Stephen J. Lee, Energy for Growth Hub
Jay Taneja, UMass Amherst

November 2024

Tracking the Modern Energy Minimum using the Open Energy Maps machine learning platform

The Energy for Growth Hub's "Modern Energy Minimum" (MEM) provides more ambitious thresholds for the provision of electricity services beyond current goals for universal electricity access. Such thresholds can better incentivize international resource deployment towards economic growth and productive uses of energy. The MEM reflects global targets of at least 250 kWh of annual residential electricity consumption for every person and country-level averages of at least 750 kWh annual non-residential per capita electricity consumption.

Progress towards the *non-residential* part of the MEM can be evaluated using the IEA's World Energy Statistics and Balances database and other resources providing country-level non-residential electricity sales and population data. Unfortunately, assessing *residential* MEM attainment isn't as straightforward. Energy authorities do not report the percentage of residential customers exceeding the MEM residential threshold. This may be due to a number of reasons: customers are not always metered and utilities do not always keep comprehensive consumption records. In addition, leading multinational organizations like the World Bank and International Energy Agency do not focus on detailed electricity consumption metrics beyond aggregated sales and access rates.

To address this gap, we employ the Open Energy Maps (OEMaps)[1] machine learning (ML) platform to produce estimates of electricity consumption across the majority of buildings in Africa. OEMaps employs globally available remote sensing features describing the characteristics of individual buildings and trains models characterizing building-level electricity access (Figure 1)[2] and electricity demand (Figure 2)[3] using infrastructure data and historical metered consumption data. We combine such access and demand estimates to infer existing consumption, and apply household size assumptions to estimate country-level progress towards the Modern Energy Minimum's residential consumption threshold (Figure 3). We validate our approach using statistical methods and we're currently expanding our estimates to create a fully auditable 'living map' of progress at high resolutions globally.

[1] Open Energy Maps™ website: www.openenergymaps.org.
[2] OEMaps estimates building-level electricity access, defined as the probability that it has a grid or grid-quality off-grid connection.
[3] OEMaps estimates building-level electricity demand, defined as a probability distribution of the building's electricity consumption assuming it has a grid or grid-quality off-grid connection.

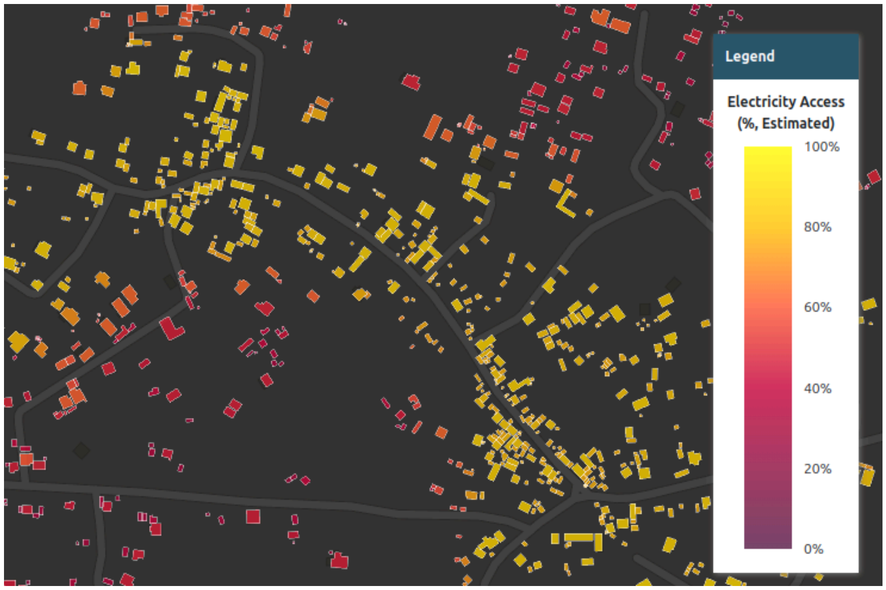


FIGURE 1: OEMaps provides estimates for building-level electricity access (i.e. how likely a building is connected to the grid) at country- and continent-scales. Purple and red building colors reflect low expected electricity access rates, and orange and yellow colors reflect higher expected rates.

FIGURE 2: OEMaps also provides estimates for building-level electricity demand. Building colors correspond to expected values for electricity demand, where blue and green buildings represent lower-demand buildings, and yellow and red buildings represent higher-demand buildings.

