August 2020

Working Paper

009.2020

M-LED: Multi-sectoral Latent Electricity Demand Assessment for Energy Access Planning

Giacomo Falchetta, Nicolò Stevanato, Magda Moner-Girona, Davide Mazzoni, Emanuela Colombo and Manfred Hafner

Future Energy Program
Series Editor: Manfred Hafner

M-LED: Multi-sectoral Latent Electricity Demand Assessment for Energy Access Planning

By Giacomo Falchetta, Fondazione Eni Enrico Mattei (FEEM) and Cattolica University
Nicolò Stevanato, Fondazione Eni Enrico Mattei (FEEM) and Politecnico di Milano, Department of Energy
Magda Moner-Girona, European Commission, Joint Research Centre (JRC)
Davide Mazzoni, Fondazione Eni Enrico Mattei (FEEM)
Emanuela Colombo, Politecnico di Milano, Department of Energy
Manfred Hafner, Fondazione Eni Enrico Mattei (FEEM), John Hopkins University SAIS and Sciences Po PSIA

Summary

Globally about 800 million people live without electricity at home, over two thirds of which are in sub-Saharan Africa. Ending energy poverty is a key development priority because energy plays an enabling role for human wellbeing and economic activities. Planning electricity access infrastructure and allocating resources efficiently requires a careful assessment of the diverse energy needs across space, time, and sectors. However, because of data scarcity, most country or regional-scale electrification planning studies have been based on top-down electricity demand targets. Yet, poorly representing the heterogeneity in the electricity demand can lead to inappropriate energy planning, inaccurate energy system sizing, and misleading cost assessments. Here we introduce M-LED, Multi-sectoral Latent Electricity Demand, a geospatial data processing platform to estimate electricity demand in communities that live in energy poverty. The key novelties of the platform are the multi-sectoral, bottom-up, time-explicit demand evaluation and the assessment of water-energy-agriculture-development interlinkages. We apply the methodology to the country-study of Kenya. Our findings suggest that a bottom-up approach to evaluating energy needs across space, time, and sectors is likely to improve the reliability and accuracy of supply-side electrification modelling and therefore of electrification planning and policy.

Keywords: Electricity Access, Energy Demand, Rural Development, Bottom-up Modelling, Sub-Saharan Africa, Multi-sectoral Approach, Water-Energy-Food-Environment Nexus

JEL Classification: Q4, Q41, O13

Address for correspondence:
Giacomo Falchetta
Fondazione Eni Enrico Mattei
C.so Magenta, 63
20123 Milan
Italy
E-mail: giacomo.falchetta@feem.it

M-LED: multi-sectoral latent electricity demand assessment for energy access planning

Giacomo **Falchetta**[1,2*], Nicolò **Stevanato**[1,3], Magda **Moner-Girona**[4], Davide **Mazzoni**[1], Emanuela **Colombo**[3] and Manfred **Hafner**[1,5,6]

1. Fondazione Eni Enrico Mattei (FEEM), Corso Magenta 63, 20123, Milan, Italy
2. Cattolica University, Lgo. Gemelli 1, 20123, Milan, Italy
3. Politecnico di Milano, Department of Energy, Via Lambruschini, 4, 20156 Milan, Italy
4. European Commission, Joint Research Centre (JRC), Ispra, Italy
5. John Hopkins University SAIS, Via Beniamino Andreatta, 3, 40126 Bologna, Italy
6. Sciences Po PSIA, 27 Rue Saint-Guillaume, 75007 Paris, France

*Corresponding author, email address: giacomo.falchetta@feem.it

Abstract

Globally about 800 million people live without electricity at home, over two thirds of which are in sub-Saharan Africa. Ending energy poverty is a key development priority because energy plays an enabling role for human wellbeing and economic activities. Planning electricity access infrastructure and allocating resources efficiently requires a careful assessment of the diverse energy needs across space, time, and sectors. However, because of data scarcity, most country or regional-scale electrification planning studies have been based on top-down electricity demand targets. Yet, poorly representing the heterogeneity in the electricity demand can lead to inappropriate energy planning, inaccurate energy system sizing, and misleading cost assessments. Here we introduce M-LED, *Multi-sectoral Latent Electricity Demand,* a geospatial data processing platform to estimate electricity demand in communities that live in energy poverty. The key novelties of the platform are the multi-sectoral, bottom-up, time-explicit demand evaluation and the assessment of water-energy-agriculture-development interlinkages. We apply the methodology to the country-study of Kenya. Our findings suggest that a bottom-up approach to evaluating energy needs across space, time, and sectors is likely to improve the reliability and accuracy of supply-side electrification modelling and therefore of electrification planning and policy.

1. Introduction

Electricity is a direct input to virtually every economic sector. An abundant, affordable, and reliable provision of power is a necessary condition for human livelihoods to thrive. This involves the achievement of nearly all the UN Sustainable Development Goals (SDGs) (McCollum et al., 2018; Nerini et al., 2018). Recent statistics on electricity access show that globally just under 800 million people (about 10% of the global population) live without access to electricity, more than two-thirds of which are in sub-Saharan Africa (IEA et al., 2020). Even in areas reached by electricity infrastructure, a large latent demand often persists (Fabini et al., 2014; Falchetta et al., 2020; Poblete-Cazenave and Pachauri, 2019).

In the context of energy planning to eliminate energy poverty, the assessment of the long-run electricity demand plays a crucial role (Leea et al., 2019). Namely, the choice of the most efficient electricity supply option and the size of the local generation capacity and storage system strongly depend on the assumed local demand. In turn, this demand is defined both by the hourly load curve and its peaks, and by the total energy consumption. The link between the target demand and electricity supply planning becomes very evident when carrying out country or regional scale studies with Geospatial Electrification Models (GEMs). GEMs are data-intensive computer-based tools that can support policymakers in the integrated evaluation of the most suitable and cost-effective technology for providing electricity access to all settlements (Adkins et al., 2017; Cardona and López, 2018; Kemausuor et al., 2014; Korkovelos et al., 2019; Mentis et al., 2017; Moner-Girona et al., 2019, 2016; Morrissey, 2019; Ohiare, 2015; Parshall et al., 2009; Sanoh et al., 2012; Szabo et al., 2011; van Ruijven et al., 2012). Thanks to growing data collection and management facilities, bottom-up techno-economic electrification analysis has become widely available (refer to the *Global Electrification Platform* and the *WRI's Energy Access Explorer*). Differently from approaches based on linear programming, GEMs do not aim at locally optimising energy systems for specific communities. Their main characteristic is that they allow to identify – country or region-wide – the optimal set-up (i.e.

the technology with the lowest local levelized cost of electricity) for providing electricity access at each settlement, along with the generation capacity and investment requirements. The cost-optimal set-up depends on the local energy resources and existing infrastructure.

Yet, most of the GEM-based literature has been strongly supply-side oriented (Morrissey, 2019). Studies have focused mainly on residential energy services when defining the demand of settlements lacking electricity access, and have so far exhibited limited capacity of accounting for the electricity demand from services and productive uses driven by the presence of farms, small businesses, commercial activities, healthcare facilities, and schools. In these studies, the residential demand itself has mostly been calibrated with regional average residential electricity consumption levels of urban and rural consumers (Mentis et al., 2017; Szabo et al., 2011; Szabó et al., 2016; van Ruijven et al., 2012), with little within-country heterogeneity. Archetypical demand targets include – for instance – values for sub-Saharan Arica from the World Bank Multi-Tier Framework (Bhatia and Angelou, 2015) or specific per-capita consumption levels defined by decision makers under a medium-run time horizon (usually 2030, the Sustainable Development Goals target year).

Many studies exploiting GEMs based on such top-down characterisation of the demand have concluded that decentralised energy solutions will play a prominent role in guaranteeing that SDG 7.1.1 (the universal electricity access target) is met. For instance, the Africa Energy Outlook 2019 (IEA, 2019) argues that mini-grids and stand-alone systems will serve 30% and 25% of those gaining access, respectively. This means that for more than half of the households, the electricity access problem could be solved thanks to decentralised energy technologies. Yet, care is required in the interpretation of these results. The number and size of non-residential consumers in a community can have a crucial effect (Peters et al., 2019) on the total long-term energy demand, the peak loads, and consequently a direct effect on the optimal energy technology mix (diesel generator, PV, wind, biomass, hydro or hybrid technologies), on the optimal technology set-up (i.e. the choice between grid extension, mini-grid, or standalone solutions) and on the overall cost-benefit analysis of electrification (Brüderle et al., 2011; Morrissey, 2018; World Resources Institute, 2020). An inadequate or generic formulation of the demand might lead to inefficient allocation of budget and sizing of electricity infrastructure (Riva et al., 2019). Moreover, enabling services for the community and productive uses of electricity beyond basic household needs – such as energy use in agriculture, small businesses, and healthcare and education facilities – is crucial to unleash local economic development (Riva et al., 2018). While substantial uncertainty persists over the structural welfare impacts of household electrification programs (Urpelainen, n.d.), there is robust evidence of the positive effect of electricity provision on time spent by household members in income-generating activities (Bernard, 2010; Bos et al., 2018; Rathi and Vermaak, 2017; Van de Walle et al., 2013). In turn – provided a set of conditions is satisfied – the electricity input might improve the income of the whole community (Peters and Sievert, 2016).

Different approaches have been introduced so far to tackle these limitations. For instance, the adoption of more detailed and heterogeneous household consumption profiles (Trotter et al., 2017), the use of system dynamics (Riva et al., 2019), or the life-cycle assessment of embodied energy in goods and services that contribute to providing what is defined *decent living energy* (Rao et al., 2019). Yet, only few (Moner-Girona et al., 2019, 2016; Narayan et al., 2018; Zhang and Zhang, 2019) of the existing GEM-based studies have so far taken into consideration field-validated load profiles or accounted for the existence of services and productive activities to estimate local energy demand requirements. To tackle these challenges, it has been argued that planning tools need to be improved, and evidence-based projections of electricity consumption need to be used (Blodgett et al., 2017; Moner-Girona et al., 2018). The main roadblock to deliver a standard methodology to estimate electricity demand stems from the data-intensiveness of the estimation, the uncertainty over the quality of the existing data and about the different scenarios that forecast the energy demand growth over time, as well as the computational challenge to produce a high-resolution output.

To advance the state-of-the art in the characterisation of the demand for electricity and ensuring that insights drawn from GEMs are suitable to empower communities in the context of electrification planning, here we introduce the open-source **M**ulti-sectoral **L**atent **E**lectricity **D**emand (**M-LED**) geospatial data processing and assessment platform. **M-LED** enables an estimation of electricity demand in communities that live in energy poverty. The key novelty of the platform is its multi-sectoral, bottom-up, high spatio-temporal resolution evaluation, which altogether advances the state-of-the-art on latent electricity demand characterisation. Here, by latent demand, we refer to demand which would exist if the infrastructure and techno-economic conditions to supply it would be met. Secondly, besides modelling different non-residential sectors including the agriculture, service, and productive activities, the platform includes a more detailed assessment of residential demand – representing heterogeneous appliances ownership and usage patterns and allowing for stochastic variability in the demand. Thirdly, the M-LED platform enables a characterisation of the seasonal and hourly variation in the demand from different sectors is of crucial importance for properly planning the energy system and assessing the complementarity of variable renewable energy sources supply curves with the demand. Finally, the multi-sectoral approach includes an assessment on the water-energy needs and the nexus implications for agriculture-related activities. This encompasses an analysis of the potential revenues and costs from the potential agricultural productivity growth thanks to artificial irrigation as a result of the provision of electricity.

The remainder of the paper is structured as follows. Section 2 introduces the methodology of the M-LED platform to carry out the multi-sectoral, bottom-up, high spatio-temporal resolution electricity demand evaluation. Section 3 presents an application of the platform for country-study of Kenya. Section 4 discusses the relevance of the results both from an energy modelling and a policy perspective, and it and highlights potential future applications of the M-LED platform.

2. Materials and methods

2.1. The Multisectoral Latent Electricity Demand assessment platform

An integrated electrification plan must identify and target population catchment areas (in this study defined as *clusters*; refer to **SI-A1**) and the different local electricity consumption drivers. These include residential demand, productive activities, and several service-provisioning facilities. The M-LED platform is an open-source, bottom-up platform designed to characterise power requirements across different sectors. The platform combines openly available geospatial information, modelling instruments, and scenario analysis to support a sectoral-inclusive electrification planning (see the **SI** for a detailed description of the underlying Materials and Methods). The input data sources are openly accessible and are reported in **Table SI1**. The data processing procedure collates field and remotely sensed observations. The lack of data or uncertainty over future evolutions over certain sectors is tackled with explorative modelling. **Figure 1** offers an overview of the workflow. The methodology is based on an array of Python-based open-source GIS algorithms (from Quantum-GIS, GRASS-GIS, SAGA-GIS, and GDAL) complemented by intermediate R scripting.

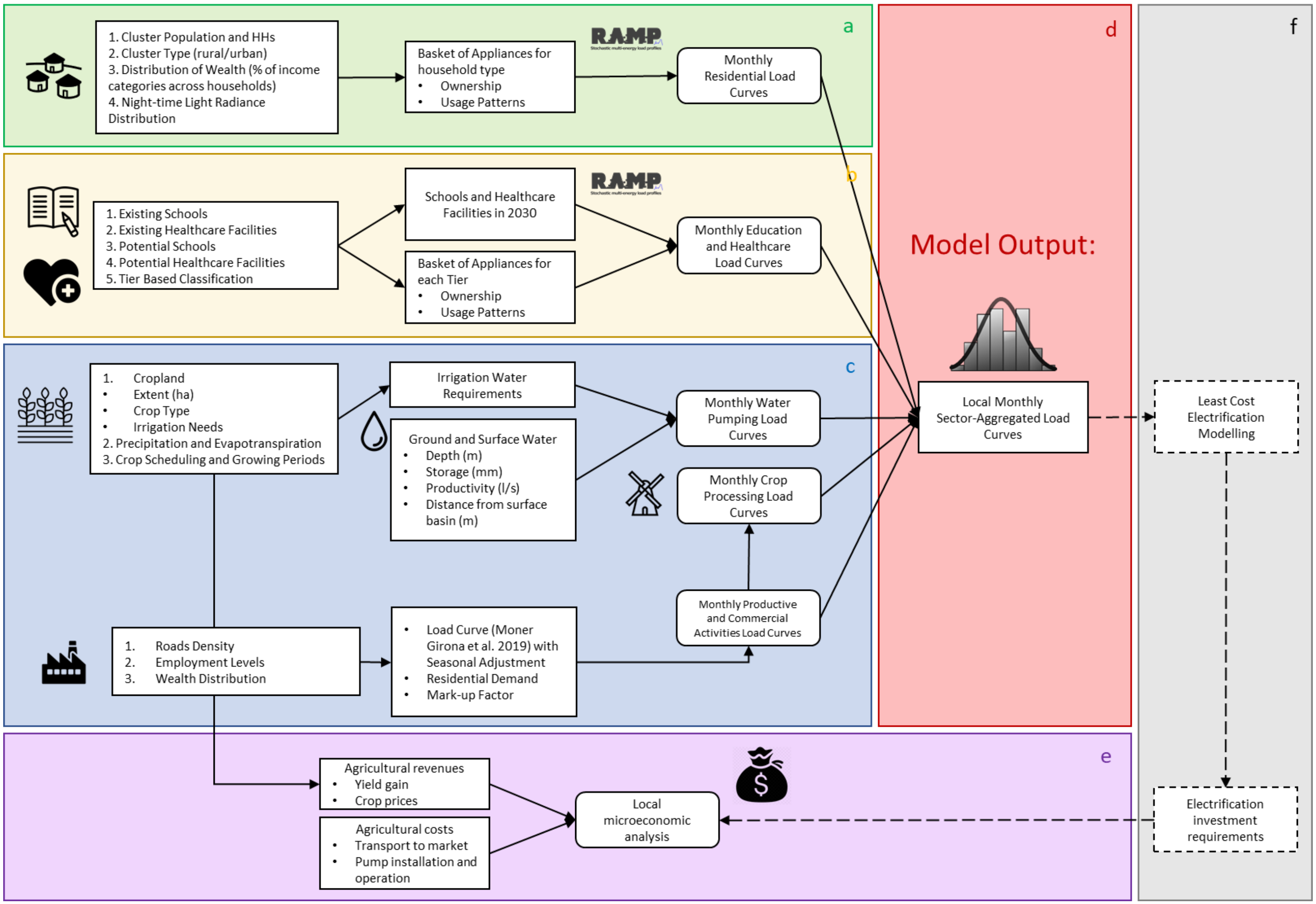
a
1. Cluster Population and HHs
2. Cluster Type (rural/urban)
3. Distribution of Wealth (% of income categories across households)
4. Night-time Light Radiance Distribution
Basket of Appliances for household type
• Ownership
• Usage Patterns
RAMP
Monthly Residential Load Curves
b
1. Existing Schools
2. Existing Healthcare Facilities
3. Potential Schools
4. Potential Healthcare Facilities
5. Tier Based Classification
Schools and Healthcare Facilities in 2030
Basket of Appliances for each Tier
• Ownership
• Usage Patterns
RAMP
Monthly Education and Healthcare Load Curves
c
1. Cropland
• Extent (ha)
• Crop Type
• Irrigation Needs
2. Precipitation and Evapotranspiration
3. Crop Scheduling and Growing Periods
Irrigation Water Requirements
Ground and Surface Water
• Depth (m)
• Storage (mm)
• Productivity (l/s)
• Distance from surface basin (m)
Monthly Water Pumping Load Curves
Monthly Crop Processing Load Curves
1. Roads Density
2. Employment Levels
3. Wealth Distribution
• Load Curve (Moner Girona et al. 2019) with Seasonal Adjustment
• Residential Demand
• Mark-up Factor
Monthly Productive and Commercial Activities Load Curves
d
Model Output:
Local Monthly Sector-Aggregated Load Curves
f
Least Cost Electrification Modelling
Electrification investment requirements
e
Agricultural revenues
• Yield gain
• Crop prices
Agricultural costs
• Transport to market
• Pump installation and operation
Local microeconomic analysis

