Energy for Sustainable Development 25 (2015) 75–86

Contents lists available at ScienceDirect

Energy for Sustainable Development

Analyzing grid extension and stand-alone photovoltaic systems for the cost-effective electrification of Kenya

Marianne Zeyringer [a,b,c,*], Shonali Pachauri [d], Erwin Schmid [c], Johannes Schmidt [c], Ernst Worrell [b], Ulrich B. Morawetz [c]

[a] Institute for Energy and Transport, European Commission, Directorate-General Joint Research Centre, Petten, The Netherlands
[b] Copernicus Institute of Sustainable Development, Utrecht University, Utrecht, The Netherlands
[c] Institute for Sustainable Economic Development, University of Natural Resources and Life Sciences (BOKU), Vienna, Austria
[d] Energy Program, International Institute for Applied Systems Analysis (IIASA), Laxenburg, Austria

ARTICLE INFO

Article history:
Received 25 July 2014
Revised 20 November 2014
Accepted 31 January 2015
Available online 13 March 2015

JEL classification:
D12
C53
C61
O13
Q41
R58

Keywords:
Electrification
Sub-Saharan Africa
Optimization model
Exponential model
Tobit model
Geographically explicit

ABSTRACT

The declaration of 2014–2024 as the Decade of Sustainable Energy for All has catalyzed actions towards achieving universal electricity access. The high costs of building electric infrastructure are a major impediment to improved access, making stand-alone photovoltaic (PV) systems an attractive solution in remote areas. Here, we analyze the cost-effective electrification solution for Kenya comparing grid extension with stand-alone PV systems. We use micro-data from a national household survey to estimate electricity demand for households that are within reach of electricity infrastructure and to predict latent demand in unconnected households. These regional demands are used in a spatially explicit supply model to seek for a least cost electrification solution. Our results suggest that decentralized PV systems can make an important contribution in areas, with low demand and high connection costs. We find that up to 17% of the population can be reached cost-effectively by off-grid PV systems till 2020.

Introduction

The achievement of the United Nations Millennium Development goals is strongly associated with access to electricity. This is also reflected in a recent declaration by the United Nations General Assembly for the decade 2014–2024 as the Decade of Sustainable Energy for All (United Nations General Assembly, 2011). In 2011, 45% of the urban population and 82% of the rural population did not have access to electricity in sub-Saharan Africa. The rate of electrification in Kenya is currently below the average of sub-Saharan Africa. 81% of the households (42% in urban and 93% in rural areas) have no access to electricity in their dwellings (Organisation for Economic Co-operation and Development and International Energy Agency, 2013). Consequently a large majority of the population still relies on firewood for cooking and paraffin for lighting (Kenya National Bureau of Statistics, 2005). Change is slow, since incentives to invest in rural areas are low due to high connection costs, low latent electricity demand and low incomes.

This article identifies least-cost options for electrification of households in Kenya. Many households cannot access electricity due to non-availability of electric infrastructure and thus their demand is unknown. For cost-effective planning of electricity infrastructure, which in many developing countries involves a choice between grid extension and the implementation of stand-alone systems, it is crucial to estimate electricity demand. We use detailed micro-data from the Kenyan Integrated Household Budget Survey (KIHBS) of 2005/2006 to estimate latent demand for electricity in Kenyan households (Kenya National Bureau of Statistics, 2005). In a second step, we use the demand model to predict electricity demand for all districts of Kenya in the year 2020. The data generated serves as an input into an electricity supply optimization model determining whether electric grid extension or the implementation of off-grid photovoltaic systems is the cost-

* Corresponding author at: UCL Institute, University College London, Central House, 14 Upper Woburn Place, London WC1H 0NN, United Kingdom.
E-mail addresses: M.Zeyringer@ucl.ac.uk, marianne.zeyringer@gmail.com (M. Zeyringer).

http://dx.doi.org/10.1016/j.esd.2015.01.003
0973-0826/

effective solution for each grid cell. We consider off-grid PV as being representative for stand-alone systems, in general. The choice between grid extension or off-grid supply is a longer term decision not taken by individual households. By contrast, a national planning authority does not need to decide on the specific option(s) for off-grid supply; this can even be done at a household or community level.

This article contributes to recent scientific literature on electricity planning for countries with low electrification rates. In particular, we are interested in determining household electricity demand in areas where electricity supply is currently not available and in assessing if grid based supply or a combination of photovoltaic panels and batteries are more cost-effective in covering demand. There exists a significant body of literature on electricity demand estimation in developing countries. However, the literature on residential electricity demand in developing countries is more limited. Some examples of literature at the household level describing factors determining fuel choice are the following: Davis (1998) analyzes a household survey in South Africa to identify the effects of access to electricity of rural households on fuel choice. Masera et al. (2000) use data from a case study in a Mexican town and from a large-scale survey from four states of Mexico to find a model describing the transition from traditional to modern fuels. They show that a multiple fuel model or in other words the accumulation of energy options describes this move better than the standard energy ladder model. Tatiétsé et al. (2002) evaluate households' actual electricity energy needs in three Cameroonian cities. Their aim is to improve distribution grid planning in order to prevent frequent network interruptions and non-profitable investments. They form three classes based on criteria such as socio-professional category, income level and dwelling type. They carry out a survey collecting data on several characteristics affecting electricity consumption. From data on household appliances they calculate load profiles. Filippini and Pachauri (2004) estimate price and income elasticities of urban households' electricity demand using disaggregate household level survey data for India. The motivation of their research is to get an understanding of the key factors that influence electricity demand at the household level. Pachauri (2004) performs an econometric analysis using household survey data from India and finds that household socio-economic, demographic, geographic, family and dwelling attributes influence the total household energy demand. Ekholm et al. (2010) use a choice model to analyze the determinants of fuel consumption choices for heterogeneous household groups in India incorporating factors such as preferences. Louw et al. (2008) use sampled household data to assess the parameters affecting the electricity usage in electrified households for South Africa. Previous studies develop methodologies to explain energy consumption in developing countries, but they are very limited for sub-Saharan Africa. Moreover, none of the papers estimate currently uncovered electricity demand due to non-availability of supply. On research for electricity planning for countries with low electrification rates, until recently, there have been only a few isolated studies comparing grid connected versus stand-alone systems (Kaundinya et al., 2009; Narula et al., 2012). Studies for sub-Saharan African countries are still very limited. Zvoleff et al. (2009) determine costs of rural electrification through settlement patterns for several of the African Millennium Villages sites. Kocaman et al. (2012) propose an algorithm to minimize overall electricity infrastructure costs based on spatial distributions of demand points studying nine sites in sub-Saharan Africa. Szabó et al. (2011) compare electrification costs of distributed solar and diesel generation with grid extension for Africa through means of a cash flow model and geographic information systems. Levin and Thomas (2012) determine the percentage of population for which decentralized systems would be the cost-effective solution for 150 countries. They compare the costs of centralized generation calculated via a minimum spanning tree to decentralized generation. Fuso Nerini et al. (2014) undertake a techno-economic analysis to determine the best electrification solutions for the Amazon region in Brazil. They then evaluate those electrification options in a multi-criteria-analysis based on weights obtained from an interview process. Rosnes and Vennemo (2012) build an optimization model to determine the least cost supply options covering demand in 43 sub-Saharan countries. Deichmann et al. (2011) and Parshall et al. (2009) propose a geographically explicit methodology in order to decide between stand-alone systems and grid electrification based on cost minimization in Kenya. Kemausuor et al. (2014) uses a similar approach for Ghana. These models are strong on the engineering side, but lack a detailed analysis and modeling of the demand side: Fuso Nerini et al. (2014) use one average consumption of a 'standard household'. For new connections Rosnes and Vennemo (2012) use two average annual consumption categories, one for rural households and one for urban households. Levin and Thomas (2012) assume that newly electrified regions will consume electricity at the same rate as the currently electrified population. Deichmann et al. (2011) assume a fixed quantity of electricity demanded for two categories of households: rural and urban, whereas Parshall et al. (2009) propose four different exogenous electricity demand values depending on income and population distribution. Kemausuor et al. (2014) also use four average demand categories.