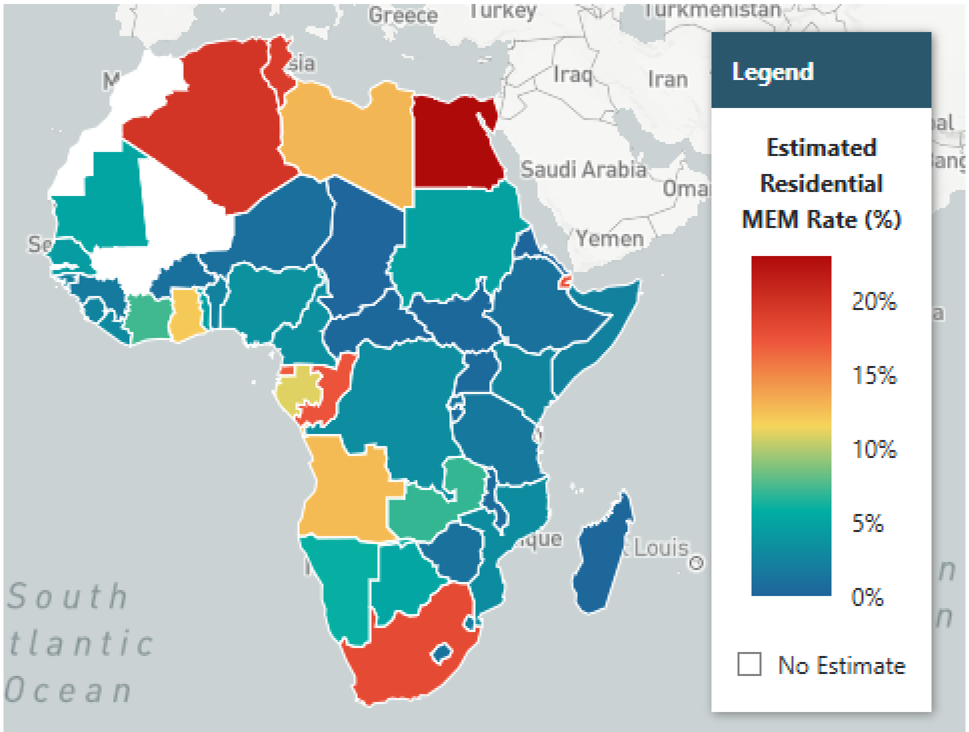


FIGURE 3: We combine OEMaps estimates of building-level electricity access and demand to estimate existing consumption, and apply household size assumptions to estimate country-level progress towards the Modern Energy Minimum's residential consumption threshold.

Leveraging remote sensing data, inter-institutional collaboration, and geospatial ML

We built a first-of-a-kind system for estimating building-level electricity access and electricity demand employing large remote sensing datasets, diverse metered consumption training data, and ML models associated with the OEMaps platform.[4]

- **State-of-the-art remote sensing data pipelines.** We aggregate, process, and join massive geospatial datasets comprising globally-available building-level features. These geospatial datasets include Google and Microsoft building footprints datasets, high-resolution satellite image tiles, land use land cover (LULC), VIIRS nighttime lights, and Ookla internet connectivity speeds datasets.

[4] Lee, S.J., *Multimodal Data Fusion for Estimating Electricity Access and Demand*, Doctoral Thesis, Massachusetts Institute of Technology, 2023; IEA (2024), Africa's electricity access planners turn to geospatial mapping, IEA, Paris https://www.iea.org/commentaries/africa-s-electricity-access-planners-turn-to-geospatial-mapping, Licence: CC BY 4.0

  - In essence, we compile rich remote sensing descriptions of the majority of built structures in Africa. These features are indicative of energy consumption.
  - All else equal, residential buildings are more likely to have electricity access and high electricity demand if they: have larger footprints, are set in more urban neighborhoods, are built out of high-quality materials, are situated on paved roads, have nearby street lighting appearing in nighttime lights imagery, and are collocated with internet users with fast internet connectivity speeds.
  - While human intuition lets us reason about the directionality of remote sensing feature impacts on electricity access and demand, our ML models are optimized to quantitatively characterize building-level electricity access likelihoods and probability distributions for electricity demand.
- **Diverse ground truth data from years of coalition building.** We train our ML systems in areas where we can partner to obtain access to georeferenced infrastructure and metered consumption data. We currently train our systems in regions of East, South, and West Africa. Our partnerships spanning nearly a dozen key organizations avail a one-of-a-kind dataset for model training and analysis.
- **We train our electricity access model using low voltage distribution network topology data and the geospatial locations of metered consumers.** We extract positive class labels for buildings that are on the low voltage distribution network and have an electricity meter. We extract negative class labels for buildings that are distant from the low voltage distribution network and do not have an electricity meter.
- **We designed an application-tailored model to deal with the unique nuances of estimating building-level electricity demand.** Estimating electricity demand proves to be more challenging than estimating electricity access.
  - While we assess grid-quality electricity access as a binary metric, electricity demand can span nonnegative numbers across several orders of magnitude.
  - In addition, we typically have difficulty geolocating consumers in metered consumption datasets. Many low- and middle-income regions do not have mature address geocoding systems (which would allow us to convert addresses to geospatial longitude/latitude coordinates) and utility-provided geospatial coordinates are often noisy: longitude and latitude values for a meter may be tens of meters away from known building locations. As a result, it is often not possible to map a given utility customer to a building in our remote sensing database with high certainty.
  - We address these challenges by building an application-tailored model: the Load Inference through Lightweight Data Fusion (LItLDF) model. LItLDF characterizes demand across multiple orders of magnitude via parameterizing long-tailed probability distributions. Instead of providing a single point estimate for a building's demand, LItLDF provides a probability distribution across a wide range of possible values.

