However, the granularity of data, such as a breakdown of connections by residential and small commercial customers, may not be always available. For example, some utilities only track the voltage or power at which customers are connected to. In this case, low-voltage connections could represent a good proxy for total households connected to the grid.

The data collection process should ideally rely on a standardised questionnaire that would go to all relevant distribution or retail grid operators.

Guidance on data for grid connections

Topic	Notes
Main data sources	• Electric utilities and/or distribution companies (DisCos) • Electricity sector regulators
Counting preferences	The sum of registered residential connections, including metered and unmetered customers, that have accounts (i.e. fixed-price unmetered connections).
Types of connections to count	In order of preference: • Connections/customers that are classified either as domestic, residential or household. • Low-voltage (LV)/low power connections/customers, which can include small and medium-sized enterprises (SMEs), government offices, educational and health infrastructures, but only if customer types are not available. • All customers if types and voltage levels are not provided. This might require the use of additional data and assumptions to model the estimated number of households connected. Alternative estimation methods when no customer data is available: • Projected number of connections based on length of medium-voltage (MV) line infrastructure. • Estimated number of connections based on feeder load.
Time frame	Connections operational on 31 December of reporting year or relevant to any other period of reporting (e.g. semi-annual). In addition, revisions to historical years must be collected.

Challenges and best practices

Informal connections: It is possible that some households are illegally connected to the grid. The default approach is to exclude these connections from access counting. Making electricity access more affordable for households by reducing fees, offering discounts for low-income customers and providing other financial incentives would reduce the number of informal or illegal connections. (See box below, “Managing common connection counting challenges”).
Sub-metering: In some countries multiple customers can be served by a single official connection, which can distort the collection of access data. Regulators should legally require providers of sub-metering services to report the number of sub-metered.
Backup supply: It is important not to count more than one connection for the same household for the purpose of tracking access. In these cases, it is recommended to count only the grid connection and exclude backup supply. Backup systems should be collected separately, which can be done by requiring companies selling stand-alone systems to collect data from their customers and report it to the Entity.

Going further

Quantity of consumption/level of energy service: Different tariffs for varying levels of household demand are often applied. Distribution companies could report the number of residential connections by tariff tier as well as the average electricity consumed by customers in each tier group. This would give better insight into how many households are living at different levels of energy services as well as help assess tariff design to ensure social or low-income rates are set at levels that balance affordability for the consumer and healthy utility operating margins.
Geolocation: Connections disaggregated by jurisdictions and by rural or urban areas can drastically improve the use of access data for electrification planning and policy-making. The level of geographical disaggregation and the definition by location should be agreed to among the users and the owners of the data such as policy makers, energy planners, access tracking entities and distribution companies. Importantly, many utilities already know the location of customers and with some additional granularity in the data collection and management process they could provide important geographical information (e.g. residential connections by village).
Quality of supply: From typical indicators such as the System Average Interruption Duration Index (SAIDI) or the System Average Interruption Frequency Index (SAIFI) to estimating how many residential connections are impacted by power cuts of specific durations and frequency can improve the understanding of the quality of electricity supply (see Appendix, “Other data and indicators”). Utilities and distribution companies can collect and use data on the number and duration of power cuts per node/distribution station combined with the number of customers served by the same nodes to also come up with geolocated supply quality information.
Affordability: The administrative Entities in charge of tracking access to electricity can collect and disseminate data on electricity tariffs and connection charges/fees. This can also be done simply by providing a link to the utilities or regulators official tariff reporting system. Utilities/distribution companies often apply kWh tariffs that increase with the electricity consumed and/or the power contracted with the aim of improving affordability for lower consumption households. Affordability is one of the main barriers for closing the access gap and tracking it is fundamental.

Counting mini-grid connections

The Entities responsible for tracking access to electricity are encouraged to work with mini-grid operators or other data providers to collect the number of households connected to them. Similar to grid connections, the number of customers by type (e.g. residential, commercial, and industrial) may not be available but data on customer consumption levels or capacity are often gathered and can be helpful for improving the tracking of the quantity of supply.

The IEA recommends periodically tracking data from all operational mini-grids, gathering information on location, type (e.g. solar/diesel hybrid or solar PV with battery storage), its nominal capacity (in kW), and, most importantly for tracking access, the number of customers served. Counting connections from mini-grids installed in the past without validating their continued operation can lead to accounting mistakes. It has been reported that a significant number of mini-grids in the African continent are no longer operational after a rather short period because the cost of maintaining equipment and regular maintenance was not accounted for when the systems were installed, resulting in their shutdown.

Mini-grid owners or operators should be legally required to report/submit data to the national Entity through a legally binding formal reporting system. The administrative entity tracking access can work with regulators to require mini-grid operators to periodically provide operational data in order to be licensed.

Guidance on data for mini-grid connections

Topic	Notes
Main data sources	• Competent licensing authority for mini-grids (e.g. energy regulator). If regulations for licensing mini-grids are missing, governments must introduce them. • Mini-grid operators. • National mini-grid programmes. • Technical and financial partners (programme donors). • Mini-grid industry associations such as the Africa Minigrids Developers Association (AMDA). • Technology and data platform providers, whether smart meter or software companies. Administrators may triangulate for consistency but beware of double counting the same connections. Keeping a registry of operational mini-grids would help avoid double counting. Comparing mini-grids' connection data with latest surveys could also help in identifying double counting and other issues to refine the estimates.
Counting preferences	The sum of registered residential connections, including metered and unmetered customers, that have accounts (i.e. fixed-price unmetered connections).

Topic	Notes
Types of connections	In order of preference: • Connections can be classified as domestic, residential or household and, where possible, by average annual consumption. • Total customers if user types and voltage levels are not provided. Alternative estimation methods when customer data is not available: • Estimation based on the number of households or population of specific areas (e.g. village). This would require an estimate of the share of the households connected to the mini-grid to avoid inflating the number of people benefitting from access. • Modelled number of connections based on the power and estimated capacity factors of the mini-grid.
Time frame	Connections operational on 31 December of reporting year or relevant to any other period of reporting (e.g. semi-annual). In addition, revisions of historical years must be collected.

Challenges and best practices

Double counting connections: Sometimes mini-grids are installed in areas served by the main national grid either because the grid arrived at a later stage or as a support to a weak grid supply. In these areas some households might be connected to both the main grid and a mini-grid. For counting purposes, it is important to avoid reporting these households twice. Mini-grid operators are often aware if their customers are already connected to the grid and this information could be directly collected. Otherwise, a discount factor can be applied to mini-grid connections operating in villages or areas served by the grid. Such a discount factor must come from additional studies and research or ad hoc surveying. For mini-grids connected to the main grid, the reporting is done by the company officially distributing electricity (i.e. billing) to the final consumer. This would determine the classification as either a grid or a mini-grid connection.

Sub-metering and informal/illegal connections can also be a challenge for mini-grids and similar approaches as illustrated above for grid connections can be used (See box below, “Managing common connection counting challenges”).

Going further

Geolocation: in the case of mini-grids, location-based tracking can be relatively easy since the absence of transmission lines implies that connections happen close to the mini-grid. Directly tracking all operative mini-grids including their location is the preferred option.

Quality of supply: Properly sized mini-grids provide relatively good electricity services – and can have uptime periods ranging from 90% and 97%. However, it remains fundamental to track the quality of electricity provided by mini-grids in terms of service uptime.

Affordability: Similar to grid connections, electricity tariffs and connection charges/fees applied by mini-grid operators can be collected. Together with consumption levels and household income they are key for understanding energy affordability.

Quantity of consumption/level of energy service: Mini-grid operators and other data providers often track the consumption of each customer through meters and smart meters. In many other cases, they are aware of the total electricity distributed and the number of customers, enabling them to estimate at least the average consumption per household served. Implementing new procedures to capture this information would be valuable.

Counting stand-alone system connections

Stand-alone systems include a number of different products, but most commonly are fossil fuel generators and solar home systems. As these are most often sold, or distributed to consumers, capturing these connections relies on counting systems delivered by the distributors. This presents a number of challenges in counting SASs, which include collecting data from a large number of decentralised private actors, avoiding double counting of systems used for backup power or replacing an existing unit, the second-hand resale of these units, and discounting units that reached end of life or that do not meet the minimum definition of access.

To cope with these challenges, the IEA recommends that the Entity create a licensing framework for companies to operate in the country and access incentives and count only systems from licensed distributors. Within the licensing framework, there should be reporting requirements that ideally would include information on the size of units distributed, the nature of the installation (i.e. backup power versus primary connection), and location. Industry associations can also play an important role in tracking sales in the country.