Figure 1: Conceptual framework of the M-LED platform. (a) Residential demand: *clustering and appliance baskets generation and parsing;* ***(b) Healthcare and educational demand:*** *geolocation and classification of facilities and appliance baskets generation and parsing****; (c) Agricultural uses (irrigation and crop processing) loads*** *(see Figure 3 for further details),* ***micro enterprises and commercial demand:*** *drivers of energy demand assessment and sectoral demand calculation;* ***(d) Model Output:*** *Cluster and Sector -specific yearly Load Curves with month level seasonality and 1-hour resolution.* ***(e) Assessment of costs and revenues of increased agricultural productivity****.* ***(f) The produced output is intended to be fed in geospatial electrification models for more effective energy planning:*** *this is not included in the present work.*

The platform exploits the RAMP (*Remote-Areas Multi-energy systems load Profiles*) model (Lombardi et al., 2019), which supports the creation of stochastic, seasonal-heterogeneous energy demand profiles. The underlying stochastic process lies in the structure of the bottom-up model adopted for load profile generation **(Figure 2)**. The structure consists of three different layers of modelling: the *User Type*, the *User* and the *Appliance* layer. "*The first layer consists in the definition of a set of arbitrary User types (e.g. Household, Commercial activities, Public offices, Hospitals, etc.). Each User Type is subsequently characterised in terms of the number of individual Users associated to that category (second layer) and in terms of Appliances owned by each of those Users (third layer). The three-layer structure allows to independently model the behaviour of each jik-th Appliance, so that each individual ji-th User within a given i-th User Type will have a unique an independent load profile compared to the other Users of the same Type. The aggregation of all independent User profiles ultimately results in a total load profile, which is uniquely generated at each model run. Multiple model runs generate different total load profiles, reproducing the inherent randomness and unpredictability of users' behaviour and allowing to obtain a series of different daily profiles*" (quoted from Lombardi et al., 2019).

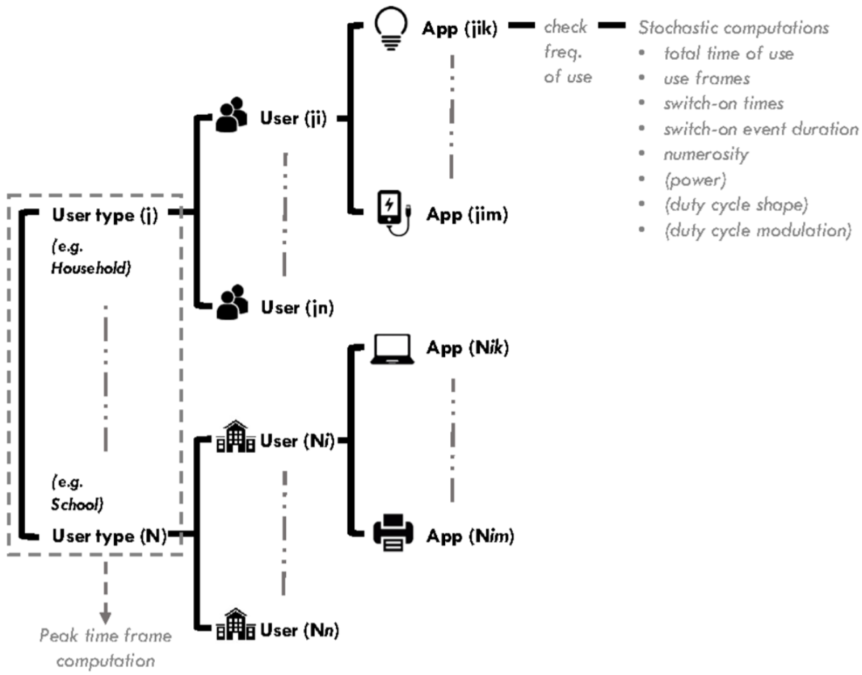

Figure 2: Schematic framework of the RAMP stochastic load generation model. Source: Lombardi et al. (2019).

The M-LED platform generates instantaneous electricity demand load curves (rendered at a one hour time step) and then derives the monthly (seasonal-varying) and yearly-aggregated consumption levels. The outputs consist of georeferenced layers for the estimated latent (i.e. currently unsupplied) electricity demand within population clusters from a set of residential, productive activities, and services. Residential, health, and education load profiles are computed following a probabilistic distribution starting from field campaign or literature-validated appliance ownership and use patterns under an array of scenarios (See SI). Agricultural (irrigation and crop processing) and micro enterprises loads are assessed combining techno-economic modelling and literature estimates.

An applicative example for Kenya is provided (with accurate country-specific data and a comprehensive assessment on the water-energy needs for agricultural activities). The key added value of the M-LED methodology is that its results will allow carrying out supply-side planning of energy access systems according not only to the energy resource availability but also to the local specific community and productive load profiles, including daily, weekly, and seasonal variation, which can significantly affect system design (Huld et al., 2017). On top of that, the M-LED geospatial analysis allows to identify agricultural productivity growth hotspots where investment can be prioritized to leverage the strongest welfare impact. For instance, the platform estimates the increase in the revenues from the potential boost in the per-hectare yield due to artificial irrigation, which in turn might compensate for the low ability-to-pay for energy services of rural dwellers (Blimpo and Cosgrove-Davies, 2019).

The code and data for the M-LED platform are hosted at the public repository https://github.com/giacfalk/M-LED, which also includes a maintained documentation at https://github.com/giacfalk/M-LED/wiki.

2.2. Residential, services, and micro-enterprise loads

Residential electrification plays a crucial role for human wellbeing, for instance by enabling telecommunications, conserving food fresh, indoor air circulation and cooling, and night-time activities. In fact, most electrification efforts and targets, including SDG 7.1.1, focus on bringing electricity to all households. Yet, also a large number of healthcare and education facilities face significant constraints in their activity because they are unable to operate appliances that are crucial for guaranteeing the wellbeing and development prospects of local population (Adair-Rohani et al., 2013; Sovacool and Vera, 2014). Finally, the provision of electricity can foster small entrepreneurial activities such as small shops, mini-markets, handcraft and telecommunication services retail (Bose et al., 2013; Kariuki, 2016; Manggat et al., 2018) which can represent a significant leverage for broader socio-economic development (Kongolo, 2010).

With regards to residential electrification, to tailor infrastructure efficiently it is necessary to distinguish among different household types. A relevant example is the introduction of ESMAP's *Multi-Tier Framework for Measuring Energy Access* (Bhatia and Angelou, 2015). To estimate household demand in a flexible way, the M-LED framework is designed to ensure a large degree of heterogeneity in residential power demand. We construct 5x2 = 10 archetypical types (five in urban areas, and five in rural settlements) of households by electrical appliance ownership and use patterns. These are designed starting from a systematic screening of the literature (Adeoye and Spataru, 2019; Blodgett et al., 2017; Kotikot et al., 2018; Lee et al., 2016b; Monyei et al., 2019; Monyei and Adewumi, 2017; Sprei, 2002; Thom, 2000) about electricity consumption in developing countries and parametrised based on data from recent field visits in Kenya by the authors and their team (2019). The empirical screening provides the rationale to compile tables of appliances and usage patterns (refer to **SI-A2**) for each household type. A total number of 22 appliances is selected and modelled across 11 dimensions (ownership, number of appliances per user, appliance power, number of daily functioning windows, windows start and end times, percentage of variability of windows start and end times, daily functioning time, percentual of random variability of daily functioning time, minimum time the appliance is kept on after switch-on event, percentage of occasional use, weekend or weekday use). These dimensions are summarised in Table 1. In order to account for seasonality of the load in the residential sector, the climate variable is taken into account and the months of January and December are considered the hottest in the country, while June and July the cooler. The appliances related with thermohygrometric well-being inside the households, namely fans and air conditioning systems, are modelled according to this climatic variability. In detail June and July are assumed to have no use of such appliances, and the other months gradually increase their use up to a full use in the months of January and December. Given the proximity to the equator of the country, dusk and dawn times are considered to not vary significantly enough to justify seasonal variation of time of use of appliances and lights. The entire set of modelled appliances, users and user types with relative parameters are reported in **Supplementary File F1**.

Table 1: Dimensions considered in the stochastic demand assessment

Dimension	Description	Range
Ownership	Category of User that owns the appliance	User Type
Number of appliances per user	Number of that specific appliance owned by the user	Non-negative [-]
Appliance power	Nominal power of the specific appliance, allows for a random variability in a defined range for thermal appliances	Non-negative [W]
Number of daily functioning windows	Number of time "windows" in which the appliance is used during the day	1-3 [-]
Window start and end times	Hours of start and end of time windows in which the appliance can be used	00:00 – 23:59
% variability of window start and end times	percentage of allowed random variation of the length of the usage windows	0-100 [%]
Daily functioning time	total amount of time that the appliance is used during one day	0-1440 [min]
% of random variability of daily functioning time	percentage of allowed random variation of the total daily time of use	0-100 [%]
minimum time the appliance is kept on after switch-on event	minimum amount of time the appliance stays on after has been switched on	0-1440 [min]
percentage of occasional use	probability that the appliance is used on a single day	0-100 [%]
Weekends or weekdays use	allows to constrain the usage of the appliance only in weekdays or in weekends periods	we / wd / none

Thereafter, the RAMP stochastic demand model (see **Figure 2**) is used to simulate for each of the ten household classes a representative community of *n*=100 households (to ensure sufficient stochasticity). For each cluster i, The RAMP model generates 12 month-specific load curves (in W), at a minute time-step for 365 days, from which it is easy to calculate the total residential power consumption (in kWh).

To parse the simulated energy demand profiles with each population cluster (see **SI-A3**), we firstly evaluate the statistical association between the distribution of the population with electricity access across electricity access tiers (based on validated, satellite-derived data on the prevalent tier of electricity access at each pixel (Falchetta et al., 2019) and with reference to the World Bank Multi-Tier Framework for measuring energy access (Bhatia and Angelou, 2015)) and the type of settlement (urban or rural prevalence (A.J. et al., 2019)), the local population density and the distribution of wealth within of sub-Saharan African countries (based on household survey data from the USAID DHS StatsCompiler (USAID, 2009)). Then, based on the regression coefficients we allocate each household without access to electricity enclosed in each cluster to each of the RAMP-generated demand profile archetypes. The process therefore assumes that the distribution among electricity access tiers of those who already today benefit from electric services at home in each cluster will also apply to households that will gain electricity access in the future.

The service infrastructure energy demand is modelled in a similar fashion to the residential assessment, we design baskets of appliances ownership and use for tiers of each category of facility (detailed in **Supplementary File F2**). Scientific (Giday, 2014; Olatomiwa et al., 2018) and gray (Action, 2013) literature on the theme exists, but is often generic and usually scarce when it comes to sub-Saharan Africa. Thanks to a field campaign conducted under the supervision of the authors in primary schools and rural healthcare facilities of Kenya and based on a survey and empirical observation of the appliance ownership and use, energy consumption, and pupils or hospital beds hosted, we are able to reconstruct the field energy demand data in the RAMP model and allocate it to the (latent) demand of clusters where similar facilities are located. Information on operational healthcare facilities is based on open-data on the location and characteristics of public[1] healthcare facilities (Maina et al., 2019). Similarly, open-data for the position and size of schools is retrieved ("Kenya Open Data Initiative - Humanitarian Data Exchange," n.d.). We classify healthcare facilities into five tiers following the criteria presented in the **SI A-4** and the facility type explicated in the original dataset (Maina et al., 2019). Once information about the location and typology of healthcare and education facilities is compiled, we calculate the density of facilities of each tier in each cluster. Based on this information, we estimate the total local sectoral demand exploiting the 1-minute resolution, tier-heterogeneous, monthly-seasonal demand profiles calculated in RAMP. The seasonality of school facilities is indeed dependent on the national school calendar[2], and has been modelled accordingly.

Approximating the residual productive demand from microenterprises (in the context of developing countries defined as small businesses employing few, generally household-related, people and with a limited turnover) is challenging task because of the lack of granular country or region-wide data, which makes it impossible to model at an appliance, plant, or facility level. Proxy estimation approaches have been introduced (Moner-Girona et al., 2019; Parshall et al., 2009). Here (see **SI-A5**) we propose a model based on employment, infrastructure proximity, and wealth to create a bottom-up

[1] To date there is no comprehensive publicly available dataset of private healthcare facilities in sub-Saharan Africa.
[2] https://publicholidays.co.ke/school-holidays/2020-dates/

residential demand multiplier factor ranging between +30% and + 60% (Moner-Girona et al., 2019). In particular, we carry out a PCA (principal component analysis) to create a composite index based on relevant drivers of productive activities (such as road density, accessibility, employment levels and wealth distribution). The composite index is used to define the local residential demand multiplier factor, which is used to derive the yearly productive demand on top of the residential demand. The baseline load curve (share of demand at each hour of the day over the total daily demand) of micro productive activities is assumed to follow the same path of that described in Moner-Girona et al. (2019) for Kenya, which in turn is derived on real metered data. A seasonal variation is imposed on the baseline load curve, so that each monthly curve follows the same monthly relative mark-up observed in the residential demand.

2.3. Load curves for agricultural productive uses: the relevance of the WEF nexus (Water, Energy, Food security)

Currently in sub-Saharan Africa more than 90% of total cropland is rainfed (Xiong et al., 2017), with the figure standing at about 95% in Kenya (The World Bank, 2019). Together with the lack of fertilisation, the unmet irrigation water demand implies a situation of sub-optimal production, in what has been defined the *yield gap* (GYGA, 2017; Mueller et al., 2012). Moreover, the bulk of the production is either for subsistence purposes or is sold to wholesale markets unprocessed. This is because of the lack of crop processing facilities in most small and medium farm businesses (Sims and Kienzle, 2017, 2016), also because of the lack of energy supply to power those plants, as well as due to market accessibility issues. Most farms thus sell their production to few large processing plants or supply it directly to wholesale markets, where crops are shipped abroad for overseas processing in larger-scale and more efficient plants. The transition from rainfed to artificially irrigated agriculture through surface or groundwater electrical pumping thus provides a relevant example of how an electricity input could dramatically boost rural productivity. Moreover, generating value added through local crop processing (Kyriakarakos et al., 2020) and retaining it among farms would considerably boost local socio-economic prospects, with the potential to set a positive feedback involving the entire local rural community. To enable these uses, the provision of energy is necessary (Barnes and Floor, 1996; Cabraal et al., 2005; Kirubi et al., 2009; Pueyo and Maestre, 2019), along with the purchase of machineries and infrastructure. In fact, currently 85% of the global population without electricity access is concentrated in rural areas (IEA et al., 2020). While planning energy solutions which can comprehensively enable agricultural uses might increase the required power capacity and upfront investment, it might also render them economically attractive because of the significant reduction in the payback time of those investment thanks to the increased rural productivity (Kyriakarakos et al., 2020).

Following this paradigm, here (**Figure 3**) we estimate the energy requirements to enable sufficient artificial irrigation (**SI-A6-7**) and raw crop processing to more refined crop products (**SI-A8**), with the final objective of evaluating the potential local economic gains (**SI-A9**). For irrigation modelling, Agricultural land, hydroclimatic factors, and cropping patterns information is conveyed in a set of agroclimatic equations to estimate daily irrigation water requirements in each cluster. Then, a groundwater pump model estimates the required power and flow rate of the pump as a function of the groundwater dwell characteristics and of the irrigation requirements.