- We account for unknown meter-to-building mappings during training by encoding constraints analogous to Kirchoff's circuit laws within LItLDF. If a meter is within close proximity to multiple candidate buildings, we consider each building to have the potential to account for part of all of that meter's consumption value. This gives LItLDF the flexibility to learn about behind-the-meter connections. Conversely, if a building is within close proximity to multiple meters, we allow its consumption to reflect up to the sum total of metered values during training. This gives LItLDF the flexibility to model apartment buildings with multiple units. The generalized form of our constraints employed allows for training despite uncertain meter-to-building mappings.

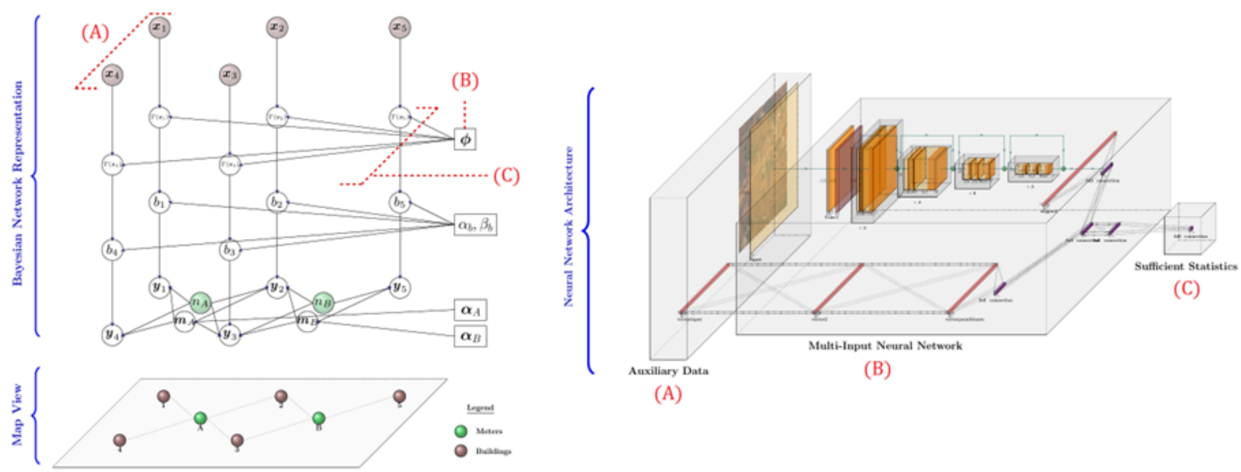


FIGURE 4: We developed physics-informed ML techniques specifically for training electricity demand models with imprecise meter location data. We employ a multi-input neural network (right) to extract relevant demand information from satellite imagery and remote sensing data while we develop a probabilistic graphical model (left) to account for potentially uncertain meter-to-building mappings. Figure from Lee, 2023.

A step change-improvement in capabilities for monitoring global energy poverty

We provide detailed electricity access and demand maps that reflect significant new capabilities and improvements for accurately monitoring global poverty. Our outputs allow tracking the MEM at scale.

- **Statistical tests provide confidence in our estimates.** Our building-level electricity access model compares favorably to baselines presented in the literature. In our tests focused in East Africa, our model achieves 80.7% accuracy while the next closest baseline achieves 70.9% (Lee, 2023). Our building-level electricity demand estimates also outperform baselines as depicted in Supplemental Figure 1.