Even with strong reporting, the IEA recommends using discount rates to account for natural turnover of these systems, which should be calibrated based on regional market characteristics.

Guidance on data for stand-alone systems

Topic	Notes
Data sources	• Competent licensing authority for stand-alone systems (e.g. energy regulators). If regulations for licensing are missing, governments must introduce them. • SAS companies as distributors/installers (preferably licensed). • Industry associations. • Off-grid electrification programmes. • Technology and data platform providers, whether PayGo software solutions or industry associations. • Custom databases (import of PV products). However, these should be used only for reference and double checking as it can lead to significant counting errors. May triangulate for consistency. If collecting data from both electrification programmes and companies, only sales taking place outside of the programmes needs to be collected from the latter.
Counting preferences	The sum of households equipped with at least one operative stand-alone system of enough size to provide the minimum level of energy service. For solar home systems acquired in the free market, supply data are available as sales or subscriptions (e.g. PayGo). For SHSs under PayGo contracts, companies must be required to report active systems and, when possible, for both free market and programme frameworks.
Type of connections	At least one stand-alone system of capacity equal to or higher than a defined minimum.
Time frame	Systems still in use by 31 December of the reporting year. This includes collecting sales data back to the assumed lifetime of systems or clearly noting in metadata otherwise.

Challenges and best practices

Which products to include: Not all stand-alone systems can provide the minimum energy services defining access to electricity. In the case of off-grid solar systems, the IEA recommends to include only for SHS (>= 10 Wp) in access counting and to track separately different size ranges. OGS systems below 10 Wp (solar multi-light systems and lanterns) need to be tracked but independently from access.[2]
Even small fossil fuel generators can provide the level of energy services that meet IEA access definitions. They can be accounted for with household surveys. Most often they are used as backup solutions where the grid is not reliable.

Counting connections from SHSs under PayGo contracts: The preference is to require companies to report the number of active customers. Although, this exercise can be facilitated under electrification programmes by including in tenders the requirement of periodical reporting of active customers, it can be more challenging for free market contracts.
The IEA recommends creating licensing frameworks entitling companies to operate in the country and access incentives provided certain conditions are respected including the reporting of active customers. However, in the short term, PayGo systems might need to be tracked through new sales/contracts.

Counting connections from SHS direct sales data:[3] For systems **distributed under electrification programmes**, the Entity in charge of the programmes must require a periodic verification of systems still in operation and preferably include the maintenance and replacement of systems to ensure continuity of access. When this is met, tracking access to electricity is implied in the process. Otherwise, tracking a system distributed through programmes must follow steps required for direct sales.
To convert direct sales data to access, the number of operating systems needs to be estimated. To do so, historical sales data and assumptions/information on the lifetime of systems are required.

Discounting for end-of-life and breakage: Some stand-alone systems might breakdown or become unused before the end of their lifetime. The number of operating systems is then estimated as the moving cumulative sum of systems sold within the window of time equal to the assumed lifetime. For example, for systems with an expected lifetime of five years, to estimate the number of operative systems at the end of 2022, the sum of sales from 2018 to 2022 is used, and assumes systems distributed in 2017 are no longer operative. In the case of SHSs, a lifetime of five years might be appropriate, while smaller OGS systems are often no longer operative after around three years.

Discounting for repeat sales: Some households might use more than one stand-alone system at the same time. This discount factor might also need to be applied to free market PayGo contracts from companies reporting the number of customers if they cannot provide information of repeat sales or backup use.

Discounting for multiple connections or backup use: To ensure that SASs are not used as grid backup that can lead to double counting, a discount factor must be applied. As backup use of SAS can vary dramatically country by country, national specific discount factor must be applied. The IEA recommends counting access of these households as coming from the grid.
In the case of SHSs, GOGLA maintains up-to-date methodologies and default values for the above discount factors.

Covering the full market of unregulated sales: SAS companies operating without licensing are hard to track, and should be omitted as the default, especially as many of these systems may not meet needed specs to reach minimum access definition thresholds. The IEA strongly recommends to put licensing frameworks in place to ensure the safety of users and to make sure low-quality products do not undermine the local off-grid industry. Countries may choose to estimate the unregulated sales of off-grid systems, but should report these estimates separately and note in the metadata what methodology has been applied.

[2] GOGLA offer's guidance for dealing with the dimensions unique to OGS systems in its "Standardised Impact Metrics for the Off-Grid Solar Energy Sector".

[3] This includes systems distributed under electrification programmes and under PayGo contracts in the absence of the preferred tracking methods defined above.

Going further

Geolocation: When stand-alone systems are part of an electrification programme, households benefiting from them are known and their location must be tracked. Similarly, in the case of business models like the PayGo, companies can track the location of the system. For these cases, gathering geographical data to track where people have access through SASs can be relatively easy provided a data collection process is in place. However, the task is more complicated tracking free market sales because often the reseller is not aware where the purchaser is located – companies under licensing must be asked to gather information about their customers, including where they will install the system.

Quality of supply and level of energy service: Stand-alone systems' quality of supply and the level of energy service they can provide is directly related to their power and energy storage capacity. For example, a solar home system with a larger PV module and battery provides a higher service. Therefore, the IEA strongly recommend tracking the number of SASs by technology and power rating or size ranges. Due to the variety of SASs, it is recommended to identify which ranges can deliver similar levels of energy services and track accordingly. For example, in the case of tracking OGS systems, GOGLA defines and tracks different ranges that can be used as benchmarks for solar home systems.
Corresponding access tracking levels:

- SHS between 10 Wp and 19 Wp (Tier 1) = IEA basic bundle
- SHS between 20 Wp and 49 Wp (~ Tier 2) = between IEA basic and essential bundle
- SHS between 50 Wp and 99 Wp (Tier 2) = IEA essential bundle
- SHS above or equal to 100 Wp (Tier 2+) = IEA extended bundle

Tracked separately from access:

- Solar lanterns below 3 Wp (Tier 0)
- Multi-light systems between 3 Wp and 9 Wp (Tier 0 – partial Tier 1)

Affordability: In the case of stand-alone systems, the cost of the system, and an estimation of maintenance and operation costs (or for PayGo contracts, the deposit and periodic instalments) can be tracked through procurement documents of the electrification programmes and direct communication with retailers and industry associations. GOGLA periodically tracks the cost and the affordability of OGS systems.

Managing common connection counting challenges

When counting connections from supply data sources there are some common issues that could lead to either overstating or underestimating access. The Entities in charge should identify the issues in their countries that must be addressed as a priority, making clear notes in their reporting of how and when the problems were resolved. Below, we offer options to tackle the most prevalent challenges.

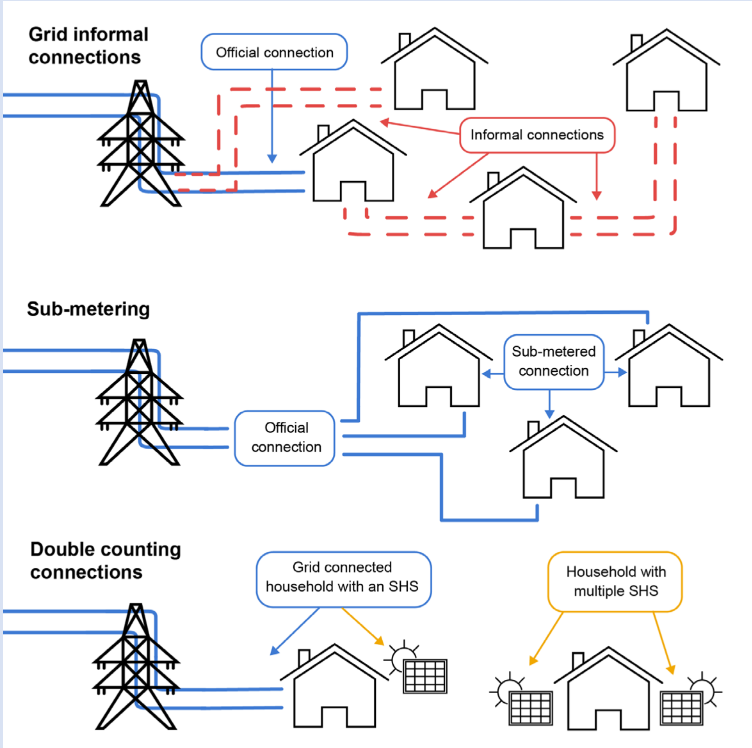
Informally or illegally connected households: Some dwellings may connect to existing distribution networks using their own wiring, splicing into the existing grid. The recommended default approach is to exclude informal/illegal connections from statistics as it is hard to estimate exactly how many dwellings may be connected in such a way. However, survey data can be used alongside non-technical loss data from utilities to estimate the magnitude of the problem. For example, Ethiopia identified through surveys that less than half of the households benefiting from grid electricity have a formal/legal connection.