For crop processing energy, an extensive literature review of crop processing energy requirements in the context of developing countries is carried out and associated to crop-specific cropland extent and average yield in each cluster. Finally, the most recent database of wholesale prices for a large basket of crops in Kenya relative the location of each wholesale market is multiplied to the local potential for yield increase of each crop, net of transportation and total (installation, operation, and maintenance) pumping costs.

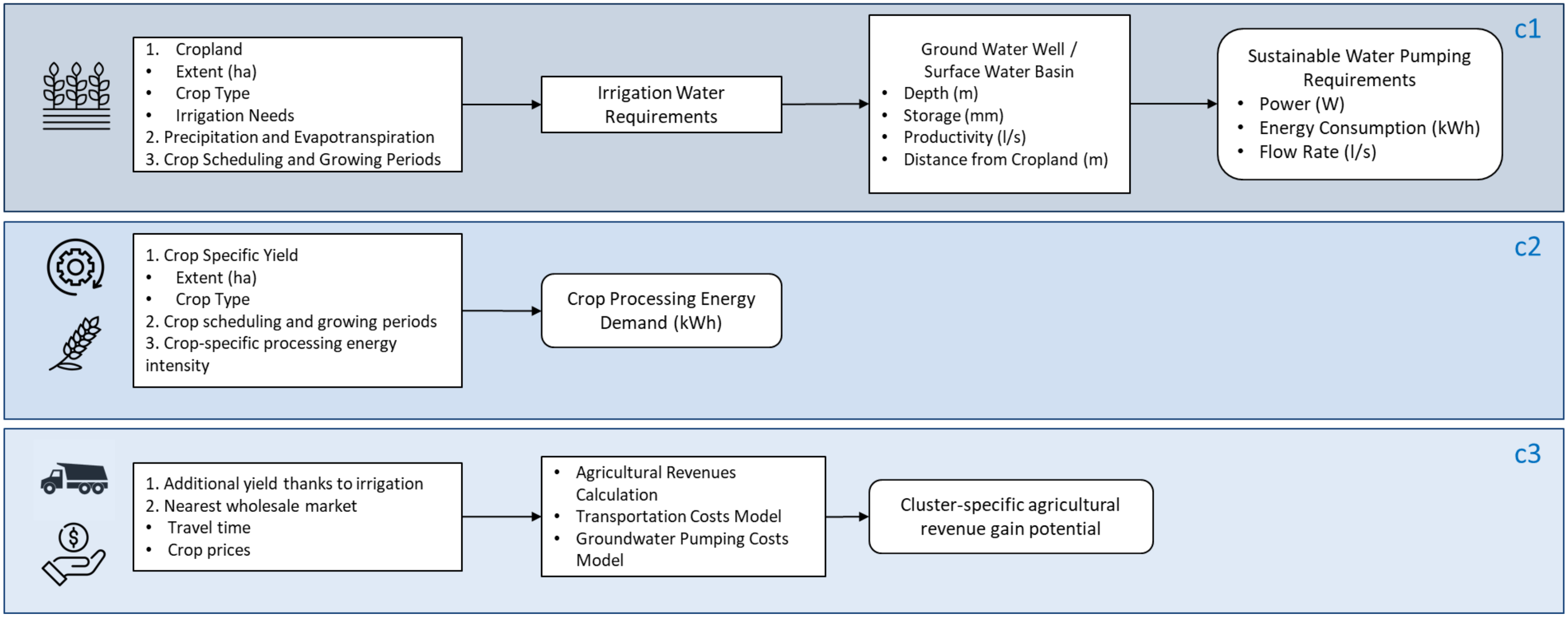

Figure 3: Workflow of the agricultural sector in the M-LED framework. (c1) *Water pumping electricity requirement estimation procedure.* ***(c2)*** *Crop processing electricity demand estimation.* ***(c3)*** *Agricultural revenues calculation.*

3. Results

3.1. An applicative example for Kenya: electricity demand revised

We select Kenya as a country-case study to provide a proof-of-concept of the implementation of the M-LED framework to evaluate sectoral, spatial and temporal energy demand heterogeneity. The selection is the result of two factors. First, data and geospatial information availability in Kenya is remarkable compared to most of SSA countries, which renders the platform implementation comparatively more accurate. Second, a large number of assessments have been carried out on electricity access planning in Kenya (Berggren and Österberg, 2017; Fabini et al., 2014; Moksnes et al., 2017; Moner-Girona et al., 2019; Parshall et al., 2009), and thus there are significant opportunities for better understanding the impacts of our multi-sectoral, bottom-up electricity demand modelling on the outputs of several electrification planning models. On top of it, the lack of available and complete field energy profile data in Kenya offers the opportunity to the M-LED to evaluate and intercompare the significance of the different demand scenarios. A selected list of these studies – all focusing on geospatial electrification analysis for Kenya but applying different tools and assumptions – include refs. (Lee et al., 2016a; Moksnes et al., 2017; Moner-Girona et al., 2019; Parshall et al., 2009).

The panels of **Figure 4** provide the resulting spatially-explicit (the original results are at a polygonal cluster-dependent resolution; here to ensure a more immediate understanding, they are plotted on a 1x1 km grid) sectoral electricity latent demand generated for Kenya with the M-LED platform. The estimated demand encompasses multiple dimensions: sectoral granularity; monthly seasonality in the demand; hourly profile; and spatial distribution of the demand.

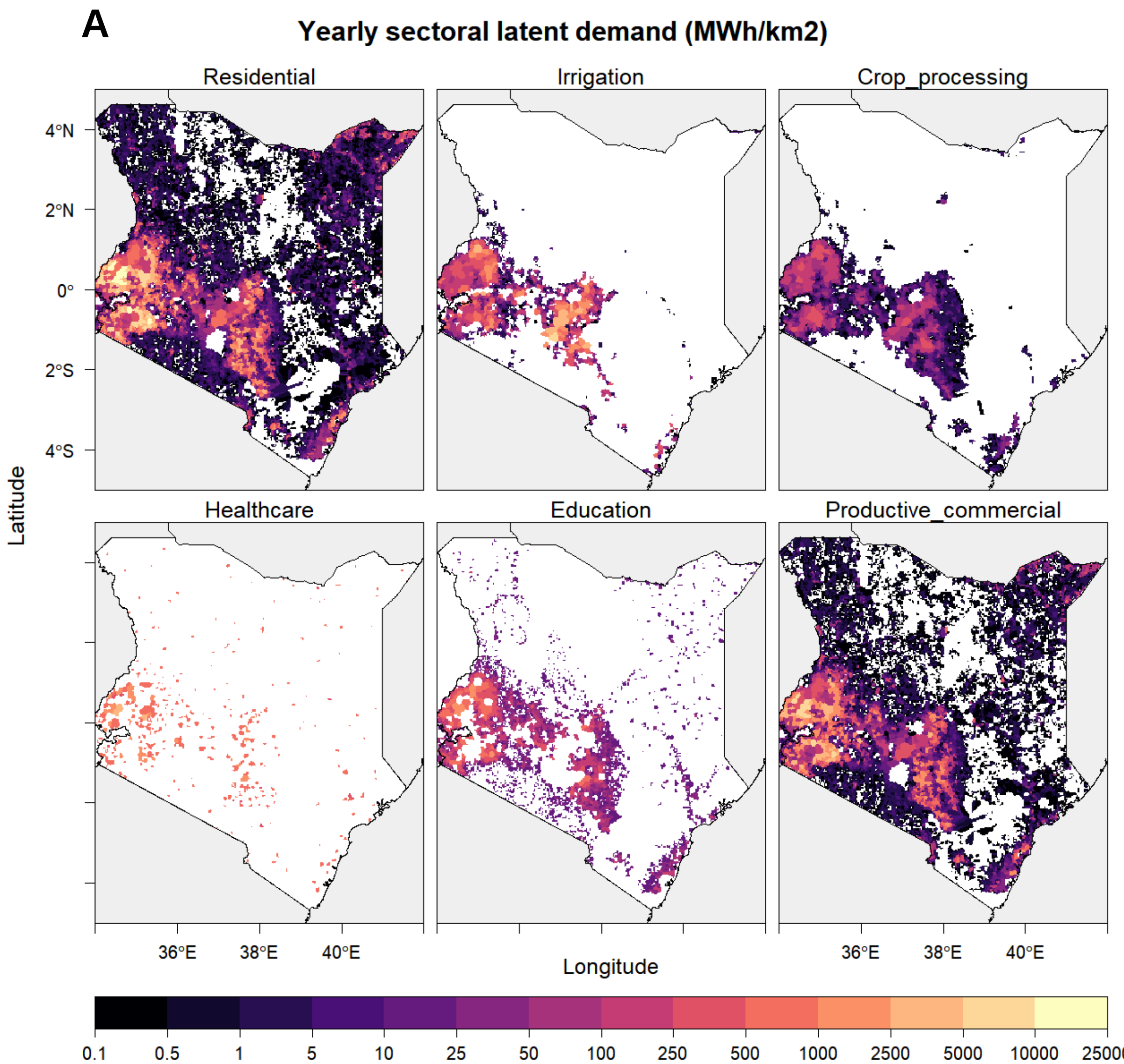

Figure 4A shows the distribution over space of yearly sectoral latent electricity demand density (MWh/year/km^2). Note that white pixels identify areas with either no population or no sectoral latent electricity demand, such as natural parks, protected areas, or cropland (for sectors different from agriculture). The results show that substantial heterogeneity is observed in the residential and commercial and micro-enterprise demand: both are highly correlated with population density, with significantly higher latent demand in south-western Kenya. Yet in some areas (e.g. in northern Kenya) commercial and micro-enterprise demand is comparatively lower than the residential demand because of lower employment and market accessibility. Irrigation and crop processing electricity demand are concentrated in the agricultural districts in the south-west of Kenya, while healthcare and education demand are more scattered across the country, although with higher density in higher density populated areas. In particular, healthcare facilities are highly sparse but at the same time exhibit a high demand density, while schools are relatively more distributed but less electricity intensive.

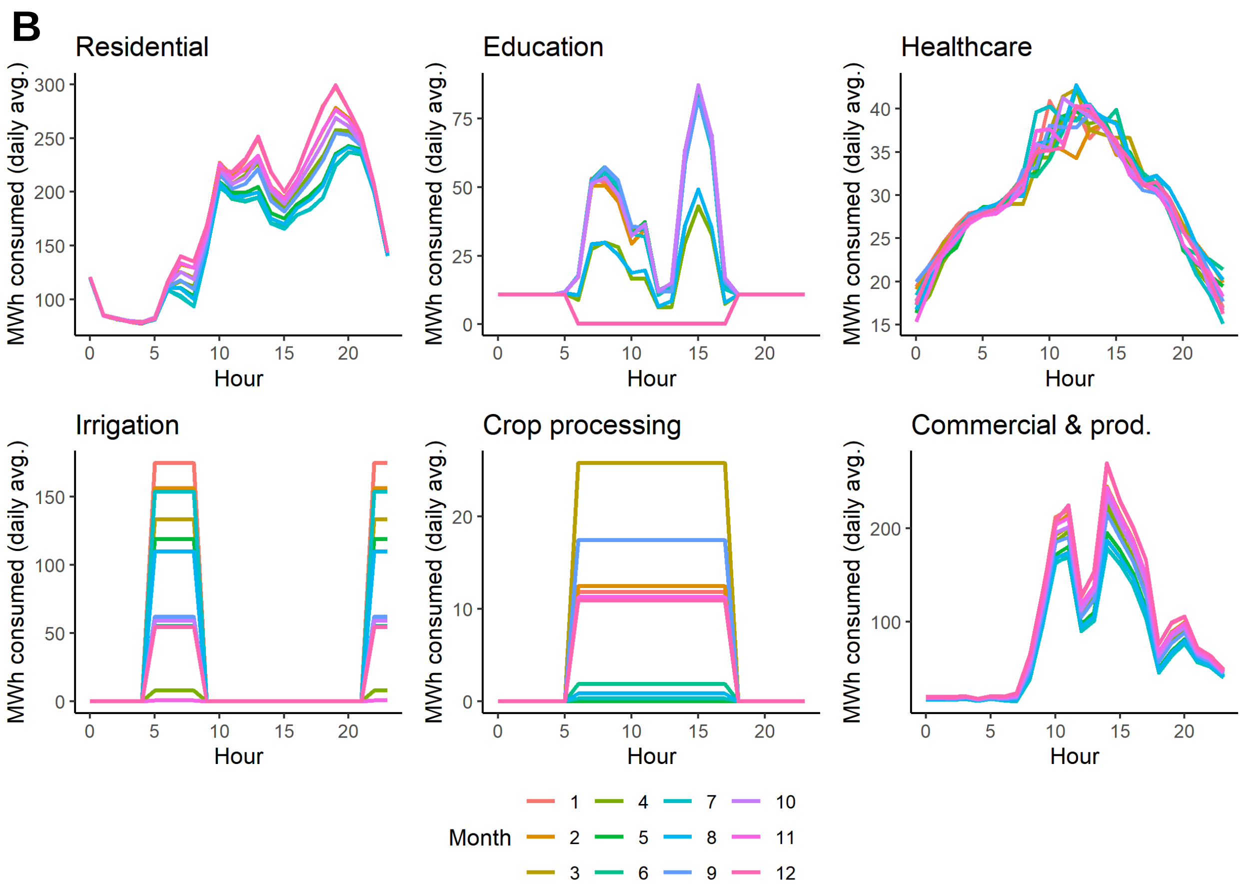

Figure 4B depicts the hourly and monthly distribution of the demand across sectors. Residential demand shows a curve with three peaks, during wake-up, lunch, and evening times. A similar polymodal distribution characterises commercial and micro-enterprise demand Most of the seasonality is explained by the variation in the use of air circulation and cooling appliances, since residual uses are rather invariant throughout the year given the proximity of Kenya to the equator. Educational centres show variation in months of year and term breaks with energy demand bimodal distribution with peaks in the morning and in the afternoon. Healthcare results show relatively little seasonal variation, with unimodal normal distribution with a peak at midday for healthcare. Agricultural-related activities show high seasonal variance in the monthly profiles, but the load of the two curves are however flat throughout the energy use windows, 5 am – 9 am and 9 – 11 pm for irrigation and 6 am – 6 pm for crop processing machinery.

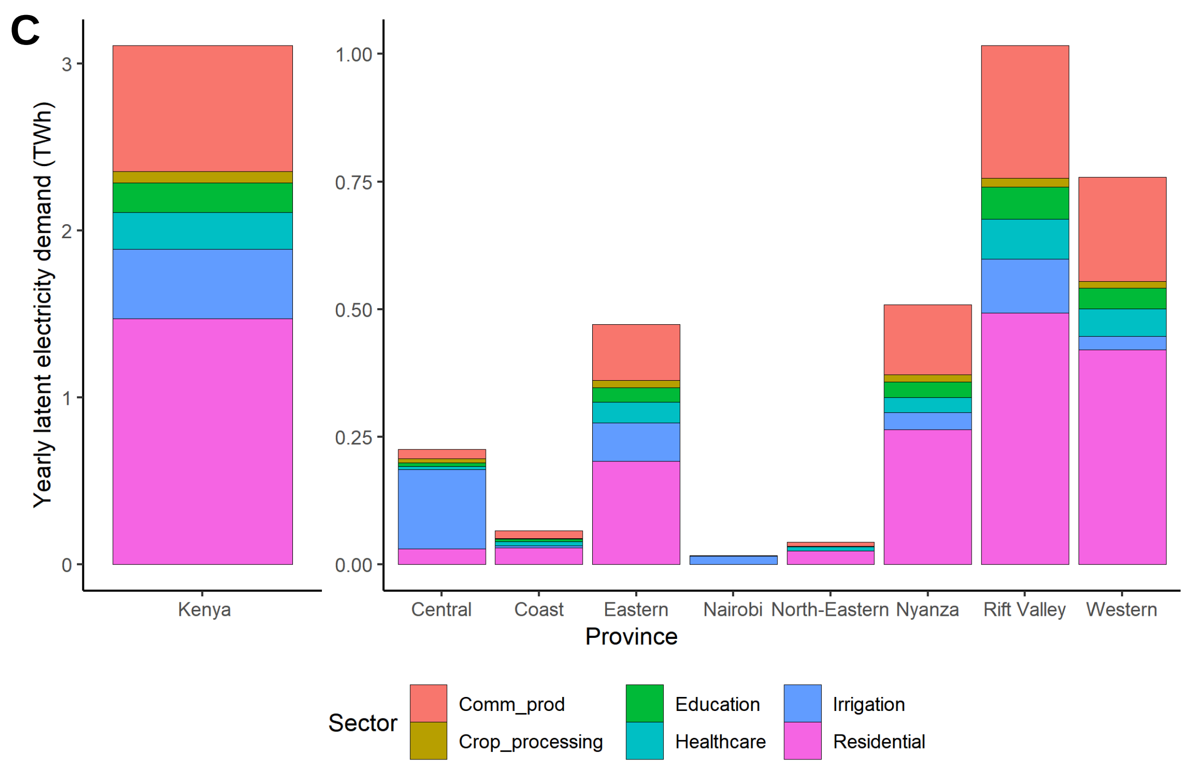

Figure 4C summarises the yearly aggregated latent demand across sectors in Kenya and its repartition among the eight regions of (visualised on a map in **Figure 4D)**. The country-wide aggregation shows that the supply requirements are unevenly split into the residential (at about 1.5 TWh/year, or 48% of the total 3.1 TWh/year), commercial activities and microenterprises (nearly 0.75 TWh/year, about one quarter of the total), healthcare (about 0.22 TWh/year, or 7%), education (0.18 TWh/year, 5.7%), irrigation (0.42 TWh/year, 13.5%), and crop processing sectors (about 0.07 TWh/year, only about 2%). Additional insights are drawn when considering the repartition of those aggregate energy requirements across the eight main regions of Kenya, as well as the shares of each sector within each region. The Rift Valley region is the region with the largest latent demand (about one third of the total latent demand), driven mainly by the residential and productive sectors; it is followed by the Western region (about 25% of the country latent demand), with a similar repartition. Notably, in the Central region irrigation latent energy is by far the first sector (>two thirds of the total).