- **Building-level estimates enable monitoring national progress towards the residential MEM.** We combine our building-level electricity access and demand estimates to infer consumption, calibrate to known consumption distributions, and apply household size assumptions to estimate country-level progress towards the MEM's residential consumption threshold. Initial residential MEM estimates as presented in Figure 3 and Supplemental Table 1. We discuss the specifics behind these assumptions in the supplement.
- **While benchmarks for MEM rates do not exist, we are still able to validate our distributions against reported national statistics.** We specifically mined dozens of annual utility reports to extract mean residential consumption figures for electrified households. By computing analogous metrics from our estimated electricity access and demand models, we are able to provide high-level validation for eight countries across Africa, as we depict in Figure 5, reflecting a mean average percentage error of 31%, weighted by the number of reported utility customers.[5]

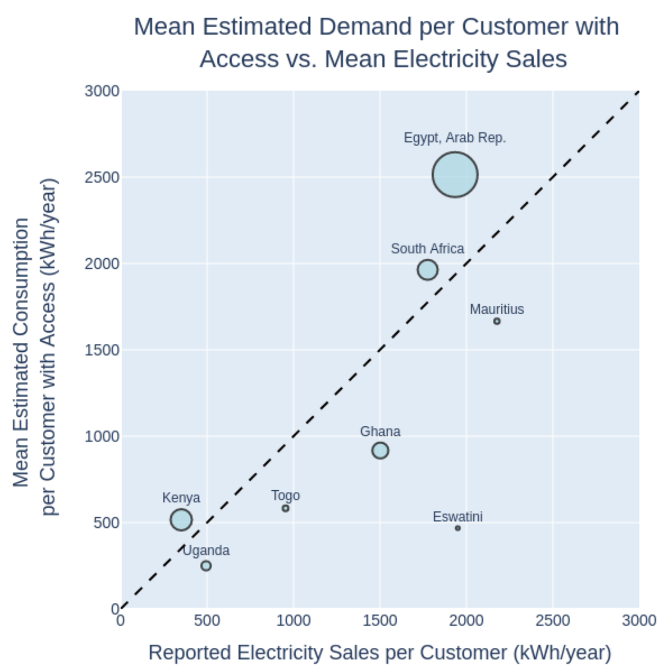


FIGURE 5: While benchmarks for Modern Energy Minimum rate estimates do not exist, we validate our estimated consumption distributions against mean residential consumption figures for electrified households from annual reports. We assess a 31% mean average percentage error for this comparable metric. The size of the circles represent the number of reported utility customers in national statistics.

[5] We extracted residential consumption customer counts and total electricity sales from public reports from the following organizations: [Kenya] Kenya Power & Lighting Company; [Uganda] UMEME Limited ; [Togo] Autorité de Réglementation du Secteur de l'Electricité ; [Ghana] Energy Commission of Ghana; [Eswatini] Eswatini Electricity Company; [Mauritius] Central Electricity Board, Government of Mauritius; [South Africa] Eskom Holdings SOC Ltd and City of Cape Town; and [Egypt] Ministry of Electricity and Energy.

- **We find our estimated MEM rates are comparable to national poverty rate statistics.** This is notable since these metrics are derived and reported fully independently of one another. We depict relationships between MEM rates and $3.65 and $6.85 per day poverty rates in Figure 6. Both plots reflect notable anticorrelation: with increasing estimated MEM rates, countries have lower shares of their population below both established income-based poverty lines.