Multiple customers/households per connection point: Multiple dwellings can be legally connected through a single official connection and pay directly to a landlord or a small private company. This is often referred to as sub-metering and can include different payment schemes (e.g. metered or fixed fee). For the main utility this is traditionally captured as one single connection, underestimating access rates. It is recommended that the Entity tracking access works with the country's regulators to legally require providers of sub-metering services to report to the utility the number of "downhill" connections, which is the number of sub-metered connections minus the primary grid meter. Results from household surveys are also helpful to track sub-metering access by providing the total number of households connected to the grid that can be used as correction/discount factors for future accounting.

Backup generation: Where grid electricity supply is not reliable, many dwellings may purchase a diesel generator or a solar-battery system for backup. Counting these devices leads to overestimating access. Information on the number of households using backup systems is extremely difficult to assess, outside of survey data. Best practice is to require distributors of stand-alone systems to collect data during installation if the system installed is the dwelling's primary or backup energy source. Applying default discount factors is not recommended since the prevalence of backup generation can vary greatly across geographies. For example, the share of households using OGS systems as grid backup ranges from between 3% in Rwanda to 70% in Nigeria. Representative surveys such as the MTF can help estimate a discount factor specific to a country or area.

Stand-alone system stacking. Some households may use multiple stand-alone systems at the same time leading to overestimations. For example, this is when a household owns two solar home systems. As with backup systems, it is recommended to require SAS companies to collect the information about the specific use of the systems from their customers. However, this issue is typically less pronounced than tracking the use of backup systems.

Connection counting challenges

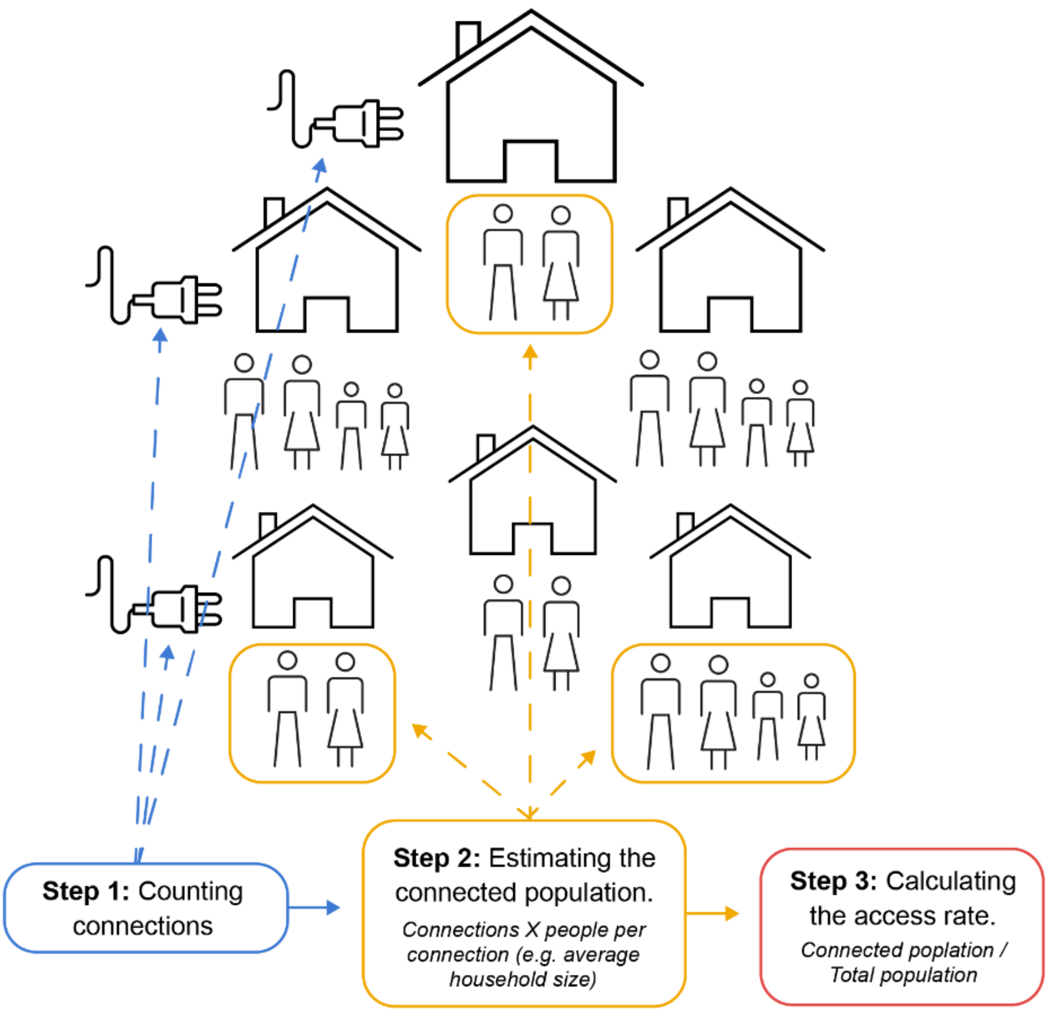


IEA. CC BY 4.0

Household sizes and population assumptions

After tabulating the number of connections, household sizes are used to estimate the number of people benefitting from each connection, which then can be used to calculate total access rates (see figure below). However, using the national average household size can contribute to large differences between supply-side data estimates and survey-based approaches. Since household sizes can vary substantially, especially between rural and urban areas, we recommend using average household size by jurisdiction, when possible.

From counting connections to computing access to electricity rates



IEA. CC BY 4.0

Countries should use the latest population estimates from national statistics offices but in the absence of those they can refer to international organisations (e.g. UN population data). Population data is prefilled by the IEA in its template for access data collection.

Guidance on data for household size/people per connection

Topic	Notes
Main data sources	• National statistic offices (censuses, household surveys). • International organisation such as the UN World Population Prospects (UNWPP).
Time frame	Most recent available for whichever source is used. It needs to be matched with connection data as much as possible. Surveys are not run every year but national statistics offices often provide estimates for years in between surveys.

Challenges and best practices
Matching geographical household sizes with connections: Sometimes the availability of household sizes by region (jurisdiction or rural/urban areas) does not match the connections' data geographical split (as reported by energy distributors) and in other cases only total national household sizes are easily available from national statistics offices. The administrative entities in charge of tracking access to electricity must work with data providers (e.g. electric utilities) to agree on common geographic definitions.

3.2. Data validation

During the validation process, raw data will need to be compared, harmonised and checked for accuracy. Developing standard procedures, check lists and data tools helps makes the validation process more thorough and robust. It also ensures a uniform approach and methodology is implemented by the different people involved in the data validation process. When problems are found, the administrative Entity should take actions, including:

- Log the identified data issues and their status (e.g. pending, solved) in metadata.
- Communicate with data providers to clarify and correct issues.
- Correct the issues through new submissions and updates or, only when necessary, with estimations.
- Re-run validation procedures.

Not all problems will be caught during the first data validation phase so a second round is typically required with all resubmitted data. A process should also be in place to deal with corrections to published data when users find errors, including a clear, public record of the correction (called a corrigendum), which explains the reasons for the correction.

It is important to allocate sufficient time and resources for data verification, and to embed checks throughout the process. For instance, questionnaires sent to data providers can include an automated first set of checks that the data providers can use before submission. The IEA recommends four types of data validation checks: coverage/definition checks, internal coherence, consistency with external sources, and plausibility checks, with detailed guidelines outlined below.

Data validation checks:

Methodological documentation and definitions need to be made easily available to data providers in order to limit potential errors. These are best integrated into the data questionnaire templates (or other tools) sent to data providers and embedding units.

Coverage/definition checks verify that data collected matches with official definitions and coverage. These checks include:

- **Boundary/perimeter verification:** Verify connections data only includes the residential sector. Other boundary verification includes the technology split. For example, confirm that off-grid systems smaller than 10 Wp, such as solar multi-light systems and lanterns, are not counted in access rates but tracked separately. If a split is reported, verify that each category such as grids, mini-grids, SHSs and other stand-alone systems exclude the other ones.
- **Time frame verification:** Confirm the data covers the specific period of reference. For example, that data is for the calendar year and not the fiscal year.
- **Unit verification:** Verify submitted data are in the expected units. For example, are connection data provided in terms of connected households and not connected population? Are the data submitted in millions, thousands, or other units?
- **Geographical verification:** Corroborate that data covers the expected geographical area according to the country's jurisdictional definitions.