D

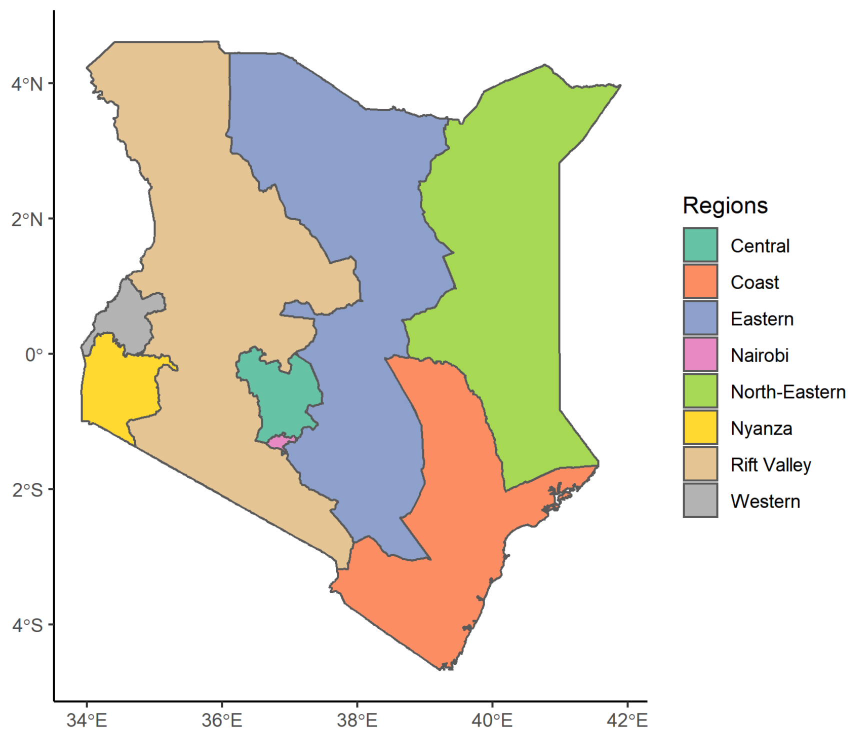

***Figure 4: Sectoral demand loads in population clusters of Kenya estimated with the M-LED methodology. (A) Maps of Kenya** representing: (i) the estimated residential demand density for households that require electrification (MWh/year/km^2); (ii) the total healthcare and education demand density for facilities requiring electrification (MWh/year/km^2); (iii) the water pumping and crop processing demand density (MWh/year/km^2); (iv) the micro-enterprise and commercial activities*

electricity demand density (MWh/year/km^2); ***(B) Total (country-wide) typical daily sectoral load profiles, by month*** *(MWh/hour in each demand cluster); (****C) Barplots comparing the yearly total regional and country-level sectoral latent electricity demand*** *(TWh/year); (****D) Map of the corresponding regions in Kenya.***

3.2. Comparison of the estimated demand with previous studies

A systematic comparison of our results with previous demand estimates found in the literature (in most cases used to parametrise geospatial supply-side electrification models) is not straightforward. This is because of the differences in both how this demand is formulated (e.g. yearly sectoral consumption in kWh or representative day load curves in W) and how it is parsed to settlements (urban/rural, poor/non-poor). Nonetheless, a number of insights can still be drawn.

For instance, Moksnes et al. (Moksnes et al., 2017) adopt tier-based values of 44 and 423 kWh/capita/year for rural households and of 423 and 599 kWh/capita/year for urban households in the two scenarios they consider. This yields to average demand values of 141 and 468 kWh/capita/year for households to be electrified. Yet, the study considers neither the temporal variability in the demand nor additional demand sectors.

Parshall et al. (Parshall et al., 2009) allocate household demand to a range of 360-1800 kWh/hh/year depending on their urban or rural status and the prevalence of poverty in the region where they are located. Productive demand is fixed across the same categorisation, with values between 50 and 340 kWh/hh/year. This results in an average yearly productive to residential demand ratio of 0.18 Irrespective of the model not encapsulating an explicit temporal dimension of the demand, a peak load is assumed across productive, service, and institutional uses of energy through a peak coincidence factor. The authors assume the following yearly total electric consumption for different facilities: clinic – 360 kWh/year, dispensary – 600 kWh/year, health centre – 2400 kWh/year, primary school (day) – 1200 kWh/year, secondary school – 2400 kWh/year, boarding school – 15,000 kWh/year. Hospitals were not included because they were assumed to already have adequate access to electricity.

Moner-Girona et al. (Moner-Girona et al., 2019) define a different load profile for each energy demand sector. each load profile is the same all year round without seasonal variability but different load peak depending on the location (i.e. number of population) year. In particular, for productive activities small-scale industrial infrastructures with a range of 1500 kWh/year to 3100 kWh/year and commercial activities with a range of 1200 kWh/year to 1800 kWh/year year are considered, while for household demand they follow the approach of (Parshall et al., 2009) to allocate Tier 3 and Tier 4 yearly consumption values, i.e. 365 and 1020 kWh/hh/year.

In the M-LED platform application for Kenya we estimate average urban and rural residential electricity demand of 62 and 842 kWh/hh/year, respectively. Yet, it must be remarked these values do not represent the heterogeneity in the demand that characterised our methodology. The country-wide average yearly productive (commercial and agricultural) to residential demand ratio of our assessment is of about 0.8, while the services (healthcare and education) to residential demand ratio is of 0.25. We calculate average healthcare facility consumption values of 2,200 kWh/year for dispensaries, 10,900 kWh/year for health centres and 124,886 kWh/year for sub-district hospitals. For schools, we estimate a value of about 6,000 kWh/year for a 700 pupils institute.

In general, this comparison suggests that the detailed characterisation of our study leads to significant differences with a number of previous studies. Firstly, including productive sectors in our characterisation increases notably both the total load of settlements and the productive-to-residential demand ratio. Secondly, it leads to a larger spread in the residential demand between urban and rural areas. Yet, when encapsulating activities such as artificial irrigation and crop processing, the gap in the demand between settlement types is reduced.

On the other hand, a visual comparison of demand maps suggests that the spatial distribution of demand hotspots is identified similarly through different approaches, provided sectors additional to the residential demand are considered. This is because the key non-residential demand drivers considered in these studies are often similar and highly correlated among each other, such as population density, urban/rural prevalence, poverty density or wealth distribution, and the geographical position of service and productive infrastructure and of crop fields. Yet, studies focussing on achieving universal electrification based on residential demand only flatten the heterogeneity in the demand. For instance, by setting a top-down rural demand they significantly underestimate the demand of rural settlements compared to urban areas;

3.3. Impacts of artificial irrigation: increasing local agricultural revenues

The M-LED platform allows the cost-benefit analysis at (partial) local micro-economic level (**SI-A9**). The cost-benefit analysis of the increased agricultural yield due to groundwater pumping serves as an applicative example to show one of the many aspects of local development that could be triggered by electrification. The analysis estimates the irrigation needs to close the yield gap by calculating the current yield (in t/ha) of each crop in each agricultural cluster and comparing it with the mean yield of the same type of crop in global areas falling in the same irrigated agro-ecological zone. The workflow then evaluates the potential local economic value added. The potential revenues for local producers are calculated by subtracting transport and total pumping costs to revenues (in turn calculated assuming the wholesale crop price in local markets). It is crucial to remark that these revenues are direct revenues to the producers, so they do not include so does not include export, taxes, and additional cost components. For the Kenyan case study, each crop wholesale price is assumed to be the 2019 price observed at the nearest wholesale market to each functional agricultural cluster (obtained from NFAIS, the National Farmers Information System of Kenya; http://www.nafis.go.ke). The transportation costs of crops from field to wholesale markets are calculated including the fuel consumption, truck rental, and time cost of carrying the extra agricultural production to the market following the shortest path based on recent accessibility maps (Weiss et al., 2018). The pumping costs are calculated estimating a multivariate regression of total pumping costs (including installation, operation, and maintenance components) on the well depth, the pump yield, and their interaction based on real field data from Xenarios and Pavelic (2013).

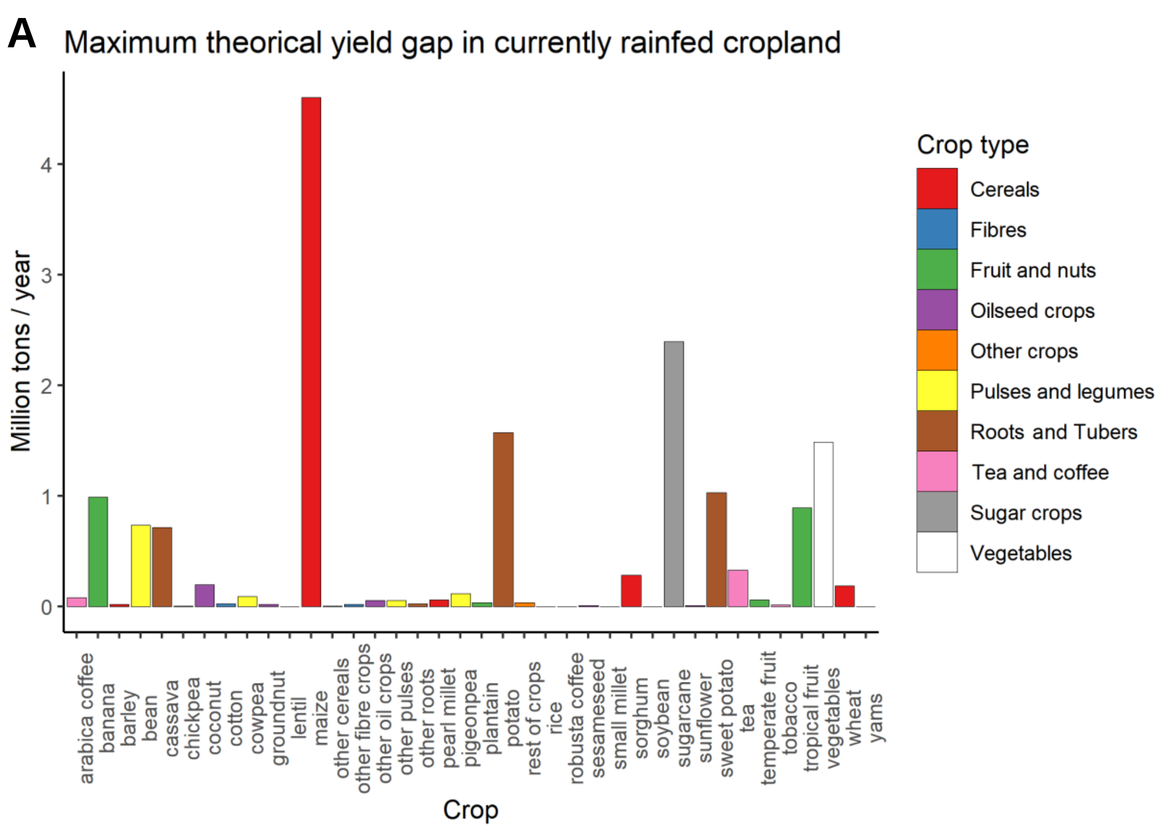

Figure 6A shows the maximum theoretical yield gap in current cropland for each specific crop. These aggregated values express the national yield gain potential if cropland was optimally irrigated, fertilised and managed (the latter two components are not modelled in this study). The results show that significant increase in the crop production exists for maize (>6 million tons/year), potatoes (>2 million tons/year), sugarcane (about 2.5 million tons/year), and bananas and fresh vegetables (both at about 2 million tons/year). Yet, the effective profitability of this potential yield gains is a function of several factors: crop prices at wholesale markets, groundwater pumping costs, and transportation distance and time (and thus costs) to these markets. In electrification supply-side analysis, agricultural revenues can be compared to the local electrification investment requirement to assess what would be the payback time of the local electricity access investment if it was covered by the additional agricultural yield generated thanks to electrification itself.

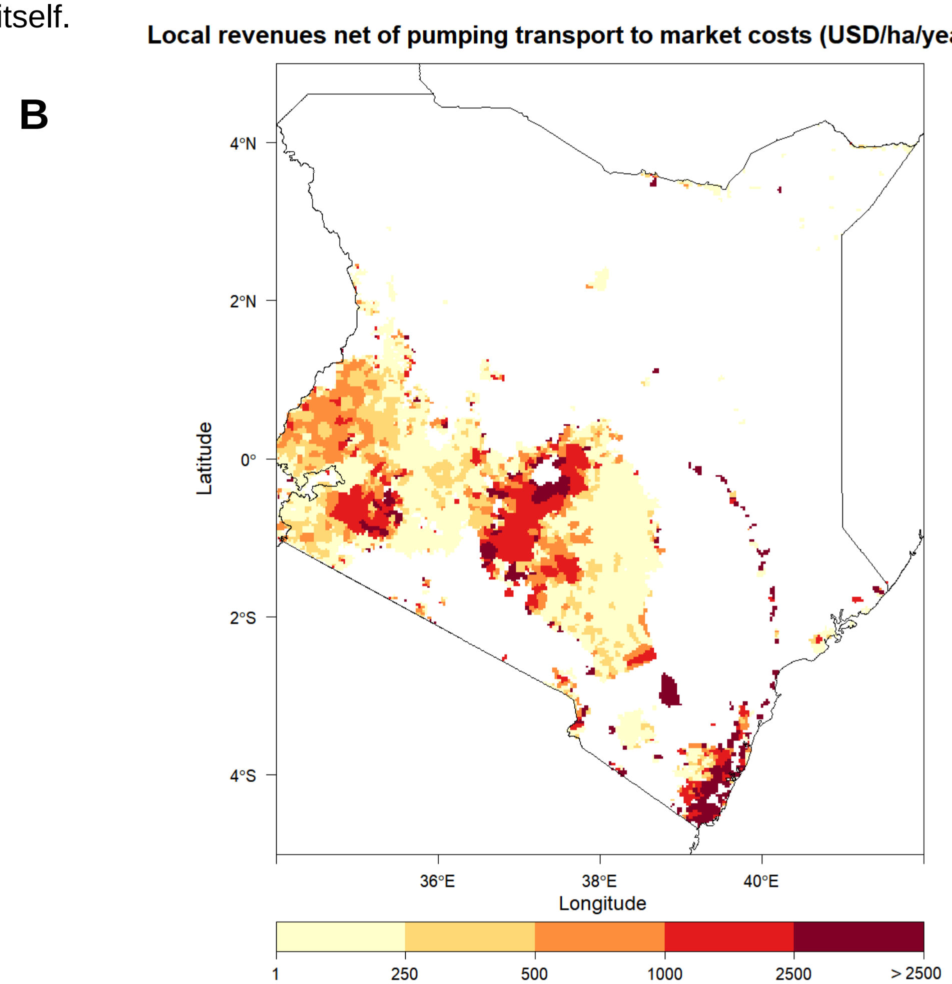

Figure 6: Potential revenues from the increased agricultural yield thanks to artificial irrigation. (A) Maximum theoretical yield gap for each deployed crop in Kenya (yield comparison with and without artificial irrigation) **(B)** *Map of Kenya showing the total revenue gain (in USD/ha/year) at current crop-specific market prices including subtraction of field to market transportation costs and groundwater pumping total costs.*

Figure 6B summarises the results of the model-based assessment with a map of Kenya plotting the potential revenues (net of transportation and pumping costs) from the increased agricultural yield thanks to artificial irrigation at each cluster. The map shows that in rural Kenya there are vast areas with gain potential of up to 2500 USD/ha (especially in the already comparatively more profitable agricultural district in western Kenya), and even larger areas with more modest but widespread revenue growth potential. These potential gains are very relevant especially if compared to the current income levels of rural Kenya. The proportion of Kenyans living on less than the international poverty line is in fact at 36.1% ("Poverty Incidence in Kenya Declined Significantly, but Unlikely to be Eradicated by 2030," n.d.). The poverty line is set at 1.90 USD per day in 2011 PPP; thus, assuming an average household size of 3.5 (United Nations, Department of Economic and Social Affairs, Population Division, 2019), as an yearly household income of 2,427 USD. Overall, summing all the potential revenues in the country, a total potential of $4.9 billion/year results, about 5% of the 2019 Kenyan GDP.