The literature review shows that there is no study combining thorough demand estimations with a supply side electrification model. Furthermore, from the literature and to our knowledge our approach of estimating latent electricity demand due to non-availability of electricity supply is novel. Similar econometric approaches to latent demand estimation have been applied in other sectors such as Briand et al. (2010) for water in households but not to estimate electricity demand. With this research we are attempting to close the gap of predicting electricity demand at a district level and using it in a supply-side cost optimization model to choose between grid extension and stand-alone PV systems.

The article is structured as follows. The Data and methodology section provides data and methodology for the demand estimation and the supply optimization model. In the Results section, we present the results of both models. The Discussion and conclusions section provides policy conclusions as well as an outlook on future research.

Data and methodology

Kenya with a geographical area of 569,250 km^2 is located in Eastern Africa on the equator. The country is divided into eight provinces and 46 districts. The Kenyan Bureau of Statistics states a provisional number of 40.7 million inhabitants for 2012 (Kenyan Bureau of Statistics, 2013). According to the KIHBS, the average population density varies depending on the district between 2.5 inhabitants per km^2 and 4500 inhabitants per km^2. One out of five Kenyans lives in urban areas (Kenya National Bureau of Statistics, 2007). The public Kenyan Power and Lighting Company (KPLC) is responsible for transmission, distribution and retail of electricity. About 80% of national electricity is generated by the state owned Kenyan Electricity Generating Company (KenGen) (Kenya Electricity Generating Company, 2014). KenGen sells the electricity to the Kenyan Power and Lighting Company (KPLC) (Kenya Electricity Generating Company, 2014). The total installed power generation capacity amounts to 770 MW of hydropower, 610 MW of thermal energy, 200 MW of geothermal energy and 26 MW of cogeneration (provisional numbers for 2012) (Kenya National Bureau of Statistics, 2014). In the Rural Electrification Master Plan (REM), the government aims at an electrification rate of 33% until 2018 and 40% until 2020 (Ministry of Energy et al., 2009).

Fig. 1 gives an overview of the data and methodology that are described in more detail in the Electricity demand estimation of Kenyan households and Electricity supply optimization model sections. We use an exponential regression model in order to predict electricity demand for households without access to electricity. The predicted electricity demand in every grid cell (2000 km^2) serves as an input parameter in a supply-side optimization model, which determines the

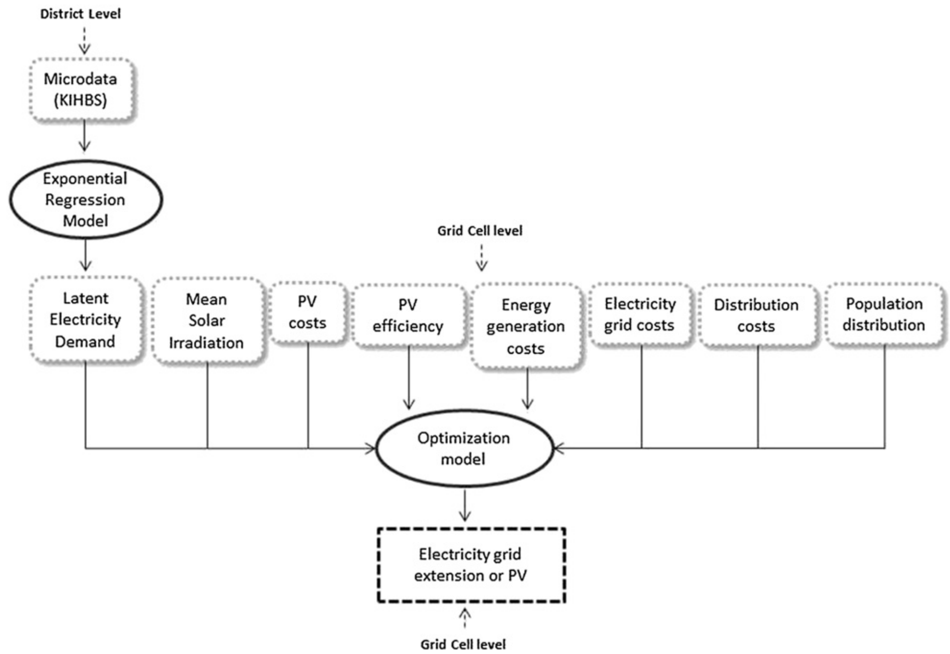

Fig. 1. Flowchart of the input data and methodology.

extension of the electricity grid and the introduction of stand-alone PV by minimizing costs in every grid cell.

Electricity demand estimation of Kenyan households

We use the Kenyan Integrated Household budget survey (KIHBS) data (Kenya National Bureau of Statistics, 2005), which is the best nationally representative data available to study the current patterns of household electricity demand. There are some caveats with the use of the data, but the survey data was systematically examined, manually corrected for outliers, and checked for logical inconsistencies wherever possible to eliminate measurements errors. We dropped one observation where electricity demand was not available, two observations where electricity consumption was zero, but which reported electricity expenses, and three observations where the head of the household was younger than 16 years. The same dataset was previously used by Lay et al. (2013) to study the determinants of Kenyan households' choices of lighting fuels. The stratified sample consists of 13,340 households surveyed between 2005 and 2006 and contains the sampling selection probabilities for each household. One of the 21 modules of the questionnaire is designed to give information on household energy use. For our descriptive analysis, we distinguish households by rural and urban expenditure quintiles and assess differences in electricity access and demand among these groups.

Early studies on electricity demand estimation in western countries primarily focused on estimating electricity from equipment stocks (Berndt, 1991). The situation in developing countries differs, because many households are not connected to the grid and consequently do not own electric appliances. Several approaches are possible to estimate latent electricity demand of households which may get connected to the grid. From a statistical point of view, a randomized experiment or matching would be the most desirable. However, this requires data before and after the introduction of electricity supply (Rosenbaum, 2010). From an economic point of view, a full household model including labor and goods market is also appealing (Bardhan and Udry, 1999). However, data limitations and underidentification of such a system of equations often require strong assumptions, which are hard to justify (Angrist and Pischke, 2010). We take a microeconometric approach, for estimating demand, which is suited for cases where a large share of households has zero demand (Wooldridge, 2010).

The Kenyan household survey data contains information that allows us to distinguish areas where the electricity grid is available within 100 m of a household to other areas where electricity infrastructure is missing. When estimating electricity demand we run the regression only using data for households within 100 m of an electricity supply. The rationale is that only these households potentially have access to electricity, i.e. they can choose to use electricity from the grid. This is important information in the regression model to estimate which households actually do consume electricity, depending on their household characteristics. There are a large proportion of households that do not consume electricity although they are within 100 m of the electric grid. Including households that don't have the option to consume electricity at all, as they are beyond this distance to the grid, would change our results significantly. Moreover, the results would be incorrect as these households do not have the option to consume grid electricity, independent of their household characteristics, as they are not able to connect to the grid. The two sub-sets of households differ significantly with respect to expenditures, servants, flush toilets, age of household head, formal education, number of people living in the house, and the share of rural households. We control for these variables in the regression model. We exclude the main cities Nairobi and Mombasa from our econometric estimation. These being the two major Kenyan cities, with higher electricity supply, we assume that households located in these two cities are structurally too different to be used to predict electricity demand for households outside of them. These are also the only two districts of Kenya which are entirely urban and thus compared to the rest of the country have very low rates of people employed in agriculture (1.3% and 1.1% of households in Mombasa and Nairobi compared to 68.8% of all households in Kenya). The industrial and commercial activities are consequently higher and the main economic sector in these cities is wholesale/retail/trade (Kenya National Bureau of Statistics, 2007; Knight Frank, 2014). Education is more accessible proven by the high percentage of population being able to read and write as well as the number of children attending school and the low distance to the nearest public primary school (Kenya National Bureau of Statistics, 2007). Nairobi and Mombasa attract a high proportion of national and international, wealthy, high-income inhabitants and tourism is strong (Kenya's Ministry of East African Affairs, Commerce and Tourism, Department of Tourism, 2010). House prices are much higher and in the high segment comparable to other global cities (Knight Frank, 2014). Nairobi and Mombasa, differently to the rest of the country have a higher proportion of households renting, thus spending a large amount of their non-food expenditures on rent (Kenya National Bureau of Statistics, 2007). Infrastructure such as pipes for drinking water provision, sanitation and flush toilets, electricity, roads, public transport (buses, railway), healthcare is much better developed (Kenya National Bureau of Statistics, 2007; The World Bank, 2014).