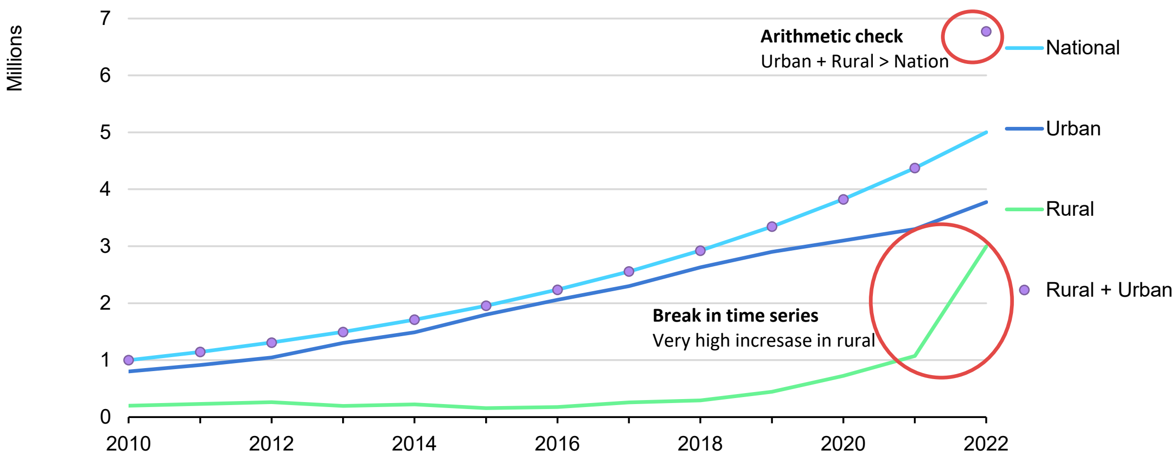
Internal coherence checks verify relationships among the data. These include:

- **Arithmetic checks:** Verify that the data all adds up. An example is verifying that total connections are equal to the sum of connections by type (e.g. grid, mini-grid, stand-alone) and by area (e.g. rural and urban). Another example is that multiplying the rural and urban access rates against the rural and urban populations, respectively, they should sum to the total number of people with access at the national level. Double-counting caused by the use of backup units is one of the issues that can cause arithmetic problems (see Section 3.1 box, "Managing common connection counting challenges").

- **Time series check:** Ensure there are no discontinuities in the data. Breaks in time series are often due to changes in methodologies and definitions from data providers or from changes in the reporting framework. It is recommended in this case to recalculate and update historic data according to the new reporting definition and methodologies to ensure the usefulness of data for analysis otherwise clear documentation explaining the breaks in the form of metadata needs to be made available to data users. Sometimes a break in a time series could be explained by a specific event such as the introduction of an off-grid electrification programme or completion of a major grid extension project, which should be noted in the metadata. In general, decreases in the number of connections are very uncommon and suspicious, and should be checked. Additionally, if the number of connections or access rate remains the same year on year, or the rate of improvement remains the same over multiple years, these may indicate placeholder data or reporter error.

In the figure below an example of a time series issue (i.e. time-series break for rural connections) leading to an arithmetic check issue (i.e. national connections lower than the sum of rural plus urban) is illustrated. The arithmetic check in this specific case would reveal that there is an issue but not where it is coming from while the analysis of the time series suggests that the problem could be coming from rural connections data because of a sudden spike, which could be due to a revision in the way the data provider assessed off-grid connections. However, there is the possibility that the time series break is justified (e.g. an acceleration of an off-grid programme) and data providers need to be contacted to confirm. If the spike in rural connection is verified and confirmed, national connections must be recalculated, otherwise rural connections can be estimated as the difference.

Example of internal consistency checks on the number of grid connections



IEA. CC BY 4.0

Consistency with external sources checks ensure that the data collected are in-line with comparable data from other sources, and that differences can be explained. For example, the IEA access to electricity data differs from data produced by the World Bank mainly because of different sources and methodology.

For supply-side access to electricity statistics, good points of comparison include:

- The results of national household surveys or censuses on access to electricity.
- The Tracking SDG7 – The Energy Progress Report 2022 on access to electricity indicators compiled by the World Bank from household surveys and censuses, although these datasets rely on estimates between years for when surveys are not conducted.
- Industry association data such as GOGLA and AMDA for SHSs and mini-grids, respectively.
- National/international programmes and studies (e.g. national electrification plans mapping or rural electrification agencies).
- Other international and regional organisations data including the UN's population data (e.g. for rural and urban population and household sizes).

Any conflicts should be discussed with other parties before being changed. Harmonising to external data points carries the risk of perpetuating out-of-date data and trends, and creating "self-fulfilling" prophesies.

These cross-checks, if done well, can help improve the robustness of data by triangulating against various different sources. For instance, household survey data can provide more details/qualitative information than data on connections and systems alone. An interesting example comes from Ethiopia, where the Ministry of Water, Irrigation and Energy (MoWIE) put in place a consistency check to compare grid connection data submitted by utilities with specifically designed and sampled surveys run from the Central Statistics office. This comparison permits the MoWIE to understand the quality of the grid connection data, to advise the utility on improvements in its data management mechanics and to refine the data when relevant.

Plausibility checks ensure that final data are reasonable and represent reality. This can range from checking the obvious to more nuanced evaluations that rely on the data checker's knowledge of the sector. Examples of plausibility checks include:

- Confirm that no negative values in access rates are reported. This can sometime happen, especially for rural access rates, when they are calculated as residual of national and urban access rates. This is a sign of underlying problems in the data coverage and consistency that might be at the level of household sizes, population or also connections data.

- Verify that access rates are not higher than 100%, which could occur, for example, as a result of significant double counting issues caused by the extensive use of stand-alone systems as backup for the grid.
- Ensure that data reported as zero represents the reality and not just the non-availability. This could occur, for example, when off-grid tracking is not available and reported as zero but the data should have been reported as "N/A" or not available, through metadata and documentation.
- Verify data falls in expected ranges and trends. For example, if data processing led to an access rate either too high or too low from expected values (e.g. is the progress in-line with current national efforts, programmes and investments) or if rural areas access rates are higher than urban areas (in most case we expect the opposite) data must be investigated in more detail to find the source of the problem.

3.3. Data processing: Computing access to electricity indicators

The data processing step combines raw data and other administrative data (e.g. household sizes, populations) to calculate final access indicators. In this phase, it is recommended to use automated calculations built into the data processing spreadsheets or other database tools to ensure consistency. We also recommend storing raw and processed data on multiple servers or a backup server to avoid any accidental loss of original data while processing.

Calculating access rates

The access rate is the primary metric for defining progress on electrification, defined as **the share of the population benefitting from access to electricity** in a certain location (e.g. a country, a jurisdiction, a village, rural area, urban area). It can be calculated by dividing the total number of people living in households with electricity access by the total population of the tracked location.

The IEA proposes a **"Recommended Formula**" which should be seen as a feasible and yet accurate reference to track access to electricity. We also show a "**Basic Formula**" that is considered as a starting point and the minimum standard by which national access to electricity tracking can be based on. These formulas identify which data, assumptions and calculations are needed to compute access to electricity rates.

Recommended formula

The Recommended Formula calculates the national access rate bottom-up from connection data, and can also provide splits by grid, mini-grid or off-grid systems. This formula can be used at the national level or by jurisdiction. The IEA recommends calculating access to electricity indicators by technology and geographical areas.