4. Conclusion and policy implications

A detailed formulation of electricity demand is a crucial factor in energy access planning. This is also reflected in the outcome of supply-side electrification models. Here we have introduced M-LED, a flexible platform for generating electricity demand curves based on a multi-sectoral bottom-up device-based approach. We have then applied the platform to the country-study of Kenya.

The analysis provided an array of novel policy-relevant insights, the crucial ones being that modelling electrification based on residential demand only is likely to strongly underestimate the total demand of settlements (and chiefly rural areas), confirming recent assessments in the literature (Moner-Girona et al., 2019). In particular, including healthcare, education, commercial and micro-enterprise, and agricultural energy uses implies (country-wide) a more than doubling of the estimated yearly latent demand *vis-à-vis* residential only. This mark-up is even greater in agriculture-intensive rural areas where energy uses for irrigation and crop processing might be significant higher in relative terms. Another crucial insight is given by our hourly and seasonal-variant formulation of sectoral load curves, which could have a significant impact on the optimisation of energy systems, in particular when paired with variable renewable energy supply curves.

This paper introduces the demand estimation methodology and results. Yet, future functionalities, currently in the design stage, will link the high-resolution hourly, seasonal, and sectoral demand estimates into an array of electricity supply planning models. The new functionality will allow to carry out an independent assessment for several electrification planning models and understand the significance of considering the new multi-sectoral and seasonal dimensions.

Concerning the specific country-study of Kenya, our analysis reveals that the sectors considered in this study as additional to the residential sector constitute a very relevant share of the total latent demand in areas electricity access deficit. In aggregated terms, they account for ~53% of the yearly latent electricity demand, or 1.65 TWh/year. The ratio between residential and non-residential demand is even more pronounced in the Central region, where although the household electricity access levels are already quite high, agriculture-related activities necessitate significant electricity input which today is largely missing. Additionally, in population-dense areas productive and commercial demand also has a significant impact on the final regional demand.

On top of the detailed latent electricity demand results, the M-LED platform enables an analysis of the potential economic returns from the agricultural sector as a result of the artificial irrigation. This reveals an untapped revenue potential (net of transportation and groundwater pumping costs) of about $4.9 billion/year (about 5% of the 2019 Kenyan GDP). This suggests significant economic potential that in many areas may quickly pay back the electrification investment if properly accounted by decision makers in the cost-benefit analysis and supported by policies stimulating improved land management and fertilisation. Yet, it must be remarked that additional relevant dimensions that might affect the results of the analysis in the future include the price change of products owing to crop processing and local value creation and the efficiency gains in transport from improved road or rail transportation and logistics.

The M-LED platform is open-source and fully customisable to let the user define the bulk of the technical and economic parameters, the devices ownership and usage patterns, and the overall infrastructure. Irrespective of the large amount of work involved in the development of the M-LED platform and in the formulation of its assumptions, limitations remain. Firstly, a limited number of sectors is estimated; secondly, the data-intensiveness of the analysis implies growing uncertainty over the reliability of the database, as (despite a careful data selection and wrangling) some sources such as infrastructure and facilities location and characteristics might be outdated or biased; thirdly, while the appliance ownership and use baskets are designed after a careful literature screening supported by field campaign experience of the authors, residual cultural, service, and economic heterogeneity might not be captured in the analysis; moreover, in the supply-side analysis a relevant role is played by the techno-economic characterisation of technologies, which might however be affected by future policies such as subsidies and taxes or specific regulatory frameworks; finally, the water and agricultural analysis stands on the assumption of an optimal irrigation scheduling and local crop processing based on current cropping patterns.

Our results are potentially beneficial for policy makers, researchers, consultants, and other stakeholders involved in the electrification planning. For instance, the results could contribute to the prioritisation decisions for the allocation of limited governmental funding by leveraging consumers who are likely to have the greatest impact on increasing economic growth thanks to the provision of electricity to existing productive activities or attracting private investments in the most productive areas.

We encourage further research on the topic and improvements to the state of the M-LED platform introduced at the time of the writing of this paper. A better characterisation of potential industrial demand and a dynamic formulation of demand (with intertemporal growth based on income and other determinants) represent potential first-order improvements.

Code and data availability

The M-LED platform source code and the accompanying documentation are available at https://github.com/giacfalk/MLED. An archive containing all the data inputs for replicating the Kenya analysis will be made available on Zenodo.

References

Abatzoglou, J.T., Dobrowski, S.Z., Parks, S.A., Hegewisch, K.C., 2018. TerraClimate, a high-resolution global dataset of monthly climate and climatic water balance from 1958–2015. Sci. Data 5, 170191.

Action, P., 2013. Poor people's energy outlook 2013: Energy for community services. Rugby UK Pract. Action Publ.

Adair-Rohani, H., Zukor, K., Bonjour, S., Wilburn, S., Kuesel, A.C., Hebert, R., Fletcher, E.R., 2013. Limited electricity access in health facilities of sub-Saharan Africa: a systematic review of data on electricity access, sources, and reliability. Glob. Health Sci. Pract. 1, 249–261. https://doi.org/10.9745/GHSP-D-13-00037

Adeoye, O., Spataru, C., 2019. Modelling and forecasting hourly electricity demand in West African countries. Appl. Energy 242, 311–333.

Adkins, J.E., Modi, V., Sherpa, S., Han, R., Kocaman, A.S., Zhao, N., Natali, C., Carbajal, J., 2017. A geospatial framework for electrification planning in developing countries, in: 2017 IEEE Global Humanitarian Technology Conference (GHTC). Presented at the 2017 IEEE Global Humanitarian Technology Conference (GHTC), pp. 1–10. https://doi.org/10.1109/GHTC.2017.8239293

Aguirre, J., 2017. The impact of rural electrification on education: A case study from Peru.

A.J., F., C., C., D., E., S., F., T., K., L., M., M., M., M., P., P., P., M., S., F., S., L, Z., 2019. GHSL Data Package 2019. https://doi.org/10.2760/062975

Allen, R.G., Pereira, L.S., Raes, D., Smith, M., 1998. Crop evapotranspiration-Guidelines for computing crop water requirements-FAO Irrigation and drainage paper 56. Fao Rome 300, D05109.

Arderne, C., 2020. carderne/clusterize.

Barnes, D.F., Floor, W.M., 1996. RURAL ENERGY IN DEVELOPING COUNTRIES: A Challenge for Economic Development. Annu. Rev. Energy Environ. 21, 497–530. https://doi.org/10.1146/annurev.energy.21.1.497

Berggren, A., Österberg, S., 2017. Kenya meeting the electricity demand of 2030: An assessment of how Kenya Vision 2030 and climate change impact the optimal electricity generation mix.

Bernard, T., 2010. Impact analysis of rural electrification projects in Sub-Saharan Africa. World Bank Res. Obs. 27, 33–51.

Bhatia, M., Angelou, N., 2015. Beyond connections: energy access redefined. World Bank.

Blimpo, M.P., Cosgrove-Davies, M., 2019. Electricity Access in Sub-Saharan Africa: Uptake, Reliability, and Complementary Factors for Economic Impact. World Bank Publications.

Blodgett, C., Dauenhauer, P., Louie, H., Kickham, L., 2017. Accuracy of energy-use surveys in predicting rural mini-grid user consumption. Energy Sustain. Dev. 41, 88–105.

Booysen, F., 2002. An overview and evaluation of composite indices of development. Soc. Indic. Res. 115–151.

Bos, K., Chaplin, D., Mamun, A., 2018. Benefits and challenges of expanding grid electricity in Africa: A review of rigorous evidence on household impacts in developing countries. Energy Sustain. Dev. 44, 64–77.

Bose, T.K., Uddin, M.R., Mondal, A., 2013. Impacts of electricity access to rural SMEs. Int. J. Manag. Value Supply Chains 4, 17.
Brüderle, A., Attigah, B., Bodenbender, M., 2011. Productive use of energy–PRODUSE a manual for electrification practitioners. Eschborn Ger. GTZ.
Cabraal, R.A., Barnes, D.F., Agarwal, S.G., 2005. PRODUCTIVE USES OF ENERGY FOR RURAL DEVELOPMENT. Annu. Rev. Environ. Resour. 30, 117–144. https://doi.org/10.1146/annurev.energy.30.050504.144228
Cardona, M.J., López, Y.U., 2018. Technical-Economic Electrification Models Rural with Renewable Energies: Systematic Review of Literature, in: 2018 IEEE ANDESCON. IEEE, pp. 1–7.
Center for International Earth Science Information Network - CIESIN - Columbia University, Information Technology Outreach Services - ITOS - University of Georgia, 2013. Global Roads Open Access Data Set, Version 1 (gROADSv1). NASA Socioeconomic Data and Applications Center (SEDAC), Palisades, NY.
Daka, K.R., Ballet, J., 2011. Children's education and home electrification: A case study in northwestern Madagascar. Energy Policy 39, 2866–2874.
Elvidge, C.D., Baugh, K., Zhizhin, M., Hsu, F.C., Ghosh, T., 2017. VIIRS night-time lights. Int. J. Remote Sens. 38, 5860–5879. https://doi.org/10.1080/01431161.2017.1342050
Fabini, D.H., Baridó, D.P. de L., Omu, A., Taneja, J., 2014. Mapping Induced Residential Demand for Electricity in Kenya. Proc. Fifth ACM Symp. Comput. Dev. - ACM DEV-5 14 5, 43–52. https://doi.org/10.1145/2674377.2674390
Facebook Connectivity Lab, Center for International Earth Science Information Network - CIESIN - Columbia University, 2016. High Resolution Settlement Layer (HRSL).
Falchetta, G., Pachauri, S., Byers, Edward, Danylo, Olha, Parkinson, SImon, 2020. Satellite observations reveal inequalities in the progress and effectiveness of recent electrification in sub-Saharan Africa. One Earth.
Falchetta, G., Pachauri, S., Parkinson, S.C., Byers, E., 2019. A High-Resolution Gridded Dataset to Assess Electrification in Sub-Saharan Africa. Sci. Data.
Fischer, G., Nachtergaele, F.O., Prieler, S., Teixeira, E., Tóth, G., van Velthuizen, H.T., Verelst, L., Wiberg, D., 2012. Global agro-ecological zones (GAEZ v3. 0).
Fried, S., Lagakos, D., 2017. Rural electrification, migration and structural transformation: Evidence from Ethiopia. Unpubl. Manuscr. Univ. San Diego.
Giday, Z.G., 2014. Technical and economic assessment of solar PV/diesel hybrid power system for rural school electrification in Ethiopia. Int. J. Renew. Energy Res. IJRER 3, 735–744.
Grimm, M., Sparrow, R., Tasciotti, L., 2015. Does electrification spur the fertility transition? Evidence from Indonesia. Demography 52, 1773–1796.
GYGA, 2017. Global yield gap atlas.
Hijmans, R., Garcia, N., Wieczorek, J., 2018. GADM: database of global administrative areas. Version 36.
Huld, T., Moner-Girona, M., Kriston, A., 2017. Geospatial Analysis of Photovoltaic Mini-Grid System Performance. Energies 10, 218. https://doi.org/10.3390/en10020218

IEA, 2019. Africa Energy Outlook 2019.
IEA, IRENA, United Nations Statistics Division, 2020. Tracking SDG7: The Energy Progress Report 2020.
Jägermeyr, J., Gerten, D., Schaphoff, S., Heinke, J., Lucht, W., Rockström, J., 2016. Integrated crop water management might sustainably halve the global food gap. Environ. Res. Lett. 11, 025002.
Kariuki, D., 2016. Rural Electrification and Microenterprises Performance: Some Lessons from Muranga County Kenya. Int. J. Econ. 1, 31–45.
Kemausuor, F., Adkins, E., Adu-Poku, I., Brew-Hammond, A., Modi, V., 2014. Electrification planning using Network Planner tool: The case of Ghana. Energy Sustain. Dev. 19, 92–101. https://doi.org/10.1016/j.esd.2013.12.009
Kenya Open Data Initiative - Humanitarian Data Exchange [WWW Document], n.d. URL https://data.humdata.org/organization/kenya-open-data-initiative (accessed 6.23.20).
Kirubi, C., Jacobson, A., Kammen, D.M., Mills, A., 2009. Community-Based Electric Micro-Grids Can Contribute to Rural Development: Evidence from Kenya. World Dev. 37, 1208–1221. https://doi.org/10.1016/j.worlddev.2008.11.005
Kongolo, M., 2010. Job creation versus job shedding and the role of SMEs in economic development. Afr. J. Bus. Manag. 4, 2288.
Korkovelos, A., Khavari, B., Sahlberg, A., Howells, M., Arderne, C., 2019. The Role of Open Access Data in Geospatial Electrification Planning and the Achievement of SDG7. An OnSSET-Based Case Study for Malawi. Energies 12, 1395. https://doi.org/10.3390/en12071395
Kotikot, S.M., Ajinjeru, C., Odukomaiya, A., Omitaomu, O.A., 2018. Geospatial Framework for Estimating Household Electricity Demand for Urban Infrastructure Planning in Select African Countries, in: 2018 IEEE PES/IAS PowerAfrica. IEEE, pp. 613–618.
KTH-dESA/PopCluster, 2019. . KTH - dESA.
Kyriakarakos, G., Balafoutis, A.T., Bochtis, D., 2020. Proposing a Paradigm Shift in Rural Electrification Investments in Sub-Saharan Africa through Agriculture. Sustainability 12, 3096. https://doi.org/10.3390/su12083096
Lee, K., Brewer, E., Christiano, C., Meyo, F., Miguel, E., Podolsky, M., Rosa, J., Wolfram, C., 2016a. Electrification for "Under Grid" households in Rural Kenya. Dev. Eng. 1, 26–35. https://doi.org/10.1016/j.deveng.2015.12.001
Lee, K., Miguel, E., Wolfram, C., 2016b. Appliance ownership and aspirations among electric grid and home solar households in rural Kenya. Am. Econ. Rev. 106, 89–94.
Leea, S.J., Sánchezb, E., González-Garcıaa, A., Cillerc, P., Duenasa, P., Tanejad, J., de Cuadra Garcıac, F., Lumbrerasb, J., Dalyf, H., Stonera, R., others, 2019. Investigating the Necessity of Demand Characterization and Stimulation for Geospatial Electrification Planning in Developing Countries. studies 12, 22.
Lombardi, F., Balderrama, S., Quoilin, S., Colombo, E., 2019. Generating high-resolution multi-energy load profiles for remote areas with an open-source stochastic model. Energy 177, 433–444.

MacDonald, A.M., Bonsor, H.C., Dochartaigh, B.É.Ó., Taylor, R.G., 2012. Quantitative maps of groundwater resources in Africa. Environ. Res. Lett. 7, 024009.

Maina, J., Ouma, P.O., Macharia, P.M., Alegana, V.A., Benard Mitto, Ibrahima-Socé Fall, Abdisalan Mohamed Noor, Snow, R.W., Emelda Aluoch Okiro, 2019. A spatial database of health facilities managed by the public health sector in sub Saharan Africa. https://doi.org/10.6084/m9.figshare.c.4399445

Manggat, I., Zain, R., Jamaluddin, Z., 2018. The impact of infrastructure development on rural communities: A literature review. Int. Acad. Res. Bus. Soc. Sci. 8, 647–658.

McCollum, D.L., Echeverri, L.G., Busch, S., Pachauri, S., Parkinson, S., Rogelj, J., Krey, V., Minx, J.C., Nilsson, M. ans, Stevance, A.-S., others, 2018. Connecting the sustainable development goals by their energy inter-linkages. Environ. Res. Lett. 13, 033006.

Mentis, D., Howells, M., Rogner, H., Korkovelos, A., Arderne, C., Eduardo Zepeda, Siyal, S., Taliotis, C., Bazilian, M., Roo, A. de, Tanvez, Y., Alexandre Oudalov, Scholtz, E., 2017. Lighting the World: the first application of an open source, spatial electrification tool (OnSSET) on Sub-Saharan Africa. Environ. Res. Lett. 12, 085003.