Electricity demand of connected households mainly depends on, according to microeconomic theory, income, prices, and preferences of the household (Louw et al., 2008). Household's preferences for electricity differ depending on location (urban or rural) and the number and structure of household members (e.g. family member size, age, education). Income is notoriously difficult to measure in developing countries and expenditure is often preferred (Deaton, 1995). We use non-food expenditures and the number of domestic servants employed in the household as a proxy for income. Non-food expenditures are transformed applying an inverse hyperbolic sine (ihs) transformation $\mathrm{ihs}(y_i) = \log(y_i + (y_i^2 + 1)^{0.5})$, instead of a simple logarithmic transformation since $\log(0) = -\inf$. We exclude food expenditures as some households grow their own food. The existence of a flush toilet can serve as indication of income and infrastructure availability and is therefore included in the model. Energy prices are not included in the model as they do not vary across regions (for electricity). We also exclude prices of substitutes and complements of electricity as no region specific data are available. In describing household preferences, we include age and education of the household head, the number of people living in the household, and whether the household lives in an urban or rural area (representing all characteristic being typical for rural households but can not be measured). We do not include regional dummies, because we have only a dozen observations in some of the provinces.

Generally, households have a non-negative electricity demand. Since a standard ordinary least square model can result in negative fitted values, Tobit models are frequently used in situations with limited outcomes (Tobin, 1958). In the Tobit model, under the assumption of normally distributed and homogeneous errors, censoring at zero demand is accounted for. The assumptions of normally distributed and homogeneous errors are often not fulfilled in the Tobit model. To model multiple step household decisions, the standard Tobit model has been extended to a class of multi-hurdle models (Wooldridge, 2010). Cragg's model (Cragg, 1971) is the most basic one which separates the consumption decision in two steps. The two estimation steps can be correlated in more recent variants (Blundell and Meghir, 1987). While these models are theoretically appealing, they rest on assumptions of normal and homogenous errors. Multi-hurdle models consist of several, possibly correlated equations such that non-convergence of the objective function is a frequent problem. The censored least absolute deviation method by Powell (1984) is used to avoid the strict distributional assumption of hurdle models. However, they often suffer from non-convergence in case of a large share of censored observations. A more direct approach is the prior transformation of the variables to fulfill the assumption of non-negative outcomes. Wooldridge (2010) suggests an exponential model of the type $y_i = \exp(bX_i) + u_i$. This exponential model can be estimated with a non-linear regression routine which results in strictly positive expected values. We estimate both an exponential and a Tobit model to allow for a comparison of the results.

The estimated equation for the Tobit model can formally be expressed as

$$Ed_i = \begin{cases} Ed_i^* & \text{if } Ed_i^* > 0 \\ 0 & \text{if } Ed_i^* \le 0 \end{cases}$$

where

$$Ed_i^* = b_0 + b_1 Ex_i + b_2 S_i + b_3 F_i + b_4 A_i + b_5 E_i + b_6 N_i + b_7 R_i + u_i. \tag{1}$$

The estimated equation for the exponential model can formally be expressed as

$$Ed_i = \exp(b_0 + b_1 Ex_i + b_2 S_i + b_3 F_i + b_4 A_i + b_5 E_i + b_6 N_i + b_7 R_i) + u_i. \tag{2}$$

The variable abbreviations in both models represent:

Ed_i^*	Latent monthly electricity demand per household in kW h
Ed_i	Monthly electricity demand per household in kW h
Ex_i	Non-food expenditures per household in KSh per month
S_i	Number of servants employed in the household
F_i	Flush toilet as main toilet facility
A_i	Age of household head in years
E_i	Formal education of the household head in years
N_i	Number of people living in the household
R_i	Dummy equal 1 if household is in a rural area of household i
$b_0, \ldots, b_7$	Regression coefficients

After estimation of the coefficients, we use the exponential equation to predict future demand. The survey data are from 2005/06, but investment decisions have to be based on future demand. We therefore predict demand in 2020, which serves as our projection horizon. We inform our estimation of future demands in 2020 by employing GDP (rural, urban), population (rural, urban), and share of educated population (over 15 years of age) projections from the International Institute for Applied Systems Analysis (IIASA) (K.C. et al., 2010; Riahi et al., 2012). While similar projections are available from other sources, such as the World Bank and United Nations, projections for all the variables of interest are not available from a single source, in a consistent manner. The data projected for the year 2020 can be found in Table 1. We base the predictions for 2020 on non-food expenditures from 2005/06 multiplied by the change in GDP (differentiated between urban and rural) and the education level of 2005/06 multiplied by the change in education. We then predict demand based on these variables and multiply the predictions by the population growth (differentiated between urban and rural). As the KIHBS provides the location of each household at the district level, the projections result in mean demand per district. Knowing in which district each grid cell is located, and multiplying the number of households in each grid cell with the mean demand allows us to estimate the total electricity demand in each grid cell.

The strength of the model we use lies in utilizing the distinction between those who could consume electricity (because they lie within 100 m of an electric connection), but do not to estimate the latent demand for electricity. This is an important differentiation, which has not been considered previously. Other models, like equipment stock based models or Almost Ideal Demand Models are more appropriate in other situation (if the households are connected already) or if other variables are available (e.g. region specific price data). For the purpose of estimating demand as input for the optimization model and the data available (one year cross-section survey and national population and economic forecasts) we consider the model as the best choice.

Table 1
Data for 2005 and projections for 2020.

	Unit	2005	2020
GDP	Billion USD	18.769	32.324
GDP urban	Billion USD	12.361	22.607
GDP rural	Billion USD	6.408	9.717
Education level of the population older than 15 years	Years	7.76	9.43
Population	Million	35.817	52.034
Population urban	Million	7.429	13.826
Population rural	Million	28.387	38.208

The exponential model is estimated with the function "nls" from the "stats" package (version 2.14.2) and the Tobit model with the function "mhurdle" (Carlevaro et al., 2011) of the statistical programming environment R (R Development Core Team, 2012). Weighted means are calculated with the R package "weights", version 0.70 (Pasek, 2011).

Electricity supply optimization model

In the optimization model we find the cost optimal electrification solution between stand-alone PV and grid extension for every grid cell of 2000 km^2. The aim of the model is to find the cheapest solution to cover demand for those households that currently cannot access electricity due to non-availability of electric infrastructure. One of the constraints of the model is that supply needs to satisfy all demand. The model does not endogenously decide which areas are more important to electrify. The model only chooses between on-grid and off-grid stand-alone PV as a solution to satisfy electricity demand for each grid cell based on cost alone. We assume that all regions are given the same priority in satisfying the latent demand of households.