"Recommended formula" example for rural and urban

$$\textbf{Access Rate}_{\textbf{Rural}} = \frac{\left(\sum_r^n c_{g_r} + \sum_r^n c_{mg_r} + \sum_r^n c_{sas_r}\right) * h_{rur}}{pop_{rur}}$$

$$\textbf{Access Rate}_{\textbf{Urban}} = \frac{\left(\sum_u^n c_{g_u} + \sum_u^n c_{mg_u} + \sum_u^n c_{sas_u}\right) * h_{urb}}{pop_{urb}}$$

$$\textbf{Access Rate}_{\textbf{Total}} = \frac{\left(\sum_r^n c_{g_r} + \sum_r^n c_{mg_r} + \sum_r^n c_{sas_r}\right) * h_{rur} + \left(\sum_u^n c_{g_u} + \sum_u^n c_{mg_u} + \sum_u^n c_{sas_u}\right) * h_{urb}}{pop_{total}}$$

Where:

- c_g = single grid connections – r= rural, u = urban
- c_{mg} = single mini-grid connections – r= rural, u = urban
- c_{sas} = single standalone connections – r= rural, u = urban
- h_{urban} = the urban household size or number of people per connection in urban areas
- pop_{urb} = the urban population
- h_{rur} = the rural household size or the number of people per connection in rural areas
- pop_{rur} = the rural population
- pop_{tot} = total population
- $\sum_u^n$ = summatory of all urban household connections
- $\sum_r^n$ = summatory for all rural household connections
- r = rural connection
- u = urban connection
- n = total number of connections (rural or urban)

Note: It is important to note that access rates by technology (e.g. mini-grids) also need to be calculated and reported separately. For example, SASs are treated as a single category but it is recommended to split those connections as relevant as proposed in Section 3.1, “Counting stand-alone system connections”. Similarly, for geographical disaggregation any different split by location can be applied as needed by key data users.

Currently, many governments have been calculating access to electricity rates using a more simplified calculation, or the Basic Formula, which is described below.

Basic Formula

$$\textbf{Access Rate}_{\textbf{Total}} = \frac{\left(\sum_i^n c_{g_u} + \sum_i^n c_{mg_i} + \sum_i^n c_{sas_i}\right) * h_{total}}{pop_{total}}$$

Where:

- c_g = national grid connections
- c_{mg} = national mini-grid connections
- c_{sas} = national stand-alone connections
- h = the average national household size
- $\sum_i^n$ = summatory of all households' connections
- n = total number of connections

The Basic Formula should be seen as an approach of last resort when gathered data is not sufficient to use the Recommended Formula.

Below provides an example of calculating the access rate by the Recommended Formula. The rate is first calculated by connection type, then summed to a national total. It is important to note that if certain data are not available for a specific category (e.g. stand-alone systems data not yet collected/processed), it is critical to maintain the category in the database and mark the data points as *not available (n.a.)*, while calculated totals (e.g. Total connections) need to read *n.a.* data points as zeroes.

Example of data processing steps

Type of connection	Connections (million)	Household size (people per household)	Population with access (million)	Total population (million)	Access rate (%)	Metadata
Grid	5	4	20	50	40%	
Mini-grid	1	4	4	50	8%	
SAS	n.a.	4	n.a.	50	n.a.	Not available
TOTAL	**6**	**4**	**24**	**50**	**48%**	**Exclude SASs**

Note: The table above shows an example that represents an extract of a database only for rural connections, A complete database would include additional "Types" of connections and geographical areas. Blue columns represent raw data collected, and green represents computations and the grey is metadata.

Creating local indicators

Tracking access to electricity by location can help inform electrification plans, especially to understand where unelectrified populations are located and how many are based in cities or villages with a critical density that makes them financially significant to electrify. At a minimum, the IEA recommends tracking access by jurisdiction and by rural-urban split. However, to better leverage satellite and geospatial datasets in access tracking, geo-tagging connections or at least logging connections by official town/village names can support improved planning.

Rural and urban access rates should ideally reflect standard definitions of population density for rural and urban areas, which is most useful to determine the density of dwellings to connect and the size of the community they are in. However, many regions classify specific jurisdictions as rural or urban, even if those jurisdictions include smaller rural or peri-urban communities. These definitions should be made clear in the metadata. It is also important to note any change in urban/rural definitions, especially if jurisdictional boundaries are redrawn or are reclassified as urban or rural, which could result in swings in statistics. The Entity in charge of tracking access to electricity must work with the national statistics office to align on standard geographical definitions, require data providers to report in the same terms, and record the definitions as well as any revisions in the metadata.

Where possible, the geolocation of each new connection should be logged, at a minimum with the name of the village/town (sub-jurisdiction), but could also use geo-tagging of co-ordinates. New grid connections often already includes geolocated data for lines and connections, but this information may not exist for historic connections. In the case of mini-grids, the data should be available based on the mini-grid location itself or the permitting application. The task can be more complicated when tracking sold stand-alone systems, but the needed geospatial data are most often available under PayGo or similar models where the location of the systems is tracked.

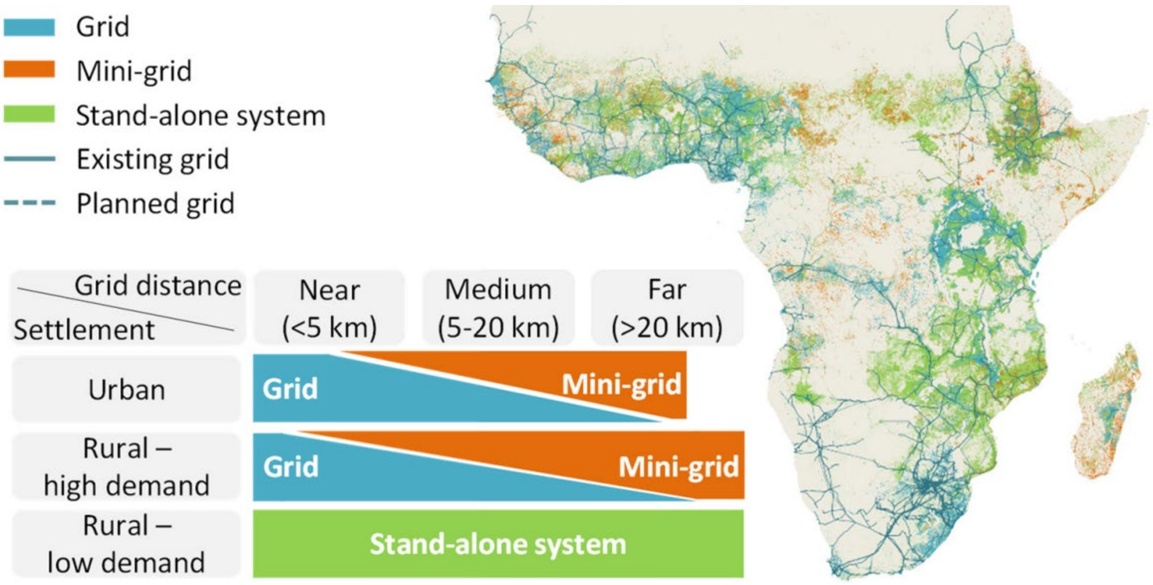
Geolocating new connections allows connection data to be collated with national statistics office data containing names and co-ordinates of villages and towns, typically maintained for surveys. Other datasets, including Google Maps, OpenStreetMap and the GeoNames initiatives, can also be used to collate this data.

Increasingly, new geospatial datasets are available for Africa and can be a useful counterpoint to local data. For example, night-time lighting satellite imagery can be advantageously combined with access statistics, as well as other information sources gathered from smartphones and smart meters, to provide a better picture of energy access and its usage. However, this should be seen as a complement to primary data as they represents a much higher level of quality and precision. (See Appendix, Other sources of data, “Remote sensing capabilities”.)

Developing and maintaining up-to-date geospatial datasets enables more countries to shift to GIS access planning tools, which are quickly becoming the standard for supporting electrification efforts. Many GIS modelling tools and datasets are available to support access planners, which the IEA maintains a list of on its website as a part of this programme. The IEA’s Africa Energy Outlook 2022 provides an example of how GIS tools can be used to determine a context-specific, least-cost electrification plan. This is done by overlaying several layers of information and performing an optimisation that considers the location of non-electrified settlements, their respective distance from the nearest national power grid and the renewable resources present on site, among other variables.

This type of analysis is very useful for performing large-scale analyses to inform individual country policies and educate both the private sector and international organisations involved in energy access about opportunities in different markets.

New connections by type in the IEA Sustainable Africa Scenario



IEA. CC BY 4.0

Source: IEA (2022), Africa Energy Outlook.

The IEA is working to improve geospatial data and tools for integrated electrification planning, in its Power Africa's Data-Driven Electrification Planning programme. A newly constructed tool aims to better estimate demand by household in different countries and regions, both when first connected and several years after having electricity. This then allows for better estimates of electricity demand based on satellite images of buildings within a community, which helps companies better size grid extensions or mini-grids, saving costs and improving the investment case for such projects.

3.4. Data dissemination

Data dissemination should be timely, in a useable format, easily understandable and accessible to all users. In the section below recommendations on what, when and how to report these criteria are provided.

Data relevance, accuracy and timeliness: What to report and when

In terms of **relevance**, the Entity in charge of tracking access to electricity must align data production and dissemination to the specific and evolving needs of relevant data users (e.g. inclusion of split by technology).