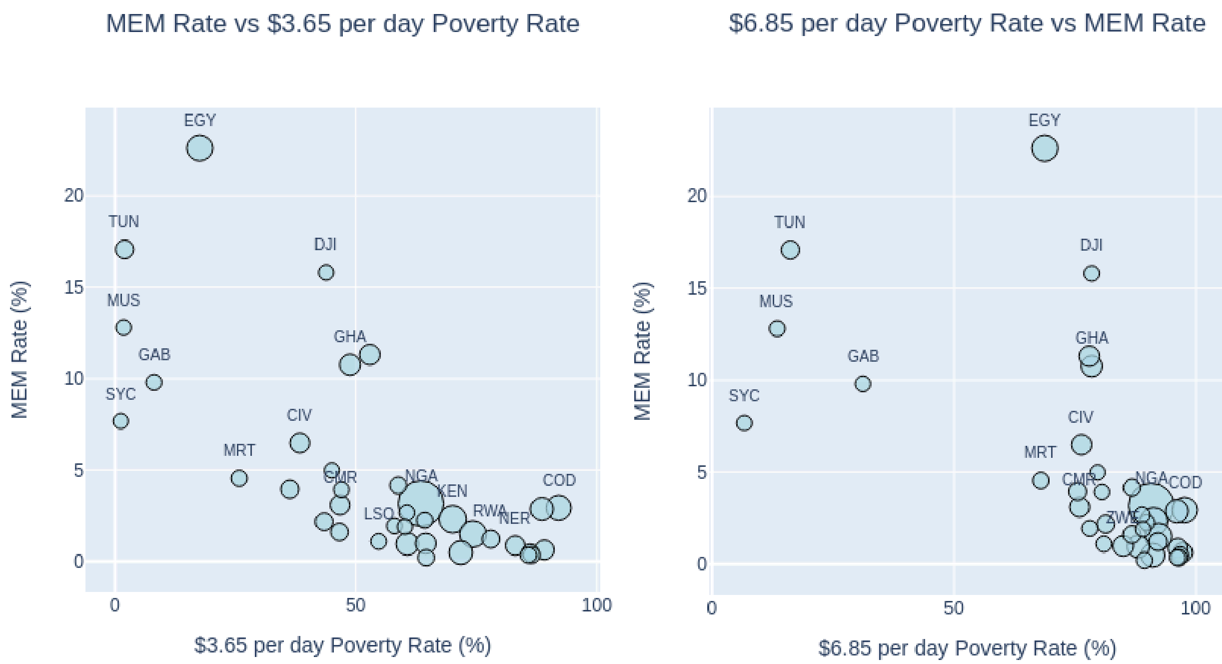


FIGURE 6: Our estimated MEM rates are comparable to reported national poverty rate statistics.

- **Our high-resolution estimates allow us to compare estimated consumption distributions across countries.** In Figure 7, we plot consumption distributions on a log-log scale for two African countries with relatively high MEM rates: Angola and South Africa, and two with relatively lower MEM rates: Benin and Kenya. This can be directly observed on the plot, as Angola and South Africa have higher shares of their population with consumption exceeding the MEM threshold.
  - Differences in consumption inequality can also be observed. Relative to South Africa, we estimate Angola to have a greater proportion of its consumers with very low consumption or no access to electricity. Because both Angola and South Africa are estimated to have otherwise similar shares of high-consumption consumers, it's apparent that Angola has greater energy consumption inequality than South Africa.
  - Relative to Benin, Kenya has greater shares of its population with both low very low consumption or no access to electricity, in addition to greater high-consumption shares. Benin conversely has greater shares of its population with intermediate consumption. We infer Kenya to have greater energy consumption inequality than Benin.

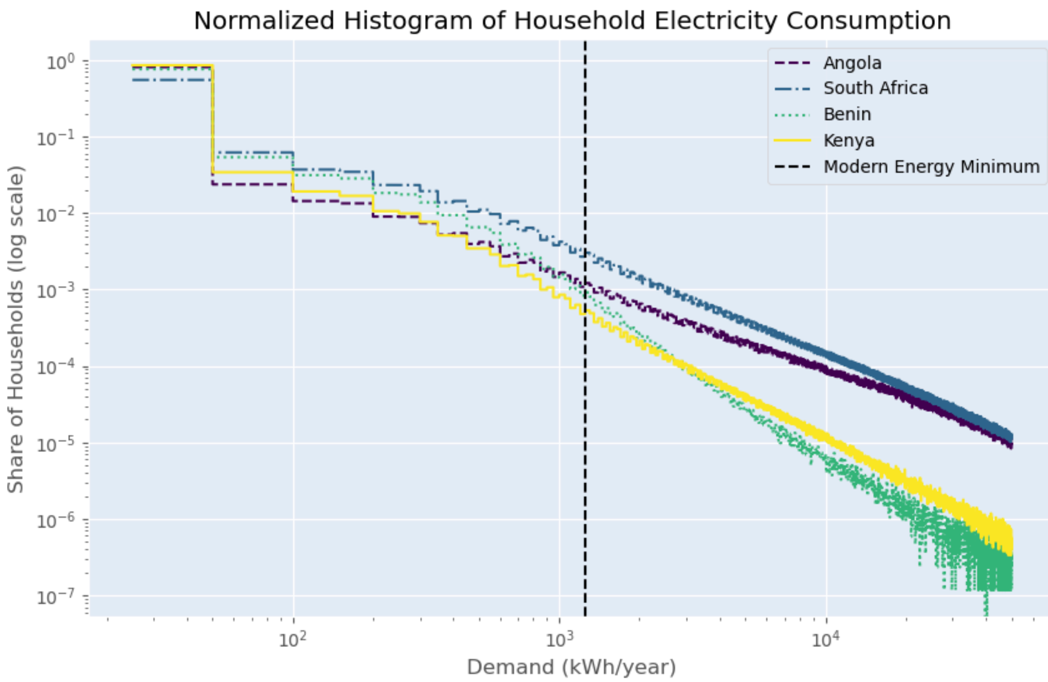


FIGURE 7: Normalized histograms of household electricity consumption for Angola, South Africa, Benin, and Kenya.