Initially, we aimed at including all options which are currently being employed in Kenya for household electrification in the analysis. In addition to the grid and solar options included in the model, diesel generation is the only other option that is used widely. Compared to other technology options (e.g. micro-hydro and wind) for which reliable data is lacking (Szabó et al., 2011); PV and diesel generation costs are well understood and the technologies can be implemented in any part of the country (Parshall et al., 2009). Spatially differentiated costs of diesel are unavailable for Kenya. To acquire meaningful results good data on spatially differentiated diesel cost (depending on the transport distance) would have been necessary as fuel consumption represents the major share of costs for these systems (Szabó et al., 2011). A test run confirmed little difference in the results for PV and diesel options. This conclusion is in line with other studies for Kenya on PV economics by Ondraczek (2014) and Szabó et al. (2011). We use off-grid PV as being representative for stand-alone options in the model. The choice between grid extension and stand-alone systems for the electrification of a grid cell requires a long-term plan, whereas the choice between PV and another stand-alone option can be taken by the individual household or community. An estimation of the relative cost competitiveness of diesel and PV stand-alone systems for a certain set of assumptions on costs and interest rate is included in the Appendix A.2 to illustrate the relative competitiveness between these two off-grid options in Kenya.

To illustrate the assumed and prevailing conditions, Fig. 2 shows a map of Kenya including the grid cells, the current electricity grid and administrative boundaries. The key data for the model are costs of the technologies and data on variables which influence costs per kW h such as population distribution, solar irradiation and PV efficiency.

Table 2 *shows the input parameters together with their values and sources. A more detailed description can be found in the* Appendix A.1.

Indices:

$i =$ supply cells

$j =$ demand cells
alias: $j = j^1 = j^2$

The investments are annualized by applying an interest rate of 6% to initial investment costs. PV prices are expected to stay stable in the future (Bazilian et al., 2013). We thus take prices from the IEA-ETSAP technology brief of 2013 (IEA-ETSAP and IRENA, 2013) for the 2020 scenario. Module prices fell to USD 950 per kW (IEA-ETSAP and IRENA, 2013). This results in costs of USD 119/per m^2 assuming that 1 kWp requires 8 m^2 of rooftop area. Cost projections for batteries are scarce or unavailable (IEA-ETSAP and IRENA, 2012). We thus use the same costs from the baseline scenario for 2020. According to the Kenyan Ministry

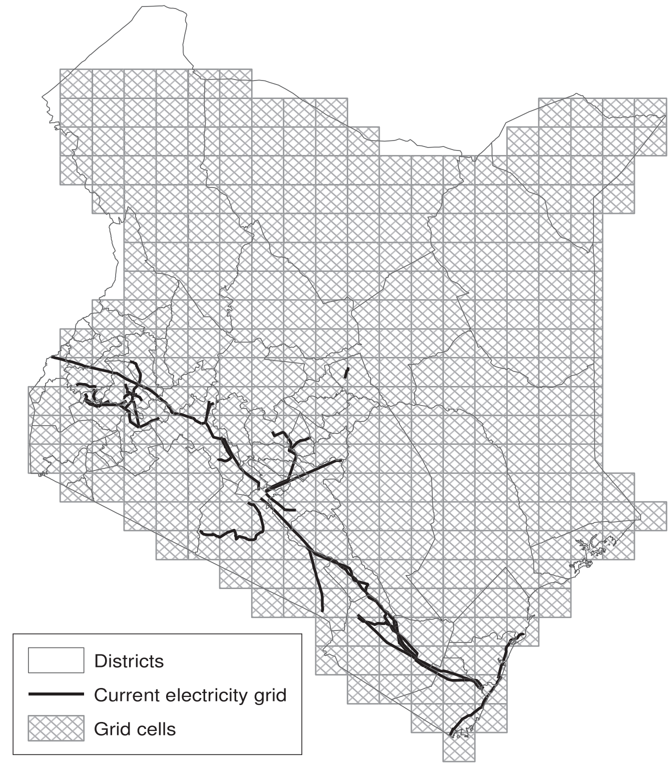

Fig. 2. Map of Kenya showing the grid cells, the districts, and the existing electricity grid.

of Energy future energy generation costs will stabilize at around USD 0.17/kW h in 2018 (Ministry of Energy et al., 2009).

The following paragraphs describe the optimization model in more detail:

Positive variables:

$x_{ij} =$ amount of grid electricity transported from i to j, where $x_{ij} \geq 0$ for all i, j (*kW h per year*)

$s_j =$ size of solar panel areawhere $s_j \geq 0$ for all j (m^2)
$u_j =$ grid electricity used in grid cell where $u_j \geq 0$ for all j (*kW h per year*).

Binary variable:

$w_{i,j} =$ investment in power transmission line from i to j where $w_{i,j} = 0, 1$.

The optimization model minimizes the following objective function:

$$Min \sum_j \left(c^e + c_j\right) * u_j + \sum_{i,j} c^t * d_{i,j} * w_{i,j} + \sum c^s * s_j. \quad (3)$$

Total costs are composed of (i) the sum of the grid electricity price c^e and the distribution charge c_j multiplied with the amount of grid electricity u_j consumed in a certain grid cell j, (ii) the costs for building the transmission grid between different cells which are determined by grid construction costs per kilometer c^t, the distance between grid cells $d_{i,j}$ and the binary variable $w_{i,j}$ indicating if a certain grid connection is built, and (iii) the costs of a solar panel c^s multiplied by the solar panel area s_j. Total costs are thus composed of electricity grid distribution costs, electricity grid transportation costs and solar panel costs.

Table 2
Parameters of the optimization model: values and sources.

Model parameters	Value	Source
$d_{i,j}$ = distance between cells i and cells j (meters)	Depending on the cells	Own calculations using ArcGIS
p_j = maximum solar panel area in cells j (m^2)	2000 km^2	Own assumption
b_j = demand for electricity in cells j (kW h per year)	Per household on average	Own model using micro-data of the KIHBS Kenya National Bureau of Statistics (2005)
c^t = costs of building of power transmission lines (USD per meter)	USD 157,470 km^{-1} on average	Kenyan Power and Lighting Company (2010) (Appendix A1.1)
c_j = distribution charge in cells j (USD per kW h)	Reference costs: USD 0.1027 kW h^{-1} Distribution charge: 1% increase in population density leads to 0.6% decrease in distribution costs	Reference costs: Ministry of Energy et al. (2009) Distribution charge: Filippini et al. (2004) (Appendix A1.1)
g = grid capacity (kW)	733 GW h	Own calculation based on 5 times maximum demand of grid cells
a_i = supply of conventional electricity in cells i (kW h)	9999 GW h	Own assumption
c^e = energy generation costs (USD per kW h)	USD 0.13 kW h^{-1}	Ministry of Energy et al. (2009) (Appendix A1.2)
c^s = solar power installation costs in cells i (USD per m^2)	USD 250 m^{-2} of modules USD 250 for a 60 watt battery USD 100 for additional equipment	Wholesale Solar, n.d.; Sollatek Electronics Kenya Limited, n.d.; Center for Alternative Technologies, n.d.; Solarlink, n.d. (Appendix A1.3)
i_j = solar irradiation in cells j (kW h/m^2/a)	The solar irradiation varies spatially and thus the yearly output is different for each grid cell. Depending on the cell, 2007 kW h m^{-2} and year on average	Data from Rigollier et al. (2004), grid cell output calculated by the authors (Appendix A1.4)
c^f = solar efficiency (%)	12.8%	Own calculation based on $E_{electric}$[a] = H_{solar}[b] * $f_{inclination}$[c] * $Fs_{olarcells}$[d] * I_{cell}[e] * PR[f] (Appendix A1.5)
Population Density — determines total demand in the grid cell	45 per km^{-2} on average	Center for International Earth Science Information Network (CIESIN), Columbia University et al. (2004) (Appendix A1.6)

[a] Produced energy per year in kW h/year.
[b] Mean sum of the yearly radiation energy in kW h/year.
[c] Inclination factor.
[d] Area of the solar cells.
[e] Module efficiency.
[f] Performance ratio.