Ensuring the data are reproducing the reality on the ground with a high degree of **accuracy and reliability** is also very important. Implementing data verification and revision procedures is a vital step during the data dissemination process.

Timeliness and punctuality should align with data users' specific needs and resources must be put in place to respect publication timelines. For access to electricity indicators produced from supply data, timeliness is a key advantage versus surveys, with countries such as Mozambique producing quarterly access rate estimates. It is recommended, at a minimum, to disseminate yearly access to electricity rates no later than six months of the next year. Time series must be disseminated together or made available with the latest data release to provide a comprehensive view of trends.

Recommended list of indicators to publish

Coverage indicators	Coverage priority	Timeliness recommended	Periodicity recommended
Access to electricity rate[4]	Essential	Year minus one or better	Annually/Biannually/ Quarterly
Number of connections	Essential	Year minus one semester or by quarter	Annually/Biannually/ Quarterly
Number of people without access	Strongly recommended	Year minus one semester or by quarter	Annually/Biannually/ Quarterly

[4] Defined as the share of the population living in households with access to electricity. However, access can be reported as the share of households with access to electricity. It is important to report how the access rate has been calculated for transparency.

Coverage indicators	Coverage priority	Timeliness recommended	Periodicity recommended
All the above split by location: - Rural and urban - Jurisdiction	Strongly recommended	Year minus one	Annually
All the above split by technology: - Grid - Mini-grid - Stand-alone	Strongly recommended		
Further split of off-grid connections by source - Diesel mini-grid - Solar PV/Diesel hybrid mini-grid - Solar PV + battery mini-grid - Micro-Hydro - Fossil generators - Solar home systems	Recommended		
Split of stand-alone systems by size (e.g. SHS by capacity size)	Recommended		

Data accessibility and clarity: How to report

After creating quality and comprehensive datasets, the data must be easily available and accessible to all users to fully maximise the value and transparency for everyone. For this reason, it is recommended that the Entity in charge provide release calendars showing when, where and how to access data. Data should be publicly released (e.g. on a website or a report) but also directly provided to the relevant data users. It is also important to communicate any data update releases to users so that they are aware new information is available and can access the most up-to-date versions.

There are different dissemination formats that can be used to release access to electricity data, including published reports, online downloadable databases and dynamic tracking web pages. In general, a combination of different convenient formats can satisfy different data users' needs – e.g. while reports provide additional information and analysis of the data, downloadable databases are very user-friendly to start new analysis.

Two good examples of disseminating access to electricity data are from Rwanda and Ghana. The Rwanda Energy Group (REG) is the Entity in charge of tracking access to electricity in the country and disseminates the data using different means, including in their annual reports, of which older versions are made available, and on a dedicated web page, which provides access rates split by grid and off-grid on a national level and by district with a two-month delay.

The Ghana Energy Commission publishes in its energy statistics publication the national, urban and rural access rates in tabular formats, including time series as well as access rates by region in the form of maps.

Although both formats provide insightful data for studying access to electricity, so far neither make available clear definitions, methodologies and metadata. So, for example, it is not possible to know if the access rates published by the Ghana Energy Commission are calculated with off-grid access captured or are not included.

Coherence, comparability, and transparency of released data: Metadata

The standards set forth in this Guidebook, if followed, are meant to ensure that access indicators are **coherent and comparable** over time and between different regions. However, there will always be small differences between how one region treats their data versus another. This is why **metadata** (i.e. data about data) is important, so people have a **transparent** record of how these data may diverge from standards.

Maintaining metadata is a fundamental practice that can provide a detailed account of where data is sourced, the conventions and definitions that are used, and other components that can affect the data. Metadata should be made available with any final published statistics. A good example of supporting metadata is the IEA energy prices dataset documentation. This includes data availability files clearly showing the coverage of the dataset as well as a specific file providing information on sources, revisions, and other relevant details to understand the data.

Example of dataset using metadata

Indicator	Area	Technology	Year	Units	Value	Source	Collection Method	Qualifiers
Access to electricity (share of the population with access)	Rural	Grid	2020	%	85	PRIMARYSUPPLY	DIRECT	OBSERVED
Access to electricity (share of the population with access)	Rural	Mini-Grid	2020	%	5	PRIMARYSUPPLY	DIRECT	OBSERVED
Access to electricity (share of the population with access)	Rural	SAS	2020	%	0	OTHERSOURCE	OTHER	NOTAVAILABLE
Access to electricity (share of the population with access)	Rural	Total	2020	%	90	OTHERSOURCE	OTHER	NONCOMPARABLE

The following list provides recommendations on which metadata and documentation should accompany access to electricity indicators:

- **Data Source:** The specific provider of each data point or set of data and, if available, the path (i.e. hyperlink or attachment) to the backup documentation. For example: “Connection data come from the national electric utility annual reports back to 2015; prior to that data were estimated with the use of surveys.”

- **Units:** Clearly identify what reporting units each indicator uses. For example, detail whether the connection data refers to million households or thousand people. This information goes in the documentation but also with any other data dissemination format as in tables or data files.
- **Definitions and methodology used:** This is very important information for data users as it explains what the indicators published refer to. It must include information such as:
  - The boundaries/perimeter/coverage used to calculate indicators (e.g. which types of connections are included in access rates).
  - The definition used for access to electricity (e.g. is it the share of population or the share of households).
  - The methodology used, for example: "Every five years results from household surveys are used to calibrate supply-data and re-estimate off-grid connections."
  - What is included in the performed estimations. For example, include information such as: "Prior to 2018 residential connections are estimated on the basis of low-voltage connections."
- **Time:** The time frame (e.g. year, half year or quarter) to which the data refers to. Traditionally, the electricity access data corresponds to the previous calendar year (1 January to 31 December) but in some cases data might be available on a fiscal year basis (e.g. April to March). Metadata needs to clearly provide this information.
- **Confidentiality:** If for confidentiality reasons some data is not available or aggregated this needs to be explained in documentation. For example, data on mini-grid connections for a specific jurisdiction/district might be confidential due to the fact that only one mini-grid player is active in the area, and this would imply releasing sensitive data to the competition. Being able to treat confidentiality at the dissemination output level would encourage data providers to share more data with the Entity in charge of tracking access to electricity.

In addition to steps detailed above, information provided in the form of separate documentation files and data points in databases can also be linked to specific metadata (e.g. qualifiers). This type of metadata should be directly attached to specific data points (e.g. as an additional column[5]) in the form of a code/value with a specific meaning (see table below). When using metadata codes (e.g. EST = estimated; CONF = confidential), documentation should be include a clear set of definitions for each abbreviation. The IEA's access statistic template includes recommended metadata notes to include along the lines of the table below.

[5] Columns in a database where each line defines only one specific data point includes defining columns as Year, Indicator Name (e.g. connections, access rate) that are fundamental to describe and isolate each data point (i.e. there is only one national access rate for 2022). Other columns that could be called descriptive columns can add important descriptions to the dataset and this include units (e.g. percentage) and other metadata (e.g. source).

Example of metadata structure with legends

Type of metadata	Metadata options
Source	PRIMARYSUPPLY: Data coming from electricity distributors and mini-grid operators.
	SURVEY: Data coming from survey or census results or from national statistics offices.
	OTHERSOURCE: Data coming from other sources not mentioned above.
Collection method	DIRECT: Data is submitted directly to the Entity through a questionnaire or via other means (e.g. direct e-mail exchanges).
	INDIRECT: Data is recovered by the Entity directly from existing publications (e.g. electric utility annual report).
	OTHER.
Qualifiers (mutually exclusive)	OBSERVED: If coming from official sources (linked to Primary supply). This is an important qualifier for zeroes to ensure the zero is a real observed value and not caused by data unavailability (see below).
	NOTAVAILABLE: Data is currently not available. This would define a zero stemming from unavailability.
	NOTAPPLICABLE: The value for this specific data point is not applicable. For example, Solar PV + Battery connection type cannot apply to grid connections.
	CONFIDENTIAL: Data point is not published for confidentiality issues. The data point can be aggregated with other values to ensure totals include it.
	PROVISIONAL: Data point is likely to be revised. This applies often to nowcasted or most recent data relying on preliminary information.
	DERIVED: The data point is derived starting from other existing or collected data. It can be, for example, the residual calculation of rural connections as the difference of national and urban connections.
	ESTIMATED: The data point has been estimated by the Entity tracking access to electricity.
	NONCOMPARABLE: Definition of data differs from the specific definition. For example, if grid connection data include also mini-grid connections as the data provider counts both of them but is yet not able to provide a split.