Moksnes, N., Korkovelos, A., Mentis, D., Howells, M., 2017. Electrification pathways for Kenya–linking spatial electrification analysis and medium to long term energy planning. Environ. Res. Lett. 12, 095008.

Moner-Girona, M., Bódis, K., Huld, T., Kougias, I., Szabó, S., 2016. Universal access to electricity in Burkina Faso: scaling-up renewable energy technologies. Environ. Res. Lett. 11, 084010.

Moner-Girona, M., Bódis, K., Morrissey, J., Kougias, I., Hankins, M., Huld, T., Szabó, S., 2019. Decentralized rural electrification in Kenya: Speeding up universal energy access. Energy Sustain. Dev. 52, 128–146. https://doi.org/10.1016/j.esd.2019.07.009

Moner-Girona, M., Puig, D., Mulugetta, Y., Kougias, I., AbdulRahman, J., Szabó, S., 2018. Next generation interactive tool as a backbone for universal access to electricity. Wiley Interdiscip. Rev. Energy Environ. 7, e305.

Monyei, C.G., Adewumi, A.O., 2017. Demand Side Management potentials for mitigating energy poverty in South Africa. Energy Policy 111, 298–311.

Monyei, C.G., Oyedele, L.O., Akinade, O.O., Ajayi, A.O., Ezugwu, A.E., Akpeji, K.O., Viriri, S., Adewumi, A.O., Akinyele, D., Babatunde, O.M., 2019. An income-reflective scalable energy level transition system for low/middle income households. Sustain. Cities Soc. 45, 172–186.

Morrissey, J., 2019. Achieving universal electricity access at the lowest cost: A comparison of published model results. Energy Sustain. Dev. 53, 81–96. https://doi.org/10.1016/j.esd.2019.09.005

Morrissey, J., 2018. Linking Electrification and Productive Use.

Mueller, N.D., Gerber, J.S., Johnston, M., Ray, D.K., Ramankutty, N., Foley, J.A., 2012. Closing yield gaps through nutrient and water management. Nature 490, 254.

NAFIS – National Farmers Information Service [WWW Document], n.d. URL http://www.nafis.go.ke/ (accessed 6.23.20).

Narayan, N., Qin, Z., Popovic-Gerber, J., Diehl, J.-C., Bauer, P., Zeman, M., 2018. Stochastic load profile construction for the multi-tier framework for household electricity access using off-grid DC appliances. Energy Effic. 1–19.

Nerini, F.F., Tomei, J., To, L.S., Bisaga, I., Parikh, P., Black, M., Borrion, A., Spataru, C., Broto, V.C., Anandarajah, G., 2018. Mapping synergies and trade-offs between energy and the Sustainable Development Goals. Nat. Energy 3, 10.

Ohiare, S., 2015. Expanding electricity access to all in Nigeria: a spatial planning and cost analysis. Energy Sustain. Soc. 5, 8.

Olatomiwa, L., Blanchard, R., Mekhilef, S., Akinyele, D., 2018. Hybrid renewable energy supply for rural healthcare facilities: An approach to quality healthcare delivery. Sustain. Energy Technol. Assess. 30, 121–138. https://doi.org/10.1016/j.seta.2018.09.007

Parshall, L., Pillai, D., Mohan, S., Sanoh, A., Modi, V., 2009. National electricity planning in settings with low pre-existing grid coverage: Development of a spatial model and case study of Kenya. Energy Policy 37, 2395–2410. https://doi.org/10.1016/j.enpol.2009.01.021

Pekel, J.-F., Cottam, A., Gorelick, N., Belward, A.S., 2016. High-resolution mapping of global surface water and its long-term changes. Nature 540, 418.

Peters, J., Sievert, M., 2016. Impacts of rural electrification revisited–the African context. J. Dev. Eff. 8, 327–345.

Peters, J., Sievert, M., Toman, M.A., 2019. Rural electrification through mini-grids: Challenges ahead. Energy Policy 132, 27–31.

Poblete-Cazenave, M., Pachauri, S., 2019. Income Growth & Household Electricity Demand: Implications of Heterogeneous Elasticities.

Poverty Incidence in Kenya Declined Significantly, but Unlikely to be Eradicated by 2030 [WWW Document], n.d. . World Bank. URL https://www.worldbank.org/en/country/kenya/publication/kenya-economic-update-poverty-incidence-in-kenya-declined-significantly-but-unlikely-to-be-eradicated-by-2030 (accessed 7.9.20).

Pueyo, A., Maestre, M., 2019. Linking energy access, gender and poverty: A review of the literature on productive uses of energy. Energy Res. Soc. Sci. 53, 170–181.

Rao, N.D., Min, J., Mastrucci, A., 2019. Energy requirements for decent living in India, Brazil and South Africa. Nat. Energy 1–8. https://doi.org/10.1038/s41560-019-0497-9

Rathi, S.S., Vermaak, C., 2017. The impacts of rural electrification on labour market outcomes in developing countries: Evidence from India and South Africa. Pap. Present. ESSA 2017 Bienn. Conf. RHODES Univ. Grahamst.

Riva, F., Ahlborg, H., Hartvigsson, E., Pachauri, S., Colombo, E., 2018. Electricity access and rural development: Review of complex socio-economic dynamics and casual diagrams for more appropriate energy modelling. Energy Sustain. Dev. 43, 203–223. https://doi.org/10.1016/j.esd.2018.02.003

Riva, F., Sanvito, F.D., Tonini, F.T., Colombo, E., Colombelli, F., 2019. Modelling long-term electricity load demand for rural electrification planning, in: 2019

IEEE Milan PowerTech. Presented at the 2019 IEEE Milan PowerTech, pp. 1–6. https://doi.org/10.1109/PTC.2019.8810727
Rockstrom, J., Hatlbu, N., Owels, T.Y., Wani, S.P., 2007. Managing water in rainfed agriculture.
Sanoh, A., Parshall, L., Sarr, O.F., Kum, S., Modi, V., 2012. Local and national electricity planning in Senegal: Scenarios and policies. Energy Sustain. Dev. 16, 13–25.
Sims, B., Kienzle, J., 2017. Sustainable agricultural mechanization for smallholders: What is it and how can we implement it? Agriculture 7, 50.
Sims, B., Kienzle, J., 2016. Making Mechanization Accessible to Smallholder Farmers in Sub-Saharan Africa. Environments 3, 11. https://doi.org/10.3390/environments3020011
Sovacool, B., Vera, I., 2014. Electricity and education: The benefits, barriers, and recommendations for achieving the electrification of primary and secondary schools. U. N. Dep. Econ. Soc. Aff.
Sovacool, B.K., Ryan, S.E., 2016. The geography of energy and education: Leaders, laggards, and lessons for achieving primary and secondary school electrification. Renew. Sustain. Energy Rev. 58, 107–123.
Spalding-Fecher, R., 2005. Health benefits of electrification in developing countries: a quantitative assessment in South Africa. Energy Sustain. Dev. 9, 53–62.
Sprei, F., 2002. Characterization of power system loads in rural Uganda. Univ.
Szabo, S., Bódis, K., Huld, T., Moner-Girona, M., 2011. Energy solutions in rural Africa: mapping electrification costs of distributed solar and diesel generation versus grid extension. Environ. Res. Lett. 6, 034002.
Szabó, S., Moner-Girona, M., Kougias, I., Bailis, R., Bódis, K., 2016. Identification of advantageous electricity generation options in sub-Saharan Africa integrating existing resources. Nat. Energy 1, 16140.
Teluguntla, P., Thenkabail, P.S., Oliphant, A., Xiong, J., Gumma, M.K., Congalton, R.G., Yadav, K., Huete, A., 2018. A 30-m landsat-derived cropland extent product of Australia and China using random forest machine learning algorithm on Google Earth Engine cloud computing platform. ISPRS J. Photogramm. Remote Sens. 144, 325–340. https://doi.org/10.1016/j.isprsjprs.2018.07.017
The World Bank, 2019. World Bank Data.
Thom, C., 2000. Use of grid electricity by rural households in South Africa. Energy Sustain. Dev. 4, 36–43.
Trotter, P.A., McManus, M.C., Maconachie, R., 2017. Electricity planning and implementation in sub-Saharan Africa: A systematic review. Renew. Sustain. Energy Rev. 74, 1189–1209.
United Nations, Department of Economic and Social Affairs, Population Division, 2019. Patterns and trends in household size and composition.
Urpelainen, J., n.d. The Quest for Impact in Household Electrification: We Still Roam in the Dark | ISEP. URL https://sais-isep.org/blog/the-quest-for-impact-in-household-electrification-we-still-roam-in-the-dark (accessed 11.26.19).
USAID, 2009. MEASURE DHS STATcompiler.
Van de Walle, D.P., Ravallion, M., Mendiratta, V., Koolwal, G.B., 2013. Long-term impacts of household electrification in rural India.

van Ruijven, B.J., Schers, J., van Vuuren, D.P., 2012. Model-based scenarios for rural electrification in developing countries. Energy 38, 386–397. https://doi.org/10.1016/j.energy.2011.11.037

Weiss, D.J., Nelson, A., Gibson, H.S., Temperley, W., Peedell, S., Lieber, A., Hancher, M., Poyart, E., Belchior, S., Fullman, N., Mappin, B., Dalrymple, U., Rozier, J., Lucas, T.C.D., Howes, R.E., Tusting, L.S., Kang, S.Y., Cameron, E., Bisanzio, D., Battle, K.E., Bhatt, S., Gething, P.W., 2018. A global map of travel time to cities to assess inequalities in accessibility in 2015. Nature 553, 333–336. https://doi.org/10.1038/nature25181

World Resources Institute, 2020. A New Nexus Approach to Powering Development [WWW Document]. World Resour. Inst. URL https://www.wri.org/news/new-nexus-approach-powering-development (accessed 4.23.20).

Xenarios, S., Pavelic, P., 2013. Assessing and forecasting groundwater development costs in Sub-Saharan Africa. Water SA 39, 529–538.

Xiong, J., Thenkabail, P., Tilton, J., Gumma, M., Teluguntla, P., Oliphant, A., Congalton, R., Yadav, K., Gorelick, N., 2017. Nominal 30-m cropland extent map of continental Africa by integrating pixel-based and object-based algorithms using Sentinel-2 and Landsat-8 data on Google Earth Engine. Remote Sens. 9, 1065.

You, L., Wood-Sichra, U., Fritz, S., Guo, Z., See, L., Koo, J., 2014a. Spatial production allocation model (SPAM) 2005 v2. 0. Available Mapspam Info Accessed June 29, 2015.

You, L., Wood-Sichra, U., Fritz, S., Guo, Z., See, L., Koo, J., 2014b. Spatial production allocation model (SPAM) 2005 v2. 0. Available Mapspam Info Accessed June 29, 2015.

Zhang, L., Zhang, B., 2019. Scenario Forecasting of Residential Load Profiles. IEEE J. Sel. Areas Commun.

Supplementary Information

A- Detailed materials and methods

A1 - Population settlements clustering

Population clusters are generated based on a processing algorithm which takes a gridded population raster layer as the main input and generates polygonal shapes as outputs. Previous applications of a similar approach are provided in refs. (Arderne, 2020; *KTH-dESA/PopCluster*, 2019). The algorithm selects high population density raster pixels (with a pop. >10 inhabitants / 900 m^2 ≈ 1,110 inhabitants / km^2) and classifies them as 'core'. The core pixels are converted to polygons and buffered by a 1 km radius to unify surrounding cores. The centroids of the resulting polygons are then extracted to identify a unique core for multi-core population areas. Finally, Voronoi polygons (the boundaries of the area closer to a given centroid than to any other centroid) are generated to cover the entire regional surface and include periphery and non-urban areas (e.g. cropland) within the reference polygon for each core centroid. The methodology allows grouping populations into boundaries that are heterogeneous in size and shape while collecting a set of neighbourhooding buildings and land. The algorithm is summarised in Figure SI1.

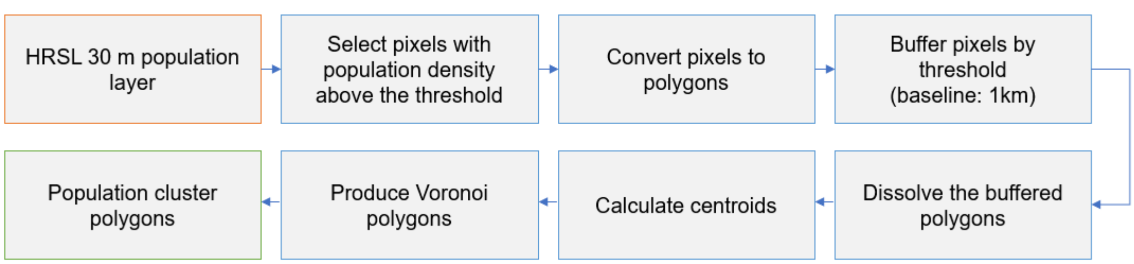

Figure SI1: *Schematic framework of the GIS algorithm to generate population clusters. The clusters represent the functional units of the GIS data processing and the demand nodes of the assessment.*

In this study, the High Resolution Settlement Layer (Facebook Connectivity Lab and Center for International Earth Science Information Network - CIESIN - Columbia University, 2016) dataset, providing population counts at a 30 m resolution based on statistical downscaling of census population based on a broad array of remotely-sensed datasets, is used. The High Resolution Settlement Layer is based on the 2015 census. Therefore, we estimate the change in the population of each grid cell from 2015 up to 2020 by applying the yearly country-level population growth rate and the share of urban population (ref. (The World Bank, 2019)). Algebraically, the union raster layer (U) of the urban and rural population's layers in year *t* is expressed by:

$$Pop_{\mathrm{t}}^{\mathrm{i}} = U(Pop_{\mathrm{t-1}}^{i\,urb}\left(1 + PGR_t^c(1 + \Delta URB_{t-1}^{t\,c})\right), Pop_{\mathrm{t-1}}^{i\,rur}\left(1 + PGR_t^c(1 + \Delta RUR_{t-1}^{t\,c})\right)) \quad \text{(Eq. SI1)}$$

where:

- $Pop_{\mathrm{t}}^{\mathrm{i}}$: the population in each cell (i) in year t
- PGR_t^c: population growth rate in country c at year t

  $\Delta URB_{t-1}^{t\,c}$ increase of share of Urban population at year t respect to previous year in country c
- $\Delta \mathrm{R}UR_{t-1}^{t\,c}$ increase of share of Urban population at year t respect to previous year in country c

The rescaling (Eq. SI1) allows to integrate the heterogeneity in the demographic change across urban and rural areas and across each country. The main limitation in this approach is that – within each country – population dynamics are homogeneous across all urban and rural areas, taking the national value of the official statistics from the World Bank (The World Bank, 2019).

Urban and rural settlements are identified at the grid-cell level using the 'degree of urbanization' method that delineates and classify settlement typologies via a logic of population size, population and built-up area densities and contiguity of the cells (A.J. et al., 2019). In our study the populations cells are classified as urban for contiguous cells with a density of 1,500 inhabitants per km^2 and a minimum of population of 5,000 inhabitants (GHS-SMOD≥2), as rural when grid cell are outside the urban clusters (GHS-SMOD≤1), or as not inhabited (GHS-POP=0). To conclude, the total population living inside each cluster is calculated with a zonal statistics algorithm (i.e. as the sum of the raster pixels falling within the polygon boundary).

A2- Residential electricity demand

Residential demand of rural and urban households in Kenya, both divided into five tiers of consumption, is computed by estimating electric appliances ownership across different tiers of consumers. The baskets of appliances are obtained through a literature review (ref. (Adeoye and Spataru, 2019; Blodgett et al., 2017; Kotikot et al., 2018; Lee et al., 2016b; Monyei et al., 2019; Monyei and Adewumi, 2017; Sprei, 2002; Thom, 2000)) supported by the authors' personal experience. The compiled database is reported in **Supplementary File F1**, where every category of users is characterized by a corresponding usage pattern of the owned appliances, differentiating every month to account for seasonality of the uses. Subsequently, the stochastic bottom-up tool RAMP (Lombardi et al., 2019) is employed to compute the load curve of each household type for each day of the year at a 1 minute time resolution. In order to avoid overlap of the peaks in this process, the simulation of the load of each tier is carried out for 100 households, taking advantage of the stochastic characteristic of RAMP, which avoids that the use of the same appliance coincides (deterministically) among users of the same category.

A3 - Allocation of population to residential consumption tiers

The next key methodological challenge requires allocating the simulated residential energy demand load curves of each household type to the population without electricity access in each cluster.