The following constraints have been defined in the model: Demand for electricity has to be fulfilled in all grid cells which is ensured by

$$\sum_i u_j + c^f * \ i_j * \ s_j \geq b_j, \forall \ j \tag{4}$$

requiring that the amount of grid electricity u_i and the amount of locally available PV electricity $c^f * i_j * s_j$ is greater than demand b_j.

The grid electricity used in a grid cell u_i and the grid electricity transferred to other grid cells $\sum_i x_{i,j}$ has to be lower than the amount of grid electricity produced in the grid cell a_i and the amount of grid electricity transferred to the grid cell from other cells $\sum_h x_{h,i}$. PV electricity cannot be transferred to other grid cells.

$$u_i + \sum_i x_{i,j} \leq + \ \ a_i + \sum_h x_{h,i}, \forall \ i \tag{5}$$

The following equation ensures that a transmission line $w_{i,j}$ of capacity g is in place if electricity is transferred from one grid cell to another:

$$x_{i,j} \leq w_{i,j} * g \ \forall \ \mathrm{i}, \ \ \mathrm{j}. \tag{6}$$

Finally, the size of the solar panel area that may be deployed is constrained by the maximum solar panel area:

$$s_j \leq p_j. \tag{7}$$

The model is implemented in the General Algebraic Modeling System (GAMS) (GAMS Development Corporation, 2009) using the solver CPLEX.

Results

Electricity demand estimation

Patterns of household electricity use

We briefly describe consumption patterns using data from the Kenyan Integrated Household Budget Survey (KIHBS). From the survey it is evident that there are households which are connected to an electric grid, but do not consume any electricity. According to the KIHBS, 17.7% of the total households have access to electricity in their household; 43.8% in urban and 3.5% in rural areas. However, only 7.2% of households report consumption greater than zero, 21.7% in urban areas and 2.3% in rural areas. Furthermore, 35% of households are reported to live within 100 m of an electricity grid. Contrary to these figures, in 2005/2006 the Kenyan Power and Lighting Company (KPLC) had 660,200 household costumers, or 9% of households were connected to the electricity grid. In the KIHBS, 10% of the households do not name KPLC as their main supplier of electricity but other systems such as community generators, solar panels, own generators or a car/motorcycle battery. This explains only a small part of the divergence in access estimates between the KPLC and KIHBS. The rest of these households might be unofficially connected through their neighbors to the national grid. The fact that the consumed amount of electricity stated by KPLC is higher than the consumed amount estimated from people having a positive electricity bill, as derived from the KIHBS, strengthens this argument: According to KPLC, total sales to domestic customers in 2005/2006 amounted to 1028 GW h. Weighted sum from the KIHBS data (2005/2006) results in an electricity demand estimate of 806 GW h for all of Kenya. This means that according to the KIHBS estimate household electricity demand for Kenya is 222 GW h lower than the official numbers given by the KPLC. When excluding Nairobi (422 GW h), the difference in household electricity demand between the KIHBS and the KPLC is lower; the KIHBS shows a demand which is 22 GW h higher than the numbers given by the KPLC (as KPLC does not report Mombasa data separately,

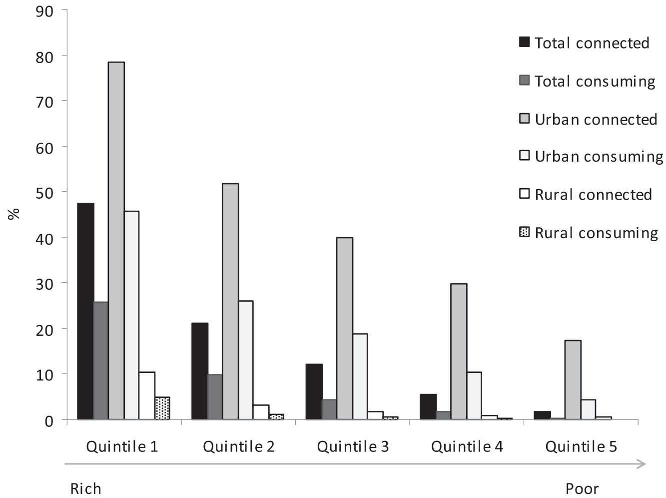

Fig. 3. Percentage of electrified households and households consuming electricity.

we deviate from other parts of the paper and only excluded Nairobi here). Fig. 3 illustrates the estimated number of households with access to electricity for each income quintile using KIHBS data. The figure also shows the difference between the percentage of households that state they have an electricity connection in the household and the percentage of households with electricity consumption (meaning electricity use greater than zero). The difference can be explained by households having a connection but not being able to afford to consume during the previous month, power outages or households which use electricity but are not billed as they are not officially connected. In the richest quintile, nearly half of all households are connected to electricity, compared to close to 0% among poor households. The percentage of connected households largely exceeds the percentage of households consuming. Further, the rate of electrification in all quintiles is much higher in urban areas. Only about 50% of the electrified households also consume electricity.

Econometric model

The mean value of electricity demand amounts to 9.63 kw H per month in our sample. A list of the key independent variables included in the model as described by Eq. (1) and (2) are included in Table 3. Table 3 also shows the difference in the sample-weighted (according to KIHBS weights) mean for the regression variables between households located within and outside 100 m of an electricity supply. The differences in the means across the two population sub-groups are significant for the variables included in the model.

Table 4 shows the results from the weighted exponential and Tobit regression. Out of the 4084 observations, 152 have been deleted due to missing explanatory variables. Common goodness of fit measures are not applicable for non-linear regressions models. For the Tobit model the pseudo (or McFadden) R^2 is 0.55. In both models, all household characteristics used to explain electricity demand are significant at the 1% level.

The coefficient of an exponential model can be interpreted as the relative change in mean electricity demand when the explanatory variable changes by one unit. The coefficient of the inverse hyperbolic sine transformed variable can be interpreted as elasticity. The elasticity of non-food expenditures is 0.57, i.e. households with 1% higher expenditures for non-food goods have 0.57% higher expenditures for electricity. The second proxy for income, the number of servants, is positive and attributes a 44% higher electricity demand to households with a servant. Similarly, households with a flushing toilet have 39% higher electricity demand. The characteristics of the household head influences the electricity demand: if the head is one year older, electricity demand is on average 2% higher and per additional year of formal education demand

Table 3
Weighted mean regression variables for households within and beyond 100 m of an electricity supply.

	Weighted mean of household		
	Within 100 m of electricity supply	Beyond 100 m of electricity supply	
Non-food expenditures per household in KSh per month	1526.63	517.45	**
Number of servants employed in the household	0.13	0.07	**
Flush toilet as main toilet facility in %	14.83	0.67	**
Age of household head in years	42.18	47.66	**
Formal education of household head in years	8.48	5.43	**
Number of people living in the household	4.33	5.61	**
Households in rural area in %	58.17	96.44	**

Note: ** indicates difference in means significant on a 5% level.

Table 4
Results of the exponential regression and the Tobit regression.