Future work

We've spent the last few years on R&D and we look forward to rolling out estimates and widely publicizing our results in 2025 and 2026. So far, we have preliminary across most African countries, and are tuning our methods before attempting to scale across all low- and middle-income countries globally. Our roadmap for the next two years includes the following steps:

- **Improve segmentation by residential and non-residential sectors.** Our initial analyses use coarse calibration and thresholding methods to distinguish between residential and non-residential buildings (see details in supplemental section S1). Future work on classifying residential and non-residential buildings will help us relax our assumptions and achieve improved reporting accuracy. We are refining an initial model we built to do this employing additional data sets including OpenStreetMap and Overture Maps.
- **Expand the scale of our analysis and calibrate our models using national consumption data.** We focused our initial work on select African countries but we are now looking to scale our analyses across all low-and-middle income countries in Africa, Asia, and the Americas. While we can run our models as-is, training data sparsity in non-African regions limits our ability to generalize. We are in the process of building a novel model calibration dataset by mining utility reports.
- **Roll out time-series estimates to infer historical access, demand, and consumption and apply time-series forecasting methods for future projections.** Most of the

geospatial data features we employ are now available historically, allowing us to run our model at scale and across time. We have not yet done so due to time and computational limitations, but have plans to roll out historical estimates for our next phase of work. Once we have these historical estimates, we plan to apply time-series forecasting methods to assess future access, demand, and consumption metrics. These model-derived forecasts will reflect a “status quo” scenario for policy-makers and decision-makers to use when planning both supply- and demand-side interventions.

Conclusion

Developing economies depend on electricity for growth, yet reliable data on where and how much electricity is consumed remains scarce. This gap hinders infrastructure investment, policy development, and decarbonization strategies, as current global metrics often reduce electricity access to simplistic, binary rates. Such measures overlook consumption levels and demand growth, factors essential for driving real economic and social impact. Without improved data, planners risk oversizing or undersizing infrastructure, wasting scarce investment dollars, missing out on benefits from economies of scale, and often increasing consumer costs and emissions. Our project addresses this gap by deploying advanced machine learning and remote sensing to create detailed, building-level electricity consumption datasets to be released in 2025-26.

Our novel approach provides governments, utilities, development banks, and donors with the first-ever public, data-driven system to track the Modern Energy Minimum that aligns with genuine economic growth needs. Using our data, planners can better allocate resources, while policymakers gain insights to design more impactful interventions. By integrating this system directly into the processes of organizations throughout the Energy for Growth Hub network, we aim to make MEM tracking a key part of global energy development strategies and ultimately accelerate global efforts to drive sustainable and inclusive development.

Supplement

Please find supplementary material below outlining key modeling assumptions and additional figures and tables referenced in the main text.

S1. Modeling Assumptions

The complexity of training building-level access and demand models, calibrating consumption values, and ultimately computing MEM rate metrics required us to make a number of modeling assumptions, outlined below.

- **Defining "electricity access" in our ML model for building-level access**. Our electricity access model employs a binary definition of electricity access; it is designed to treat grid connections and high-quality (grid-quality) off-grid connections as having "access" and no connection or low-quality off-grid connections as being "without access." The model is trained on grid connection data sourced directly from electricity utilities. Since off-grid connections are much less prevalent, their corresponding mislabeled buildings in training data can be thought to have minimal influence on the model. However, high-quality off-grid connections—such as those with bright nighttime lights or reliable internet access—may exhibit remote sensing characteristics similar to grid connections. Thus, the model may estimate higher probabilities of access for these higher-quality off-grid installations while estimating lower probabilities of access for low-quality off-grid connections.