Other data and indicators

The Entity in charge of tracking access to electricity is encouraged to expand over time the number of indicators collected, calculated and disseminated. The table below provides additional indicators that can provide a multi-dimensional perspective of the quality, availability, affordability and use of electric services.

Multi-dimensional indicators

<table>
<tr><th>Type</th><th>Indicator</th><th>Sources, challenges and practices</th></tr>
<tr><td rowspan="2">Location</td><td>Geotagged connections</td><td>In principle, electricity distributors know where their customers are located for technical and billing purpose, and this is even more true for mini-grids and for SHS PayGo models. However, exploiting this data requires additional resources and efforts to develop and maintain an up-to-date new system. The level of granularity can vary from a simpler urban/rural split or jurisdiction to a village level and finally to GPS co-ordinates, with the cost and resources required increasing.
<table>
<tr><th></th><th>Area (urban/rural)</th><th>Jurisdiction</th><th>Village, town, location</th><th>GPS co-ordinate</th></tr>
<tr><td>Customer 1</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Customer 2</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Customer 3</td><td></td><td></td><td></td><td></td></tr>
<tr><td>Customer 1000</td><td></td><td></td><td></td><td></td></tr>
</table></td></tr>
<tr><td>Other Geographic Information System (GIS) data</td><td>The periodical update and maintenance of GIS data on infrastructures (e.g. grid transmission and distribution lines, mini-grids, resources, population densities) and connections is key for effective electrification planning.</td></tr>
<tr><td rowspan="2">Quantity</td><td>Number of connections by consumption tier And/or monthly or yearly consumption</td><td>Energy regulators often require electricity distributors to apply different tariffs depending on the household level/tiers of consumption in kWh per month (e.g. social tariffs). Therefore, from the billing data electricity distributors could retrieve information on the number of customers in each consumption tier. This data exists in some form but to be exploitable additional data work and resources are often needed by the energy distribution company. Depending on the needs, this data can be prepared according to the national tariff scheme or specific consumption tiers can be estimated (e.g. for regional or international comparison purposes).</td></tr>
<tr><td>Appliance ownership</td><td>Understanding which appliances households own is a good indicator of the level of electricity and power they consume. This data can be estimated from sales figures (similar to what is done to estimate operative solar home systems) or from surveys results.</td></tr>
<tr><td>Availability[6]</td><td>Average Service Availability Index (ASAI)[7]
The ratio of the total number of customer hours that service was available during the year to the total customer hours demanded.</td><td>Most electric utilities and electricity distribution companies track this information.
$$1-\left(\frac{\sum[RestorationTime(hours)\ X\ CustomersAffected]}{(CustomerServed\ X\ 8760(hours))}\right)$$</td></tr>
</table>

[6] Availability and reliability indicators are generally not disaggregated by connection type or tariff.

[7] Sometimes referred to as the Average Service Reliability Index.

Type	Indicator	Sources, challenges and practices
	System Average Interruption Duration Index (SAIDI) The average duration in minutes of a sustained interruption (>five minutes) for the average customer during the year.	Most electric utilities and electricity distribution companies track this information. $$\frac{\sum[RestorationTime(min)\ X\ CustomersAffected]}{CustomersServed}$$
	Customer Average Interruption Duration Index (CAIDI) The average time to restore service to a customer following a sustained interruption (>five minutes).	Most electric utilities and electricity distribution companies track this information. $$\frac{\sum[RestorationTime(min)\ X\ CustomersAffected]}{CustomersAffected}$$
Reliability	System Average Interruption Frequency Index (SAIFI) The average number of times that a system customer experiences an interruption (>five minutes) during the year.	Most electric utilities and electricity distribution companies track this information. $$\frac{\sum[CustomersAffected]}{CustomersServed}$$ OR, Alternatively $$\frac{SAIDI}{CAIDI}$$
	Customer Average Interruption Frequency Index (CAIFI) The average number of interruptions (>five minutes) per customer interrupted per year.	Most electric utilities and electricity distribution companies track this information. $$\frac{Interruptions}{CustomersAffected}$$
Quality	Over/under voltage occurrences The average times the customer experiences over/under voltage.[8]	Most electric utilities and electricity distribution companies track this information. $$\frac{\sum[Overvoltages + Undervoltages]}{CustomersServed}$$

[8] The over/under voltage threshold (i.e. +/-5%) should be reported with this indicator in order to understand the standard in place for the grid operator.

Type	Indicator	Sources, challenges and practices
Safety	Deaths/Injuries per year The total number of deaths and injuries related to electric services.	Most electric utilities and electricity distribution companies track this information. $$Deaths + Injuries$$
Affordability	Connection charges Average amount paid by a customer to get a first connection to the grid.	Utilities or energy regulators have this information easily available. It is important to track and disseminate connection charges with any relevant details such as specific incentives. or difference in fees depending on the distance from existing stations, poles or areas. It is also important, when possible, to track connection costs for off-grid solutions such as mini-grid connection fees, the market cost of solar home systems (by size), the deposit cost in case of PayGo or similar business models.
	Electricity tariffs by consumption tier Variable charge (per kWh) and fixed charge for the different consumption tier.	Utilities or energy regulators have this information easily available. It is important to collect and disseminate the detail of tariffs including taxes, differentiation of tariffs by consumption levels/tiers, fixed costs and any other relevant incentives or regional differentiation. It is also important to track and publish the tariffs/costs of off-grid electricity such as mini-grids and stand-alone systems. For mini-grids, tariffs might be similar to the regulated grid while for stand-alone systems the periodic fees households pay in the case of PayGo or similar business models can also be tracked (possible by system size).
	Customer cost of service as % of median income The cost, including fixed charges, taxes, fees, capacity charges, etc. of powering a minimum energy service (e.g. the IEA essential bundle) expressed as a % of median income.	This requires combining the data above on connection costs and energy costs with household income data. Defining income for the households yet to gain access and those that have gained access but have affordability issues might be challenging. Combining household income data and connections by location (e.g. village) can be fundamental for better electrification planning and for the private off-grid sector. A solution that provides interesting affordability indicators is using household income by percentile[9] and estimating the share of the population able to afford a specific minimum energy service. To do that a threshold on the percentage of income above which paying for electricity service becomes unaffordable must be assumed. It is common practice under current policies to set this threshold at 5% of the household income for electricity affordability.

All of the above is focused on household energy, however tracking indicators for non-household access in public buildings, hospitals, productive uses, and businesses is also important. For example, knowing the location of un-electrified health centres, telecommunication antennas, mining centres or schools is key precursor information for electrification planners. This is often of great interest, as these entities can be important anchor customers which improve the economics of interconnecting a certain region or community. While the above guidelines can be easily adapted to non-household access tracking, the IEA plans to release more details on data collection and estimation for these productive end uses in a subsequent version of these guidelines.

[9] The Poverty and Inequality database from the World Bank provides per capita income data by percentile. In absence of national data, the World Bank estimates can be used to calculate affordability indicators.

Appendix

Other sources of data

Remote sensing capabilities

In addition to supplier-centred and survey-based data, remote sensing technologies offer a complementary set of methods to employ in measuring electricity access. Satellite data along with other georeferenced datasets, such as network infrastructure, customer locations and other overlays, have become indispensable tools for electrification planning, using location as the anchoring variable. It is generally accepted that “spatially explicit data on electricity access and use are essential for effective policy-making and infrastructure planning in low-income, data-scarce regions,” according to a report on electrification in sub-Saharan Africa. The pertinent question for this Guidebook is: Can satellite data provide easy, low-cost and accurate measurements of electricity access?

Night-time lights (NTL) have already been demonstrated as an effective proxy for many types of indicators including migration, income inequality, regional gross domestic product (GDP) and electricity access, with a large body of literature indicating that locations where remote sensors detect human settlement but not light are likely areas where people live without access to electricity. Through a process that combines NTL data with other critical information regarding urban and rural population density and other inputs it is possible for GIS models to render estimates not only of the binary variable of access to electricity, but also the intensity of consumption at a resolution of one square kilometre. Current practice can essentially discern activity corresponding to larger capacity isolated grids but may miss smaller mini-grids and geographic concentrations of single-family generation solutions. Publicly available NTL data often has a lower temporal and geographical resolution than other datasets, so using higher resolution data would improve the quality of the estimations.

Importantly, the relationship between electricity consumption and remote sensing data must be calibrated using survey data to account for local effects, such as the cultural propensity to engage in outside activities after dark or municipal-level decisions regarding investment in public street lighting. Research that still needs to be done on calibration for each location means that in the very near term satellite data may be used in a limited capacity as a validation check on other datasets (population surveys, enterprise customer reports), particularly if there is mismatch in one geographically delimited area but is not likely to supplant those other sources for statistical and international reporting purposes.