Explained variable: electricity demand (kW h per month)	Estimated coefficient	Standard error	Level of significance	Estimated coefficient	Standard error	Level of significance	Mean marginal effects
	Exponential model			Tobit model			
Intercept	−3.21	0.24	***	−884.97	47.18	***	−189.40
Inverse hyperbolic sine non-food expenditures per person in KSh per month	0.57	0.02	***	58.63	5.20	***	12.68
Number of servants employed in the household	0.45	0.025	***	35.28	8.79	***	7.08
Flush toilet as main toilet facility	0.39	0.06	***	123.85	10.81	***	27.47
Age of household head in years	0.021	0.00	***	2.36	0.37	***	0.54
Formal education of household head in years	0.03	0.01	***	8.46	1.20	***	1.78
Number of people living in the household	0.10	0.01	***	14.77	1.96	***	3.06
Household in rural area (yes = 1)	−0.63	0.09	***	−69.39	12.31	***	−11.32
Sigma	NA	NA		190.46	5.08	***	
Observations	4084			4084			
Pseudo R^2				0.55			

Note: *,**, and*** indicate significance on a 10, 5, and 1% significance level.

increases by 3%. An additional member in the household increases electricity demand by 10%. The mean electricity demand in rural households is estimated to be 63% lower than the electricity demand of urban households. The coefficients of the Tobit model indicate the marginal effects of the latent variable. The marginal effect can be calculated either for connected or for all households. In this context, we report the marginal effect of the actual electricity demand, i.e. the electricity demand of connected households (Long, 1997). Households with 1% higher expenditures for non-food goods have on average 0.12 kW h higher electricity demand. Households which employ an additional servant have an electricity demand 7.1 kW h higher and those which possess a flushing toilet 27.5 kW h higher, on average. Our income proxies thus substantially influence electricity demand. The characteristics of the household's head also influences the electricity demand: an additional year of age is correlated with an increased demand of on average 0.5 kW h and an additional year of formal education by 1.8 kW h. Finally, an additional household member increases demand by 3.06 kW h. The mean marginal effect of the dummy variable for rural households implies that rural households have a 11.3 kW h lower electricity demand than urban households on average.

Fig. 4 shows the predicted mean household demand from the regression, the observed weighted mean demand for all electrified households as well as all households from the KIHBS. The observed mean electricity demand from KIHBS data is distributed among the districts, ranging between 0.1 kW h and 65 kW h per household and month. The observed mean demand for electrified households ranges between 17 kW h and 342 kW h. The predicted mean demand calculated from the regression model ranges between 14 kW h and 156 kW h per

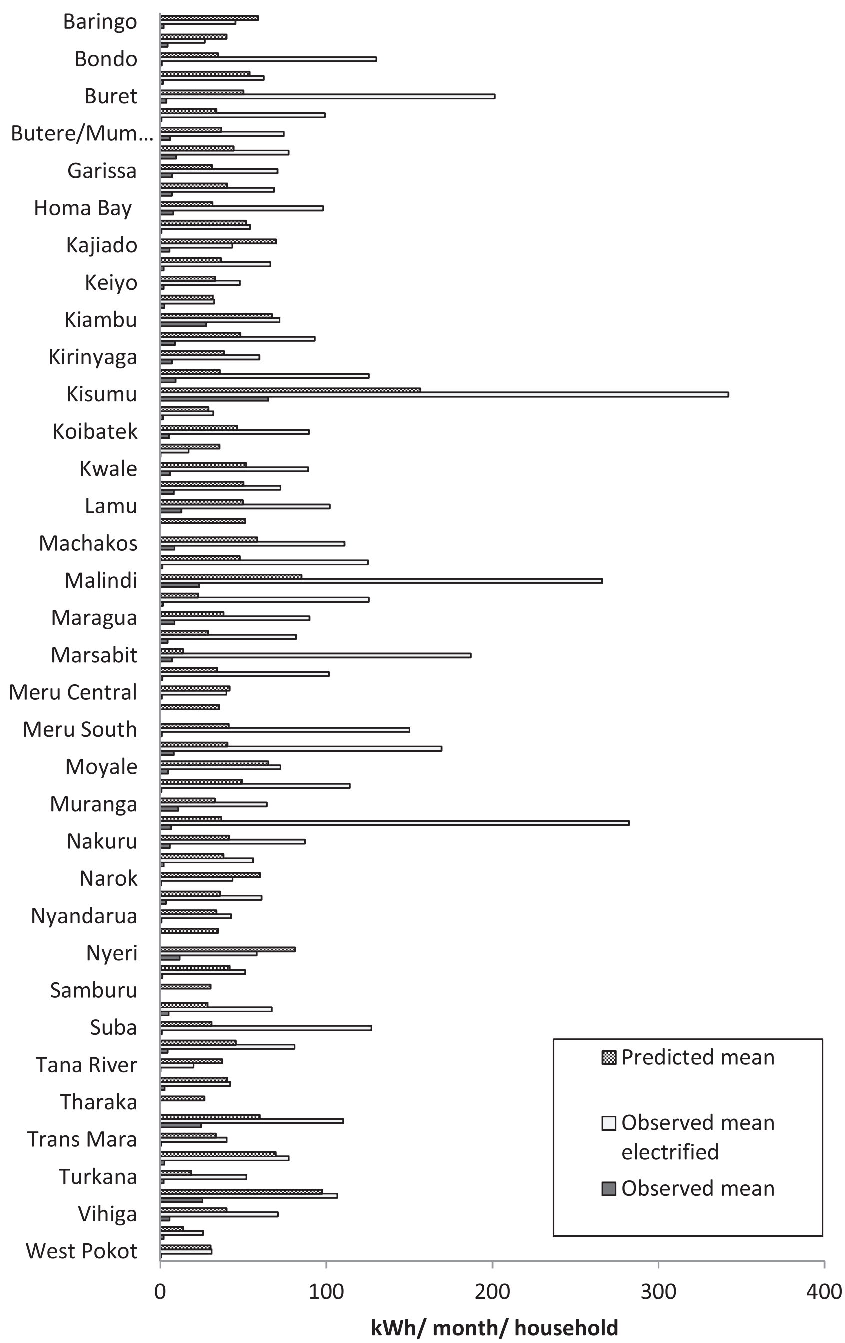

Fig. 4. Predicted mean household electricity demand, observed mean household electricity demand for electrified household and observed mean household electricity demand for all households for each district.

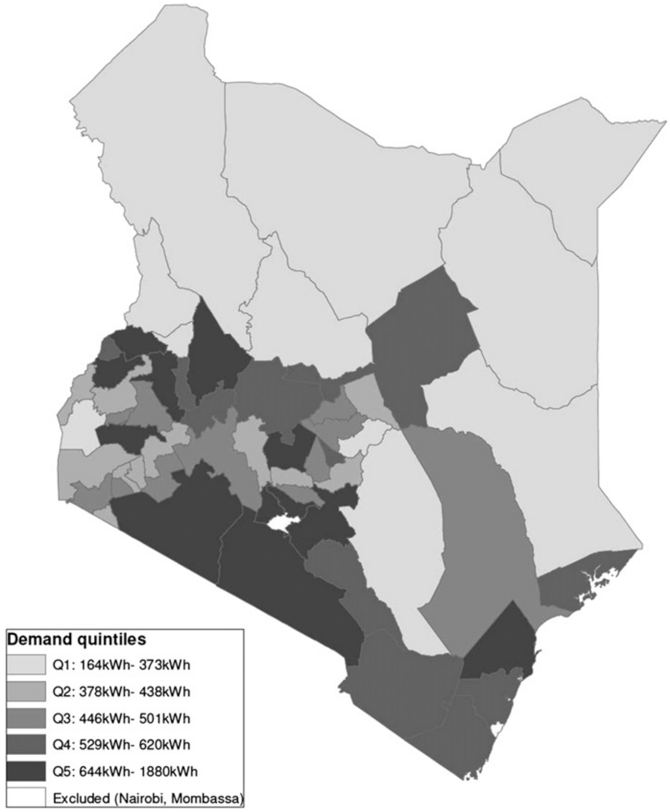

Fig. 5. Mean yearly household electricity demand per demand quintile.

household and month. The difference between the predicted and observed electricity demand can be explained by the difference in household characteristics of the newly electrified households.