Firstly, a multi-variate random forest regression machine learning model is estimated to evaluate the current association between the distribution across tiers of households who already benefit from electricity access and their characteristics throughout sub-Saharan Africa:

$$Tiershare_i = WQ, UR, \rho_i + \epsilon_i$$

(Eq. SI2)

where:

- $Tiershare_i$ is a vector of the shares of population with access in each cluster *i* that belongs to each of the four access tiers. Information about the current distribution of households with electricity access across tiers is derived from ref. (Falchetta et al., 2019). This source provides satellite-proxied field-validated estimates of the distribution of households with electricity access across tiers;
- WQ is a vector of five variables expressing the proportion of households in each wealth quintile in the province within which each cluster falls. The sum of the five variables at each cluster is thus always 1. The wealth distribution information is derived from the most recent DHS survey data for each country (USAID, 2009);
- UR is a fractional variable expressing the share of the population in each cluster that is classified as urban; it is calculated based on the 'degree of urbanization' method (A.J. et al., 2019);
- ρ_i is a vector of country fixed-effects;

- ϵ_i is a vector of residuals.

The trained model is then used to predict the propensity of households are currently without access to electricity to fall within each of the five electricity tiers once they gain electricity access. To conclude, the predicted distribution of households without electricity across energy access tiers in each cluster are matched to the corresponding load curves and the relative power consumption levels estimated in RAMP. The approach ensures that in each cluster the latent residential electricity demand depends on the current link between electricity access tiers, wealth distribution, and urban/rural prevalence within each country. The results of the regression model are reported in Table SI3.

A4- Healthcare and education demand

In order to assess the energy behaviour of primary schools and healthcare facilities in Kenya, a field campaign was conducted in the second semester of 2019 by the authors and their team with the specific purpose of interviewing personnel from public facilities about their appliance ownership and usage patterns. During this campaign 65 Primary Schools, 10 Dispensaries (Tier 1), 14 Health Centres (Tier 2) and 3 Sub-County Hospitals (Tier 3) were visited. The purpose of the field campaign was double, collecting data directly from the facility managers to better model the electrical loads and engage with local authorities to better understand and classify the different kinds of facilities into Tiers, ad understand the national plans for the public facilities in the medium term. Thanks to the field campaign it was hence possible to collect and process the data presented in **Supplementary File F2**. The reported data are then fed into the open source tool RAMP (Lombardi et al., 2019) that thanks to a stochastic bottom-up process computes the load curve of the user per each day of a year, with a one minute time resolution.

The generated load profiles are then parsed to geospatial information about the location and characteristics of healthcare and education facilities in Kenya (refs. X,Y) with the following logic:

Healthcare

- Tier 1 -> dispensary; Tier 2 -> Health clinics; Tier 3 -> Sub-district hospital; Tier 4 -> District Hospital / Provincial General Hospital; Tier 5 -> National Referral Hospital
- Facilities with missing beds number: Tier 1 -> 0; Tier 2 -> 45; Tier 3 -> 150; Tier 4 -> 450; Tier 5 -> 2000
- Number of beds in healthcare facility *i* of tier *k* * per-bed load at tier *k*

Education

Number of pupils in school *i* * per-pupil load

A5- Micro-enterprises and commercial activities demand

In the M-LED framework, we estimate the electricity demand induced by small-scale productive and commercial activities that are widely emerging in communities of sub-Saharan Africa with an availability of electric energy, such as barber shops, minimarkets, or telecommunication points. This is carried out in three steps. First, a composite index based on the productive activities drivers and energy use is constructed based on road density (with road infrastructure data drawn from ref. (Center for International Earth Science Information Network - CIESIN - Columbia University and Information Technology Outreach Services - ITOS - University of Georgia, 2013)), employment levels and wealth distribution (at the provincial level, with data from ref. (USAID, 2009)), and city accessibility (ref. (Weiss et al., 2018)) proximity is built. The indicators are aggregated using a principal component analysis (PCA). PCA is a multivariate statistical method that is used in development research to reduce the number of variables in a dataset and construct composite indices. In a PCA, the variables are weighted according to the variance explained by the first principal component (Booysen, 2002). Figure SI2 below highlights the results of the PCA:

Figure SI2: *Results of the PCA to evaluate the propensity of micro-entrepreneurial and commercial activities to operate*

Next, the PCA outcome is rescaled to the 0.3 and 0.6 range (following (Moner-Girona et al., 2019)) to create a bottom-up mark-up factor on top of the residential demand. The baseline load curve (share of demand at each hour of the day over the total daily demand) of micro productive activities is assumed to follow the same path of that described in (Moner-Girona et al., 2019), which in turn relies on ground-metered data from mini-grids in Kenya. Finally, a seasonal variation is imposed to the monthly demand loads curves: in particular, the seasonality follows the same monthly mark-up observed in the residential demand across months of the year.

Algebraically, the final sectoral demand $CommProd_{imh}$ (where *i*, *m*, and *h*, identify demand clusters, months of the year, and hours of the day, respectively) is expressed as:

$$CommProd_{imh} = (1 + PCA_i^{range}) \times Residential_{imh} \times CommProdCurve_{mh}$$

(Eq. SI3)

where:

- $CommProd_{imh}$ is the commercial and productive demand at each cluster *i* at each month of the year *m* at each hour *h*;
- PCA_i^{range} is the result of the PCA at each cluster rescaled to the 0.3-0.6 range;
- $Residential_{imh}$ is the residential demand at each cluster *i* at each month of the year *m* at each hour *h*;
- $CommProdCurve_{mh}$ are the twelve month-specific hourly curves for the sectoral demand derived from (Moner-Girona et al., 2019) and adjusted for the seasonality based on the residential seasonality variation.

A6- Irrigation water requirements modelling

In developing countries crops are mostly rain-fed and existing water storage systems exploit gravity. For instance, in sub-Saharan Africa it is estimated that over 90% of all agricultural land is rain-fed only (Rockstrom et al., 2007). The possibility to exploit electrical energy to pump water bears a huge rural productivity growth potential – if those water resources are used sustainably (Jägermeyr et al., 2016; Mueller et al., 2012). Thus, accurately predicting those water requirements and their load curve and in turn derive the electric energy necessary to pump it can shed light on the role of pumping energy in the elaboration of a rural electrification plan that might act as a trigger to rural productivity growth.

We exploit 30-m resolution GIS information on the location of rainfed cropland in Africa (Teluguntla et al., 2018) to statistically downscale 10 km resolution information on the cropping area and regime of 42 distinct crops (You et al., 2014a). First, we use the GFSAD30AFCE cropland extent product to estimate the rainfed cropland area within each cluster. Then, using the MapSPAM database and referring to the rainfed harvested (i.e. not only the physical area, but the total area accounting for multiple harvests of a crop on the same plot) cropland area for 42 types of crops, we calculate the total area for each crop type within the clusters. Since, however, the GFSAD30AFCE product has a 30 m resolution while the MapSPAM layers only have a 10 km resolution, we redistribute the area value written into each 1 km resolution pixel such that it is proportional to the share of total cropland area within each cluster over the total cropland area underlying each MapSPAM pixel. Following this approach, we are able to downscale the layers based on the 30 m layer, under the assumption that, under each 10 km pixel, for each crop cropland is homogeneously distributed in underlying pixels. While this is an assumption, it is not particularly limiting given the already high resolution of the MapSPAM layer, which limits the maximum spatial allocation error to a ~ 500 m radius, and in any case is such that the total sum of cropland in the clusters underlying the MapSPAM pixels is equal to the value reported in the MapSPAM pixel itself.

We then combine the cropland information with satellite-derived observations of precipitations and evapotranspiration (Abatzoglou et al., 2018) and information about crop scheduling and watering periods (refer to Table SI2) to accurately estimate the daily water gap that would be necessary to ensure the optimal yield is achieved in each cluster (with the caveat that we do not consider variation in fertilisation, pesticides, or land management regimes):

$$WR_i^y = \sum_m^{12} \frac{AET_i^m - PR_i^m \, \eta \, CRSHARE_i}{\eta_c}$$

(Eq. SI4)

where:

- WR_i^y is the yearly irrigation water requirement at cluster *I;*
- AET_i^m is the total monthly actual evapotranspiration in cluster *i* calculated from the processed geospatial information on each crop's harvested area (You et al., 2014b), the relative crop factors (Allen et al., 1998) – which depend both on each specific crop and the agroclimatic zone where it is being cultivated –, and the local potential evapotranspiration (Abatzoglou et al., 2018);
- PR_i^m are the monthly cumulative precipitations;
- $CRSHARE_i$ is the share of cropland area over the total cluster area;
- η is a roots absorption efficiency parameter, set at 0.6.

The artificial irrigation water requirement is increased by dividing it by an irrigation efficiency parameter η_c. This is crop-specific, as each crop is irrigated with either drip, sprinkler, or surface irrigation, for which efficiencies of 0.85, 0.6, and 0.6, respectively, are assumed. Each crop is allocated to a technology following the FAO guidelines (Allen et al., 1998), with staple crops allocated to surface irrigation, and sprinkler and drip irrigation to vegetables and sugarcane and fruit trees, respectively. Rainfall is given an absorption efficiency of 0.6. Finally, the yearly water requirement per hectare per crop is embedded into each cluster, a weighted sum between the products of such water requirement and the rainfed harvested area of each crop in that cluster is calculated. This results in monthly and yearly water requirement in m^3

within each cluster, which is the requirement necessary to attain the potential yield in currently rainfed cropland.

A7- Water pumping energy demand quantification

To quantify the electricity necessary yearly to satisfy the estimated demand for irrigation in each cluster, we set-up a groundwater pumping model based on Eq. SI5:

$$PW_i = \frac{\rho q g h}{\eta}$$

(Eq. SI5)

where

- PW is the hydraulic power requirement in W;
- ρ is the density of the fluid in $kg \cdot m^{-3}$ (here set at 1,000, for water);
- q is flow capacity of the pump in $m^3 \cdot h^{-1}$;
- g is the gravitational constant (9.81 $m \cdot s^{-2}$) ;
- h is the differential head, in m;
- η is a pumping efficiency parameter, set at 0.75.
- h is defined by calculating the average local groundwater well depth using data from MacDonald et al. (2012) – including depth, storage, and productivity.

The flow capacity of the pump q is defined as the flow capacity necessary to satisfy the local irrigation requirements in the month t with the highest requirement assuming a maximum watering of six hours per day. To translate the pumping power requirement (W) in the daily electricity demand (kWh), the following product is estimated:

$$PWh_m = \frac{PW}{1000} \times IH_m$$

(Eq. SI6)

where:

- PWh_m is the estimated electricity consumption of the pump (in kWh) in each month *m;*
- PW is the nameplate power of the pump (in W);
- IH_m are the number of irrigation hours in month *m.*

Finally, as shown in Figure SI3, to derive cluster and month-specific load curves we consider an archetypical curve with two irrigation windows per day (5am-9am and 10pm-12am) where the pump is operating, consistent with farming practices to reduce evapotranspiration:

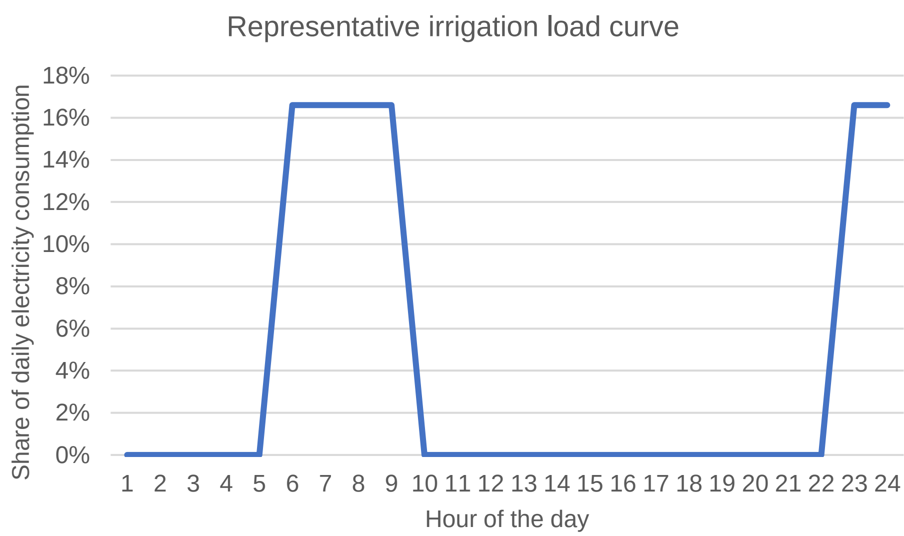

Figure SI3: *Representative load curve of the irrigation electricity demand (% of daily load at each hour of the day).*

In order to guarantee a sustainable supply of irrigation water, two constraints are set so that irrigation does not lead to the dwell's depletion.

The constraints are formulated as:

$$WW_i^d \le \frac{\left(GWProd_i * \frac{Noirrhours}{Irrhours}\right) + GWProd_i}{1000}$$

(Eq. SI7a)

and

$$WW_i^d \le \frac{GWStor_i}{Irrhours \times 3600} + \frac{GWProd_i}{1000}$$

(Eq. SI7b)

where:

- WW_i^d is water withdrawal for irrigation purposes on the day of the year *d* in cluster *I;*
- $GWProd_i$ is the average groundwater dwell productivity (in litres per second)*;*
- $GWStor_i$ is the average groundwater dwell storage (in meters)*;*
- $Irrhours$ and $Noirrhours$ are the number of hours in which the pump is operated or not, respectively, during the average irrigation day.

If the constraints are not met, the algorithm seeks to fill the watering gap withdrawing from the nearest freshwater surface (if this is within a reasonable distance threshold, set at 5 km). The surface water pumping is modelled as:

$$SFPW_i = q^i \times \frac{(32 \times WS \times SWD_i \times V)}{PD^2} \times \eta^{-1}$$

(Eq. SI8)

where:

- $SFPW_i$ is the power of the surface water pump in W
- q^i is the required water flow rate (m^3/s), obtained as the difference between the total required flow rate to meet irrigation needs and the

flow rate that can be guaranteed sustainably by the groundwater pump;

- WS is the speed of water in the pipe, set at 2 m/s;
- SWD_i is the Euclidean distance to the surface water body;
- V is the viscosity of water, 0.00089 Ns/m^2
- PD is the pipe diameter, set at 0.8 m
- η is a pumping efficiency parameter, set at 0.75.

In those instances where the irrigation demand cannot be fulfilled sustainable either by groundwater or via surface water pumping, a remark is signalled in the analysis result about the possibility to replace the existing crops or cropping schedule to relax the water stress in critical clusters. The analysis is carried out at a daily temporal resolution to account for overlapping growing seasons of crops found in each cluster and the therefore greater simultaneous water withdrawal needs.

A8- Crop processing electricity demand

To estimate the electricity necessary to mechanically process the raw crop production of each cluster, an extensive literature review of crop processing energy requirements in the context of developing countries is carried out. Refer to **Supplementary File 3** for an extensive summary of the sources accessed. Figure SI4 summarises the resulting estimates range for each crop (in kWh/kg of processed crop).

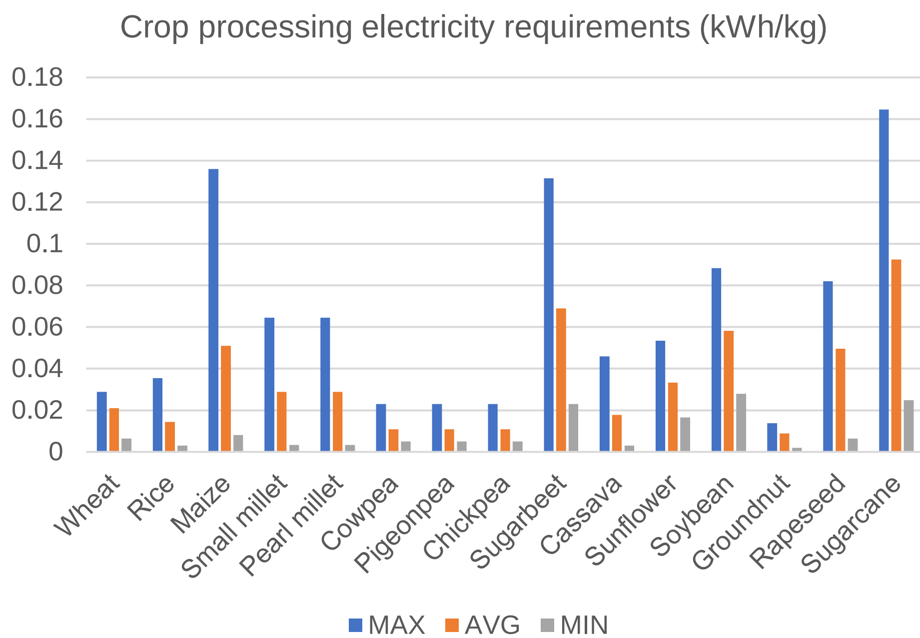

Figure SI4: *Comparison of the literature ranges for crop processing electrical requirements for the main crops considered for sub-Saharan Africa (in kWh/kg of processed crop).*

Thereafter, the yearly crop yield in each cluster *i* for each of the 42 crop classes *c* of the MapSpam database is estimated multiplying the mean crop yield (in kg/ha) of pixels falling into each cluster with the downscaled crop-specific cropland extent (in ha) of each cluster.

$$YYield_c^{im} = \overline{Yield_c^{im}} \times cropland_c^i$$

(Eq. SI9)

where:

- $YYield_c^{im}$ is the yield of each crop *c* at each month *m* in each cluster *i;*
- $\overline{Yield_c^{i}}$ is the average yield (in kg/ha) of crop c for cropland falling within each cluster *I*;

- $cropland_c^i$ is the harvested area of each crop *c* at each cluster *i* (in ha).