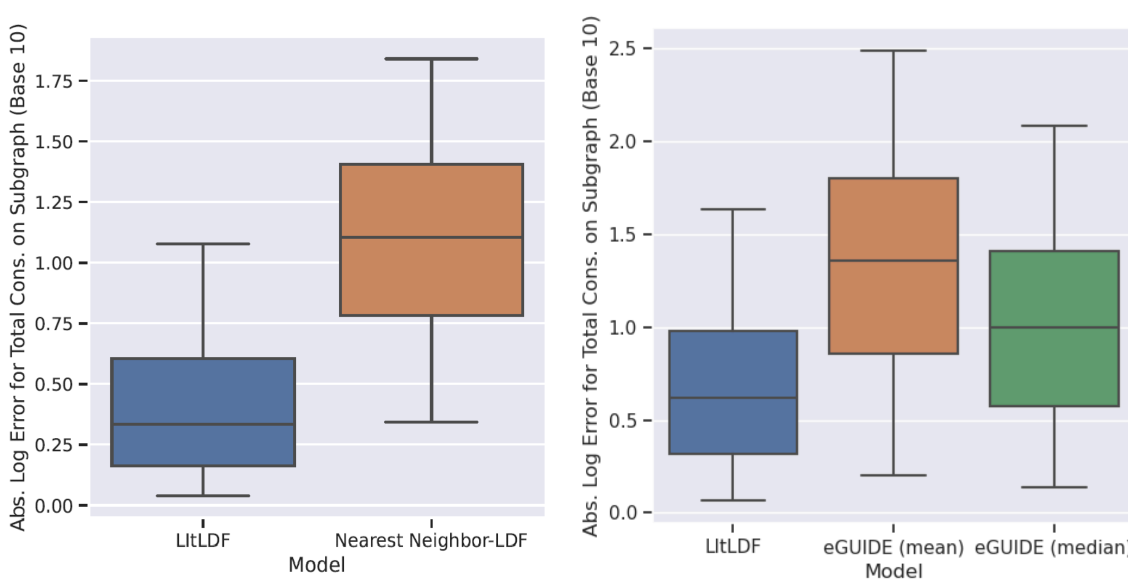
- **Calibrating for discrepancies between residential and non-residential consumers and accounting for non-electrified under-grid buildings when running LItLDF.** Our model uses linear scaling to align consumption estimates with known residential electricity consumption distributions from electrified consumers in Rwanda and Kenya. This calibration step mitigates biases introduced by the presence of non-residential consumers, under the assumption that residential vs. non-residential distributions in other African countries resemble those in these East African benchmark countries. Additionally, this process addresses inherent challenges that arise when distinguishing between customers and buildings. In the main text, we described the fact that a single meter may serve multiple buildings (e.g., detached apartments, etc.) and that multiple customers may reside within a building (e.g., apartment complexes, etc.), and that this precludes the use of traditional machine learning methods. LItLDF accounts for these complexities; however, ambiguity still remains for "under-grid" buildings lacking direct access or historical data. We assume that any residual bias from such cases can be offset through linear scaling. Finally, we apply thresholding to filter out any clear non-residential consumers from residential estimates, excluding any building with an estimated consumption exceeding 50 MWh/year.

- **Assuming household size for per-capita consumption estimates.** To convert residential building consumption to per-capita terms, as required by the Modern Energy Minimum, we assume an average of 5 residents per household. While we recognize that household size varies by country and level of urbanization, this estimate provides a consistent baseline for initial calculations. We anticipate refining this assumption in future iterations as we advance our methodology.

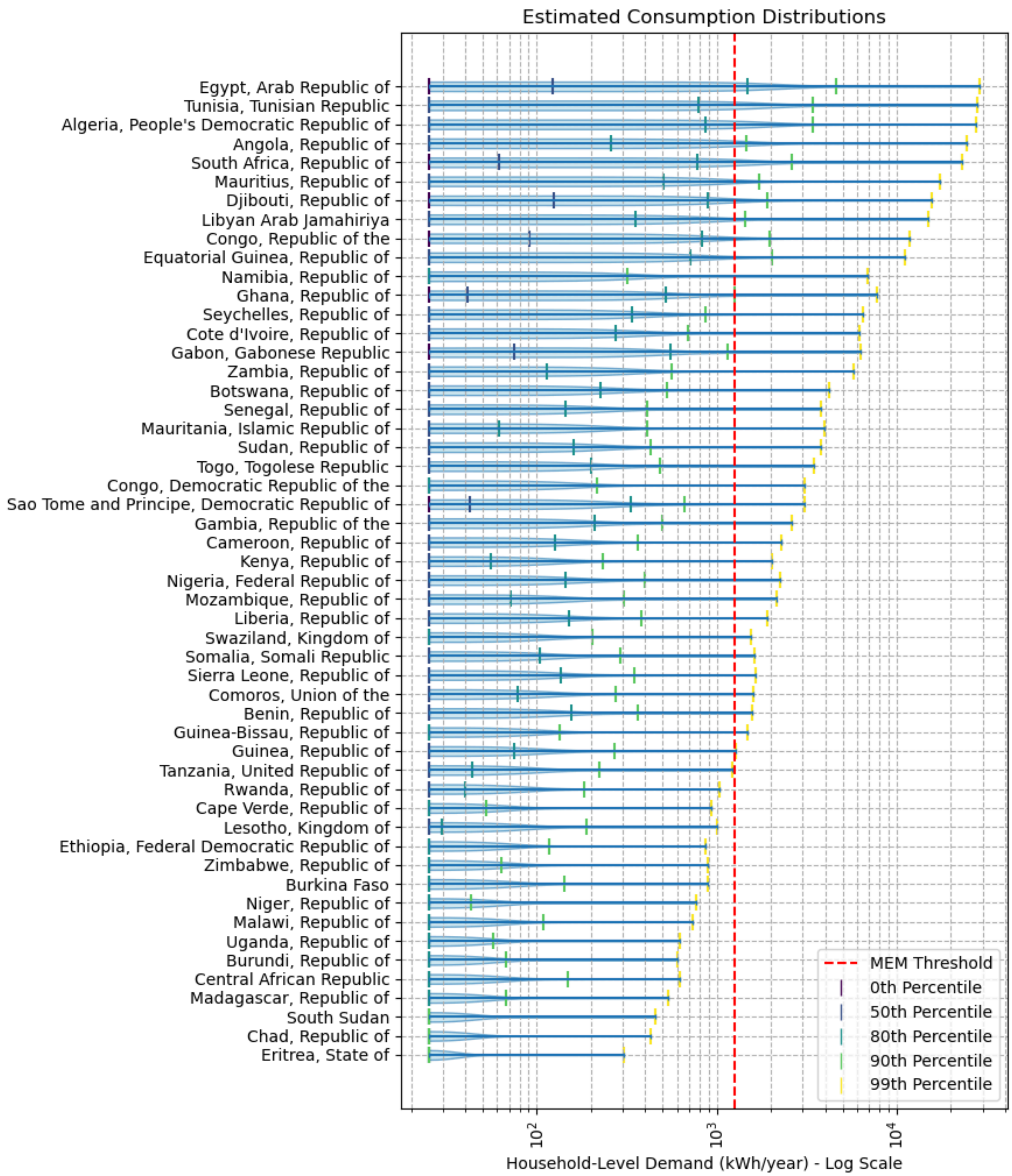
- **Excluding countries with incomplete building data from residential MEM estimates.** We were unable to generate residential Modern Energy Minimum (MEM) estimates for Mali and Morocco (including Western Sahara) due to significant gaps in building data. These missing data points prevent us from accurately estimating electricity access and consumption rates, which are essential for calculating MEM rates. Future improvements in building data coverage may allow us to include these countries in subsequent analyses.