Fig. 5 shows the predicted mean electricity demand of the districts on a map. We grouped districts by quintile of electricity demand ranging from lowest to highest. Predicted electricity demand is highest in Central-South and South-West Kenya. The lowest predicted mean electricity demand per household is in the North of Kenya.

Table 5 shows the mean values for the dependent and the independent regression variables using the same demand quintiles as in Fig. 5. Quintile 1 has the lowest value for the following variables: current electricity demand per household, non-food expenditures, percentage of households with a flush toilet and the educational level of the household head. It has the largest mean household size and the highest percentage of households located in rural areas. Quintile 5 shows the highest mean values for the following variables: current electricity demand per household, non-food expenditures, number of servants and percentage of households with a flush toilet. Quintile 5 has the lowest percentage of households living in a rural area.

Table 6 shows the household electricity demand in 2005/2006 according to KPLC and KIHBS (a more detailed description can be found under 3.1.1) and the predicted values for 2020.

Table 5
Mean values for the regression variables per demand quintile.

Quintile	Ed_i	Ex_i	S_i	F_i in %	A_i	E_i	N_i	R_i in %
1	3.37	5.38	0.07	0.01	44.38	4.38	5.61	0.75
2	6.83	6.52	0.06	0.05	45.17	8.08	4.75	0.69
3	6.95	6.48	0.12	0.05	45.83	7.11	4.84	0.73
4	5.86	6.43	0.08	0.08	45.38	6.58	5.39	0.69
5	7.28	6.83	0.13	0.14	43.35	6.92	5.02	0.59

Table 6
Real and predicted household electricity demand in 2005/2006 and in 2020 in GW h.

	Total household electricity demand in Kenya	Household electricity demand excluding Nairobi	
	2005/2006	2005/2006	2020
KPLC	1028	362	
KIHBS	806	384	
Predicted exponential		673	1483
Predicted Tobit		615	1432

The predicted annual demand in the exponential model amounts to 673 GW h and to 615 GW h in the Tobit model. The assumptions of normal and homogenous errors are not fulfilled in the Tobit model. Therefore, we use the predictions of the exponential model as an input into the optimization model.

Electricity demand from the KIHBS is 289 GW h lower than the predicted demand resulting from the exponential model. In other words, if all Kenyan residents were located within 100 m of an electricity grid the demand would amount to 673 GW h according to the prediction of our model. The difference of 289 GW h represents latent electricity demand. By 2020, annual demand would approximately double from 673 GW h in the exponential model and 615 GW h in the Tobit model to 1483 GW h in the exponential and 1432 GW h in the Tobit model, assuming growth in population, income, and education levels as outlined in the Electricity demand estimation of Kenyan households section.

Electricity supply side optimization model

Results of the optimization model show that for most grid cells PV electricity is the most cost-efficient electrification option for the year 2005/2006 (illustrated in Fig. 6). The demand covered with electricity produced by stand-alone PV amounts to only 15% of the total electricity consumption, 22% of the households but 80% of the grid cells.

For the 2020 scenario the changes in parameters (increase in demand, decrease in PV costs and increase in electricity generation

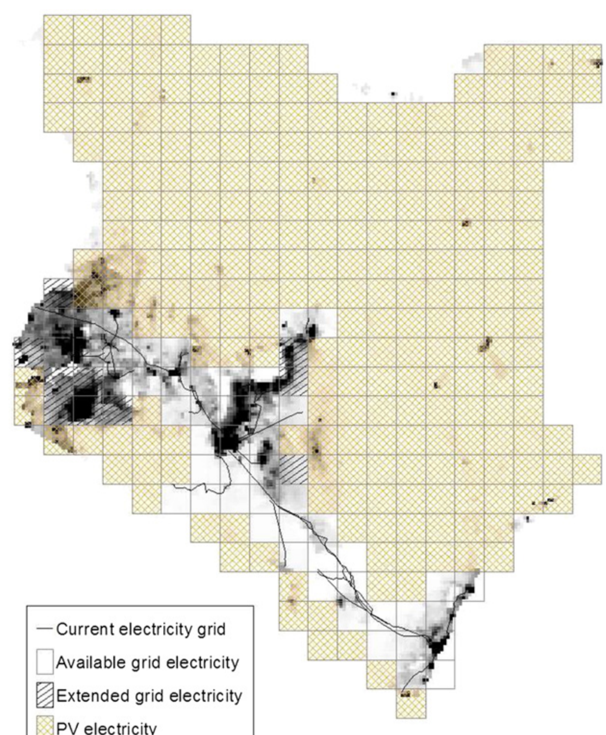

Fig. 6. Result of the optimization model.

Table 7
Mean values in cells supplied with grid and PV electricity.

	Cells with grid electricity	Cells with PV electricity
Mean number of households per grid cell	72,110	4152
Mean number of members per household	5.1	5.7
Mean monthly electricity demand per household (kW h)	73	47
Mean total demand per grid cell (kW h)	4,334,594	240,780
Mean solar irradiation (W h m^{-2} days^{-1})	1341	1407

costs) cancel each other. However, if we only increase electricity demand to 2020 values and keep all other parameters the same, six more grid cells are supplied with grid electricity compared to the scenario for the year 2006. In that case demand covered with electricity produced by stand-alone PV amounts to only 11% of the total electricity consumption, 17% of the households but 78% of the grid cells.

Differences between PV and grid electricity cells

Cells supplied with PV electricity have larger distances from the grid and lower demand than cells for which grid electricity is chosen as an optimal solution. Table 7 illustrates that the mean number of households, the mean monthly electricity demand per household, and the mean total demand per cell is lower for districts supplied with PV electricity than without PV electricity. The cells supplied with PV electricity show a slightly higher mean solar irradiation.

Costs

In the optimal solution of the baseline scenario transmission costs amount to USD 6.65 million and distribution cost to USD 82.78 million per year. Together they account for 77% of the total system costs. Total PV costs are USD 25.98 per year. In the optimal solution of the baseline scenario, costs for PV electricity amount to USD 0.243 per kW h and USD 0.143 per kW h for grid electricity.

Sensitivity analysis

In a sensitivity analysis for the baseline scenario (illustrated in Fig. 7), we analyze how a change in input parameters affects the share of PV in total electricity consumption. We vary the parameters within a range of 50%. The model shows that increasing electricity demand, solar panel costs and decreasing PV efficiency, electricity price and grid extension costs results in more districts selecting PV electricity as the cost optimal solution. Ceteris paribus, a 50% lower electricity demand increases the demand covered with electricity from PV from 15% to 21%. An increase in electricity demand by 50%, leads to 13% of the total demand then being supplied by PV electricity. A 50% change in PV efficiency leads to 7% (less efficient PV) and 27% (more efficient) of the demand being supplied by PV electricity. A change in solar panel costs has the largest effect. The share of PV electricity increases to over 45% from a cost drop of more than 50%. Solar panel cost increases have a smaller impact: an increase of costs by 50% reduces PV supply to 9% of total electricity demand. A variation in energy price by 50% leads to 23% or 11% of the electricity supplied by PV, depending on a rise or fall. A bidirectional change in grid extension cost by 50% results in 18% for higher and 11% for lower electricity grid costs of electricity demand covered with PV electricity.