The total yearly electricity consumption for crop processing (*CP*) in each cluster is then calculated as:

$$CP_{im} = \sum_{i}^{N} YYield_c^{im} \times kWh/kg_c$$

(Eq. SI10)

where:

- CP_{im} is the estimated electricity consumption for crop processing (in kWh) in each month *m* at each cluster *i;*
- $YYield_c^{im}$ is the yield of each crop *c* at each month *m* in each cluster *I;*
- kWh/kg_c are the crop-specific unit processing energy requirements (in kWh).

In a similar fashion to the irrigation load curve definition, crop processing machinery follows an archetypical load (Figure SI5) with an on/off flat curve and an operation window between :

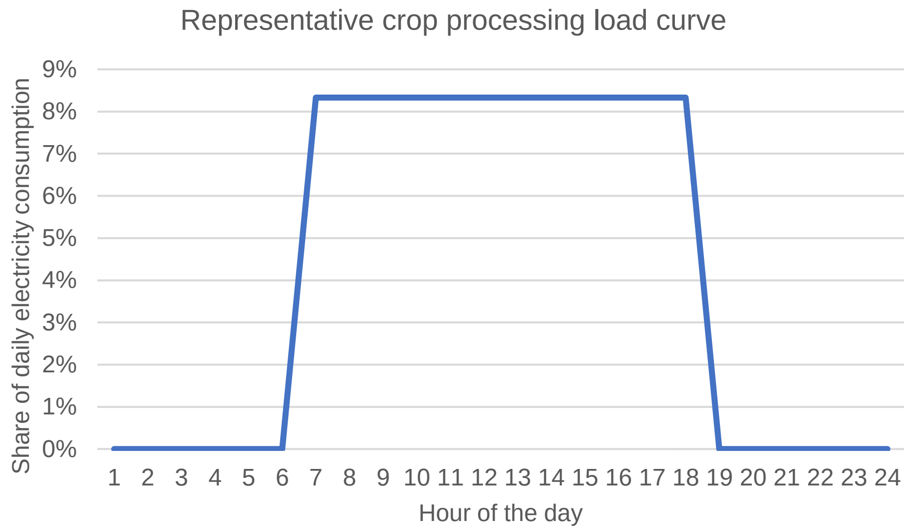

Figure SI5: *Representative load curve of the crop processing electricity demand (% of daily load at each hour of the day).*

A9- Yield gap. agricultural revenues, and costs

For the Kenya country-study, we estimate the local productivity ($kg \cdot ha^{-1}$), the mean yield for each crop is calculated using the MapSPAM rainfed crops layers. The total production in rainfed cropland is the given by the product of yield and harvested area in each cluster. To estimate the revenues stemming from the irrigation of previously rainfed cropland and the related costs, we develop a simple model of production, transportation, and wholesale. Farmers in cluster i would bear cost components F (the fixed cost for purchasing the water pump) and R (the running costs, including electricity to power the pump and operation and maintenance of the appliance), as well as T (the travel costs to the closest wholesale market, including the rent/use of the truck, the fuel, and the opportunity cost of time). In turn, they earn a revenue which is defined as the additional yield of each crop produced thanks to irrigation by the wholesale market price at which that crop is currently exchanged (according to official statistics).

Here we model transportation costs, total pumping costs, and the potential revenues from wholesale to quantify the potential locally generated agricultural revenues from the increased agricultural productivity as a result of the artificial watering. To estimate this added value, we retrieve the most recent database of wholesale prices for a large basket of crops in Kenya relative the location of each wholesale market. We then calculate what is the nearest wholesale market to each cluster, and – assuming constant prices over time – we estimate the yearly revenue.

$$Revenues_i = \sum_{j=1}^{C} P_i^j \times Yieldgap_i^{jc}$$

(Eq. SI11)

where:

- P_i^j is the local (i.e. in each cluster *i*) wholesale unit price for each crop j
- $Yieldgap_i^{jc}$ is the average difference between rainfed and irrigated yield in climate zone *c* for each crop *j.*
- To estimate and subtract transportation costs needed to generate these revenues to obtain effective profit, we calculate the following:

$$TCs_i = 2 \times (TTM_i \times Fuelcost_i \times lpermin) \times n$$

(Eq. SI12)

where:

- TTM_i is the travel time from each cluster i to the nearest market calculated in Google Earth Engine exploiting the algorithm developed by (Weiss et al., 2018)
- $Fuelcost$ is the local cost of diesel fuel derived with the approach described in (Szabo et al., 2011)
- $lpermin$ is a parameter expressing the average fuel consumption of a truck in litres per minute.

The whole product is multiplied by 2 to simulate a return journey, and by n, which is defined as the ratio of the weight of the total yield gap and the weight that a track journey can transport, thus representing the number of required journeys.

To model groundwater pumping total costs, we refer to the database for recent projects in different countries of sub-Saharan Africa compiled in ref. (Xenarios and Pavelic, 2013), selecting only mechanical electric-powered pumps. In particular, we estimate the following non-linear regression model:

$$TPC_i = h_i \times \beta_1 + y_i \times \beta_2 + h_i \times y_i \times \beta_3 + \varepsilon_i$$

(Eq. SI13)

where:

- TPC_i are total pump costs, which include both fixed upfront costs (for the installation of the pump) and operational and maintenance costs;
- h_i is the well depth (in m);
- y_i is the pump yield (in l/s).

The model yields a cost function, which is plotted in Figure SI6 for a $h_i \in (10, 50)$ and $y_i \in (1, 10)$. We then estimate total pumping costs using the model in all clusters of our analysis where groundwater pumping requirements and feasibility criteria are met .

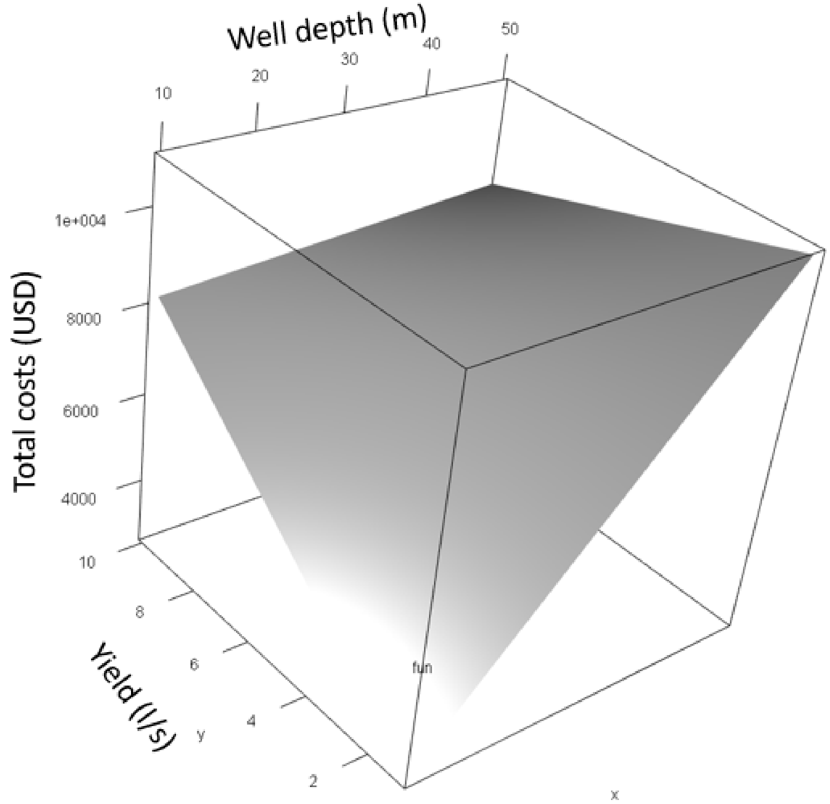

Figure SI6: 3D surface plot of the non-linear function to assess groundwater pumps total costs

A limitation of our local microeconomic economic analysis is that we do not monetise the intangible benefits of the improved local education and healthcare level as a result of the new electricity input. However, these likely imply both substantial costs savings in terms of human lives and treatment, and a greater accumulation of human capital which in the long-run can yield to significantly larger economic growth, as discussed in the relevant literature (Aguirre, 2017; Daka and Ballet, 2011; Sovacool and Ryan, 2016; Spalding-Fecher, 2005). We encourage studies targeting to quantify those indirect, long-run monetary gains. The same is true for small commercial and productive activities, and the additional value added from local crop processing. We acknowledge that electricity is likely to have a broad array of impacts through complex socio-economic linkages (Riva et al., 2018), including on fertility and migration decisions (Fried and Lagakos, 2017; Grimm et al., 2015).

Supplementary tables

Table SI1: *Main data sources in the* M-LED *platform*

Input step	**Dataset**	**Unit**	**Source (ref. number)**	**Time resolution**	**Spatial resolution**
Population clustering and residential demand	High Resolution Settlement Layer	Number of people per cell	(Facebook Connectivity Lab and Center for International Earth Science Information Network - CIESIN - Columbia University, 2016)	Annual	30 m
	Wealth distribution,	Distribution across quintiles;	(USAID, 2009)	Survey year	Province-level

	employment levels	employment rates			
	GHS-SMOD	Classification urban, rural settlement	(A.J. et al., 2019)	5 years	250 m
	VIIRS DNB nighttime lights	Radiance	(Elvidge et al., 2017)	Monthly (aggregated to annual)	30 arc-seconds
	Electricity access levels	%	(Falchetta et al., 2019)	Annual	450 m
	GADM – global administrative layers	-	(Hijmans et al., 2018)	2018	Country and provincial boundaries
Productive demand	Travel time to nearest feature	hours	(Weiss et al., 2018)	-	1 km
	Cropland extent	Land area (ha)	(Teluguntla et al., 2018)	2015	30 m
	Crop-specific harvested area and yield	Land area (km^2) Yearly yield (tonnes/(km^2)	(You et al., 2014b)	2005	10 km
	Crop processing energy demand	kWh/kg of yield processed	See Table SI3	-	-

	Crop schedule and crop factors	Days and coefficients	See Table SI2; ref. (Allen et al., 1998)	-	-
	Global Agro-Ecological Zone (GAEZ) layers	Area (ha), climate zone	(Fischer et al., 2012)	2005	0.5°
	Groundwater depth, productivity, storage	m, l/s, m	(MacDonald et al., 2012)	2012	5 km
	Surface water basins	Distance (m)	(Pekel et al., 2016)	-	30 m
Services demand	Healthcare facilities	Tier	(Maina et al., 2019)	Existing (2015) and predicted to 2030	Exact position
	Education facilities	Tier	("Kenya Open Data Initiative - Humanitarian Data Exchange," n.d.)	Existing (2015) and predicted to 2030	Count of facilities in cluster
Kenya case study	Crop wholesale prices at different markets	USD/ton	("NAFIS – National Farmers Information Service," n.d.)	1 year	Exact position

Table SI2: _Crop schedule and crop factors_

Crop	K_c 1	K_c 2	K_c 3	nd_ 1	nd_ 2	nd_ 3	nd_ 4	pm_ 1	pm_ 2	eta_ir r
Maize	0.3	1.2	0.35	30	50	60	40	1503	3010	0.6
Bean	0.15	1.15	0.35	20	30	40	20	1510	0109	0.6
Sorghum	0.3	1	0.55	20	35	45	30	2503	1510	0.6
Sweet potato	0.5	1.15	0.65	15	30	50	30	1503	0109	0.6
Tea	0.95	1	1	90	90	90	90	0101	0101	0.85
Plantain	1	1.2	1.1	120	60	180	5	0101	0101	0.85
Cowpea	0.4	1.15	0.3	20	30	35	15	1503	1510	0.85
Pigeonpea	0.7	1.05	0.95	20	30	35	15	1003	1510	0.85
Vegetables	0.7	1.05	0.95	40	60	50	15	2003	0109	0.85
Arabica coffee	1.05	1.1	1.1	90	90	90	90	0101	0101	0.85
Banana	1	1.2	1.1	120	60	180	5	0101	0101	0.85
Potato	0.5	1.15	0.75	25	30	45	30	1503	0109	0.6
Cotton	0.35	1.2	0.6	30	50	60	55	0101	0101	0.6
Cassava	0.3	1.1	0.5	150	40	110	60	0101	0101	0.85
Pearl millet	0.3	1	0.3	15	25	40	25	1503	1510	0.6

Wheat	0.5	1.15	0.33	15	30	65	40	1503	1507	0.6
Rice	1.05	1.2	0.75	30	30	80	40	0101	0101	0.6
Small millet	0.3	1	0.3	15	25	40	25	1503	1510	0.6
Sugar beet	0.35	1.2	0.7	40	70	75	35	0101	0101	0.6
Sunflower	0.4	1	0.35	25	35	45	25	1503	1510	0.85
Soybean	0.5	1.15	0.5	15	15	40	15	1503	1510	0.6
Groundnut	0.7	1.15	0.6	25	35	45	25	1503	1510	0.6
Rapeseed	0.4	1.1	0.35	25	35	45	25	0101	0101	0.6
Sugarcane	0.4	1.25	0.75	50	70	220	140	0101	0101	0.6

Table SI3: *Results of the multi-variate random forest regression for residential electricity tiers allocation*

Sample size: 297

Number of trees: 1000

Forest terminal node size: 5

Average no. of terminal nodes: 40.335

No. of variables tried at each split: 3

Total no. of variables: 8

Total no. of responses: 4

User has requested response: acc_pop_share_t1

Resampling used to grow trees: swor

Resample size used to grow trees: 188

Analysis: mRF-R

Family: regr+

Splitting rule: mv.mse *random*

Number of random split points: 10

% variance explained: 46.63

Error rate: 0.03

NOTE DI LAVORO DELLA FONDAZIONE ENI ENRICO MATTEI
Fondazione Eni Enrico Mattei Working Paper Series

Our Working Papers are available on the Internet at the following addresses:
http://www.feem.it/getpage.aspx?id=73&sez=Publications&padre=20&tab=1

NOTE DI LAVORO PUBLISHED IN 2020

1. 2020, FACTS Series, Alessandra Celani de Macedo, Nicola Cantore, Laura Barbier, Marco Matteini, Giorgia Pasqualetto, The Impact of Industrial Energy Efficiency on Economic and Social Indicators
2. 2020, 2030 Agenda Series, Nicola Comincioli, Sergio Vergalli, Effects of Carbon Tax on Electricity Price Volatility: Empirical Evidences from the Australian Market
3. 2020, 2030 Agenda Series, Marco Buso, Cesare Dosi, Michele Moretto, Do Exit Options Increase the Value-For-Money of Public-Private Partnerships?
4. 2020, FACTS Series, Ricardo Nieva, A Tragic Solution to the Collective Action Problem: Implications for Corruption, Conflict and Inequality
5. 2020, FACTS Series, Charles Fang Chin Cheng, Nicola Cantore, The Inclusive and Sustainable Development Index: a Data Envelopment Analysis Approach
6. 2020, 2030 Agenda Series, FACTS Series, Coker Eric, Cavalli Laura, Fabrizi Enrico, Guastella Gianni, Lippo Enrico, Parisi Maria Laura, Pontarollo Nicola, Rizzati Massimiliano, Varacca Alessandro, Vergalli Sergio, The Effects of Air Pollution on COVID-19 Related Mortality in Northern Italy
7. 2020, 2030 Agenda Series, Laura Cavalli, Giulia Lizzi, Port of the future : Addressing Efficiency and Sustainability at the Port of Livorno with 5G
8. 2020, FACTS Series, Federica Cappelli, Gianni Guastella, Stefano Pareglio, Institutional Fragmentation and Urbanisation in the EU Cities
9. 2020, FEP Series, Giacomo Falchetta, Nicolò Stevanato, Magda Moner-Girona, Davide Mazzoni, Emanuela Colombo and Manfred Hafner, M-LED: Multi-sectoral Latent Electricity Demand Assessment for Energy Access Planning

Fondazione Eni Enrico Mattei
Corso Magenta 63, Milano – Italia

Tel. +39 02.520.36934
Fax. +39.02.520.36946

E-mail: letter@feem.it
www.feem.it