- **Basing estimates on 2020 data with plans for future updates.** Our residential MEM estimates are centered around the year 2020, as both our remote sensing inputs and supporting datasets correspond to this period. While current estimates provide a valuable snapshot, future updates will incorporate more recent data and, ultimately, enable time-series modeling for access, demand, and consumption.

S2. Supplemental Figures



SUPPLEMENTAL FIGURE 1: We compare absolute log error metrics for the LItLDF model (our model) with baseline models when trained and validated in parts of Rwanda and tested in held-out parts of Rwanda (left plot), and when when trained and validated in parts of Rwanda and tested in held-out parts of Kenya (right plot). Our method achieves favorable performance relative to baselines.



SUPPLEMENTAL FIGURE 2: Estimated consumption distributions with 50th, 80th, 90th, and 99th percentiles visualized for household-level demand presented as violin plots by country. Households with no access are grouped with the minimum value in the range for visualization purposes.

SUPPLEMENTAL TABLE 1: Estimated MEM Rates by country

Country	Estimated Residential MEM Rate (%)	Country	Estimated Residential MEM Rate (%)
Egypt, Arab Republic of	23	Sudan, Republic of	4
Algeria, People's Democratic Republic of	18	Togo, Togolese Republic	4
Tunisia, Tunisian Republic	17	Senegal, Republic of	4
South Africa, Republic of	16	Gambia, Republic of the	4
Djibouti, Republic of	16	Nigeria, Federal Republic of	3
Congo, Republic of the	16	Cameroon, Republic of	3
Equatorial Guinea, Republic of	15	Congo, Democratic Republic of the	3
Mauritius, Republic of	13	Mozambique, Republic of	3
Libyan Arab Jamahiriya	12	Liberia, Republic of	3
Angola, Republic of	11	Kenya, Republic of	2
Ghana, Republic of	11	Sierra Leone, Republic of	2
Gabon, Gabonese Republic	10	Benin, Republic of	2
Seychelles, Republic of	8	Somalia, Somali Republic	2
Cote d'Ivoire, Republic of	6	Comoros, Union of the	2
Zambia, Republic of	6	Swaziland, Kingdom of	2
Namibia, Republic of	5	Guinea-Bissau, Republic of	2
Sao Tome and Principe, Democratic Republic of	5	Guinea, Republic of	2
Botswana, Republic of	5	Tanzania, United Republic of	1
Mauritania, Islamic Republic of	5	Rwanda, Republic of	1

SUPPLEMENTAL TABLE 1-b: Estimated MEM Rates by country

Country	Estimated Residential MEM Rate (%)
Cape Verde, Republic of	1
Lesotho, Kingdom of	1
Ethiopia, Federal Democratic Republic of	1
Zimbabwe, Republic of	1
Burkina Faso	1
Niger, Republic of	1
Malawi, Republic of	1
Burundi, Republic of	< 0.5
Uganda, Republic of	< 0.5
South Sudan	< 0.5
Central African Republic	< 0.5
Madagascar, Republic of	< 0.5
Chad, Republic of	< 0.5
Eritrea, State of	< 0.5