With respect to current approaches to measure access rate, many administrative sources use geospatial references; however, these data are not yet used as a method to generate access to electricity statistics but instead as a tool used for estimating, evaluation, and validation of statistics in the framework of electrification planning. Presently, studies show that NTL approaches to evaluate access to electricity are broadly consistent with recently published province- and national-level statistics.

In conclusion, satellite-derived estimates and other GIS data are predominantly utilised for planning; however, they have enormous potential to become important in the tracking of electricity access. A degree of methodological validity through peer review already has been demonstrated. As continued research establishes tighter correlations between remote data and ground-level access data, and perhaps also as technology advances (e.g. imaging devices, processing capabilities), satellite imaging is expected to reduce the time, cost, and difficulty of producing timely and reliable data.

Mobile phone use as a proxy for access to electricity

The high penetration of mobile telephone coverage has provided an alternate data source that can be used to understand access to electricity. In the past few years, studies have shown how this data can be turned into critical insights to support demand estimation and electrification planning. Mobile phone metadata – including app usage, location and charging patterns/characteristics – provides proxy indicators of population density, economic activity and social characteristics of a population.

Moreover, this data, which is available in real-time, can be used to estimate electricity consumption per capita, permitting access rates to be assessed at a high level of spatial and temporal resolution, and at a considerably lower cost than survey data. Similar to issues with remote sensing, data from mobile phones must first be locally calibrated using ground-level survey data. One of the current stumbling blocks to using phone data to estimate electricity access is solving privacy concerns and ensure individuals can't be identified in the datasets.

Annexes

Glossary and definitions

Centralised system: Refers to connections to the electrical grid. Typically, the source of electricity generation (a large power plant) is able to provide electricity to long-distance consumption sites.

Commercial connection: Electricity connections classified as commercial by utilities are those that use power for business activities and have their own billing tariff. Some commercial and industrial connections are medium-voltage.

Decentralised system: Refers to off-grid connections (see off-grid connections below).

Grid connection: (Also referred to as "on-grid"). A connection to an electrical grid. Electricity from electrical grids is generated from traditional power plants which use resources such as coal, oil, natural gas and nuclear technology, or from renewable sources such as hydro, wind, solar and geothermal sources, to produce the electrical power distributed on the electrical grid.

High-voltage connection: High-voltage electricity connections are for electric transmissions and can range from 100 kilovolt (kV) all the way up to greater than 1 000 kV. High-voltage is used for long-distance transmission to reduce grid losses.

Household connection: Electricity connections needed for household use such as lighting, cooking, heating, cooling, communication and entertainment. Individuals living together in the same house are considered a household. Sometimes, household connections can use the supplied electricity for non-energy intensive productive or commercial uses. This definition serves as the meaning of "household connections" referred to in this Guidebook.

Illegal, informal or unofficial connections: Electricity connections that are established unlawfully such as through meter tampering, secondary connections (connecting via a neighbour), bribery, among others, in an effort to acquire electricity at a lower cost or at no cost. Illegal connections established through meter tampering and secondary connections can be dangerous due to the absence of safety devices. Through poor wiring and electrical overload, these illegal connections can cause damage to household appliances, the entire electrical system, and potentially the people living within the home.

Institutional connection: (Also referred to as Community Infrastructure) Electricity connections needed for community purposes such as street lighting, centralised water pumping, health facilities, educational facilities, local government offices, community buildings (such as places of worship), and other shared infrastructure.

Low-voltage connection: Low-voltage electricity connections are typically for residential and retail users. Low-voltage transmission is usually for household and/or small commercial distribution at single phase and is generally capped at 1 000 volts (V) for the power lines. Typical users have end-use requirements of 120 V to 240 V at 15 to 20 amperes (Amps).

Medium-voltage connection: Medium-voltage electricity connections are for industrial and commercial users. Transmission lines for medium-voltage are between 1 kV and 70 kV and users typically have needs that require three-phase power.

Metered connection: Electricity consumed from the electrical grid that is recorded using metering equipment. The quantity of electricity used is measured and reported through a meter.

Mini-grid connection: An isolated electricity distribution network (though they may be interconnected) which supplies electricity to a small set of customers within a local service area. Mini-grid systems often rely on solar, wind, biomass, and/or diesel generators for electricity generation.

Multiple metering: (Also referred to as sub-metering). Refers to the practice of having electric connections with a secondary meter downstream from the primary metering system connected to the electric distribution companies. This allows electric services to be centralised, with a single customer responsible for the collection of service fees from individual users.

Off-grid connection: Electricity obtained from a source other than a connection to a main or national grid. This includes mini-grids and stand-alone systems. Electricity is acquired through non-renewable and renewable resources such as wind, solar, hydro, among others.

PayGo: (Also referred to as Pay-as-you-go). PayGo contracts or subscriptions for off-grid solar systems are aimed at reducing the initial upfront cost for solar energy access by allowing customers to make instalment payments to purchase time units for using solar electricity instead of paying upfront for solar home systems. PayGo contracts can be divided in two categories: **rent-to-own** or other leasing arrangements; and **energy-as-service**. Under a rent-to-own scheme, customers typically pay an upfront deposit for the system and appliances, pay back the full cost of the system through periodic instalments and at the end of the contract customers own the SHS. Under the energy-as-service model, the household enters a long-term contract for energy supply and the operator oversees maintenance and replacement.

Solar home system (SHS): A form of stand-alone systems that use solar photovoltaic power to supply electricity. In a stand-alone/SHS, the solar panels are not connected to an electrical grid; they generate electricity to charge batteries that store the energy to be used by appliances such as radio, lighting bulbs, fans etc. They exist in different sizes from 10 Wp to 49 Wp and are able to provide pre-electrification levels of access while systems from 50 Wp can provide first access to essential energy services as defined by the IEA.

Solar lanterns: Very small-scale solar-based solutions (below 3 Wp), typically a single-light point with or without phone charging capabilities. Energy is provided through solar PV and all system components are integrated within a single, usually handheld, device. Aside from providing important improvements in the welfare of the household, these systems are not considered as providing the minimum electricity access due to the limited energy services (mostly to lighting and phone charging) they can provide.

Solar multi-light system (SMLS): Small-scale solar-based solutions (from 3 Wp to 9 Wp) providing more than one light point, phone charging and, in some cases, a radio. Aside from delivering important improvements in the welfare of the household, these systems are not considered sufficient to provide basic electricity access due to the limited energy services (mostly limited to lighting and phone charging) they can support.

Stand-alone system (SAS): A form of off-grid technology that supplies electricity to single consumers (e.g. households) and that is not connected to a main or national grid nor to a mini-grid. This can include SHS or other renewable technologies as well as petrol generators.

Unmetered connection: Electricity consumed from the electrical grid that is not directly recorded using metering equipment. In an unmetered connection, electricity is disbursed by the electrical grid, and consumed by the consumer, but the quantity of electricity is not measured. The consumer is often billed on a lump sum sometimes based on the capacity (in kW) of the connection.

Abbreviations and acronyms

AMDA	African MiniGrid Developers Association
Amps	amperes
CAIDI	Customer Average Interruption Duration Index
DHS	Demographic and Health Surveys
DSO	distribution system operators
ESMAP	Energy Sector Management Assistance Program
GDP	gross domestic product
GIS	geographic information system
IEA	International Energy Agency
IRENA	International Renewable Energy Agency
LSMS	Living Standards Measurement Study
MICS	Multiple Indicator Cluster Surveys
MoE	Ministry of Energy
MoI	Ministry of Infrastructure
MTF	Multi-Tier Framework
NTL	night-time lights
PayGo	pay-as-you-go
PV	photovoltaic
SAIDI	System Average Interruption Duration Index
SAIFI	System Average Interruption Frequency Index
SDG	Sustainable Development Goals
SAS	stand-alone system
SHS	solar home system
SLS	solar lighting systems
SME	small and medium-sized enterprise
UN	United Nations
UNWPP	United Nations World Population Prospects
USAID	U.S. Agency for International Development
WHO	World Health Organisation
WHS	World Health Surveys

Units of measure

kV	kilovolt
kW	kilowatt
kWh	kilowatt hour
V	volt
Wp	watt peak

IEA Publications
International Energy Agency
Website: www.iea.org
Contact information: www.iea.org/contact

Typeset in France by IEA - Mars 2023
Cover design: IEA
Photo credits: © shutterstock