Discussion and conclusions

Major efforts and investments into its electricity infrastructure are essential in order to meet the Millennium Development Goals in Kenya. Two alternatives are available to meet this basic demand, either through the extension of the national grid or through supply with stand-alone systems. The latter option is particularly important for rural households, but might also only serve as an interim solution for certain areas where grid extension will be the more cost effective solution in the future. The model results serve as a guideline for which regions of the country grid extension is more economic and where it is likely to be more beneficial to concentrate on the implementation of the stand-alone option. We find that under current circumstances the implementation of stand-alone PV systems is the more cost-effective solution for a majority of the rural area with low population density. This finding is in line with Ondraczek (2014) who comes to the conclusion that PV is already a viable energy option for off-grid applications in Kenya. Today, and even in the near future, grid electricity is a choice mainly for districts located around the existing electricity grid with high per household demand and population density. This is mainly due to low demand and low population density in a large part of the country. It may be favorable in the short-term to focus on off-grid solutions to provide people with basic electricity to meet lighting demands and power small appliances. Depending on future cost developments, only in the longer term is grid electricity likely to serve as an affordable solution for remote areas. Our results highlight the importance of accounting for region specific features in electrification planning. Results show that it can be advantageous for planning to consider and adapt to technological developments. For instance the sensitivity analysis suggests that improvements in PV efficiency and panel price reductions have a large effect on the optimal solution. Further development of the PV market in Kenya could be accompanied by a lowering of costs due to external factors, but also to learning effects in the local market of installing PV systems and improvements in design, performance and capacities of such systems. Our conclusions are quite contrary to those of Parshall et al. (2009). They show that grid extension is the cheaper option for most areas of Kenya and the choice is mainly depending on geographic features. Deichmann et al. (2011) conclude that first stand-alone renewable energy technologies represent the cheapest option for a significant minority in rural and remote areas but not in densely populated areas. This comes closer to our results.

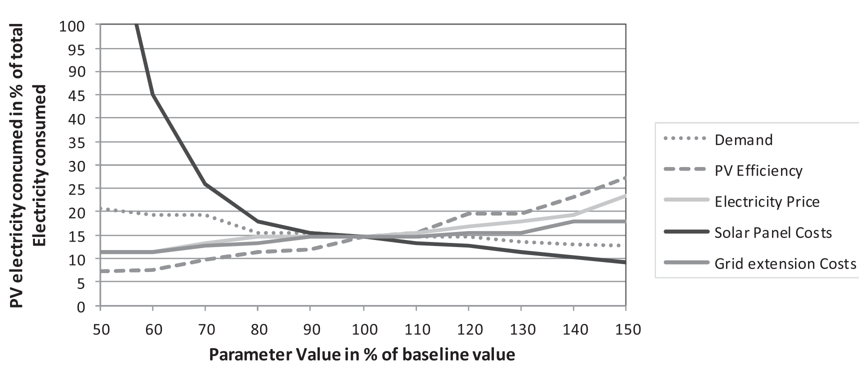

Fig. 7. Change in parameter values and effect on the share of PV electricity on total electricity consumption in the baseline scenario.

The difference in results can be explained by the use of different values: Parshall et al. (2009) use four different demand categories, from between 30 kW h per household and month for the sparsely populated/poor to 150 kW h for the densely populated/rich. Deichmann et al. (2011) use a fixed demand of 120 kW h per household and month. These values are higher than our demand values, thus this could explain differences in results. The comparison is strongly influenced by the assumed costs. The costs used by Deichmann et al. (2011) are comparable to ours: USD 90,000 for the 132 kV lines and USD 192,000 for 220 kV lines, production costs of USD 0.107 per kW h, USD 700 for a 50 W PV set (USD 12,000 per kW capital costs and USD 1956 per kW O&M costs). The costs for grid extension in Parshall et al. (2009) are lower. They amount to USD 14,098 per km for a 132 kV line and USD 90,000 for a 220 kV line of demand node connection costs and USD 10,611 per km of household connection cost. In Parshall et al. (2009) and Deichmann et al. (2011) PV system costs amount to USD 600 for a 50 W PV system assuming lifetimes for the panel of 20 years, 3 years for the battery, and 10 years for the balance.

Our approach of combining both demand and supply analysis together as done in this paper is novel. As discussed, similar studies in the past have used an engineering approach focusing on the supply side optimization, but have tended to lack the necessary detail in the estimation of demand, often assuming an average value for this. For cost-effective planning of electricity infrastructure it is crucial to estimate electricity demand and account for the heterogeneity in it across regions. The contribution of our study lies in employing the combination of a regression model to estimate household electricity demand and an electricity supply optimization model to identify least-cost electrification options. The literature review revealed that none of the papers estimate currently uncovered electricity demand due to non-availability of supply. When using the average demand of electrified households instead of a model estimation, which accounts for heterogeneity in household characteristics, the demand might not represent differences between districts. The differences in demand have been shown to be a crucial variable in the sensitivity analysis. As discussed in the previous paragraph our results differ from other studies with a similar aim which, however, do not model demand but use average consumption categories either depending on household location (e.g. rural, urban) or on income.

In its current version, the demand model does not consider prices of electricity and its substitutes. With changes in demand, one can expect changes in prices. Prices at the regional level are currently not available and elasticities could not be calculated. In the short run and for internationally traded goods such as paraffin and oil, prices are likely not to be affected by demand in Kenya.

A few caveats of our study need to be highlighted: (i) the spatial resolution may be too low to draw definite conclusions at the cell level. One needs to be cautious when taking recommendations of this model and applying them to individual towns or households. It might be advantageous to electrify the cell in general through grid electrification, but this may not hold for particular towns or households within the cell. (ii) Currently, we assume that grid electrification is always done by extension of the existing grid. It may, however, be useful to build a new electricity grid in cells which are relatively densely populated but far away from the existing grid. (iii) Besides PV, there are other off-grid solutions for electrifying rural areas such as diesel or biomass generators or small hydropower plants. We did not assess those options in detail, but a preliminary assessment of the relative cost-effectiveness between diesel and solar off-grid options (in Appendix A.2) suggests that PV is the more cost-effective stand-alone option.

Additional data and linkage with other tools could allow one to further develop the model and answer a wide range of research and policy relevant questions. First, one could carry out a dynamic assessment, which would require modeling the influence of electricity access on demand. Once electricity is available, demand may increase over time due to additional economic growth. As there are different economies of scale for grid extension and PV (grid extension getting much cheaper with increased capacity, while PV has fewer economies of scale), at some point in time grid extension is likely to be the cheaper solution. Our model is currently static and the dataset available does not allow us to model the effects of electricity access on demand. However, this would represent interesting research for the future. Second, we assume that all regions are given the same priority in supplying electricity. Other tools could be developed to help identify priority regions for electrification based on social or political factors. Finally, grid electricity may provide a different quality of electricity compared to electricity from stand-alone systems. A higher security of supply cannot be attributed to one of the two options. It depends on very specific factors such as the local capacity to maintain and repair PV, the occurrence of severe weather and generation constraints leading to blackouts, the reliability of the electric grid which may be low and the time it takes to repair line failures. However, such conditions and events depend on specific circumstances for which we don't have the necessary data and therefore, could not be accounted for in this analysis. Standalone PV systems with batteries may have rather strict capacity restrictions. However, stand-alone systems are chosen in areas with low demand density which is a consequence of low mean demand and low population density, the capacity restriction therefore will not strongly affect results. The available demand data only includes a small number of households using stand-alone systems. We are therefore not able to estimate demand of households using stand-alone systems. A difference in quality of electricity supply could lead to households making different electric equipment purchasing decisions. This would lead to a different development of electricity demand in the long-term, and perhaps even influence the choice of electrification mode directly, if households value reliability higher. The impact of changing the source of power supply is just one of many factors that may affect the (future) demand and load shape of the households, including household income, power costs, availability, type of economic activity, household composition. There is little data to support the differentiation of power demand for the source of power.

Future research that builds on our analysis and accounts for additional technological supply options, including renewable energy technologies and decentralized mini-grids, limits in transmission and generation capacities, as well as additional demand sectors, could provide more nuanced insights for future electrification planning in Kenya.

Acknowledgments

Part of the research was undertaken during the Young Scientists Summer Program (YSSP) at the International Institute for Applied Systems Analysis (IIASA). The authors would like to thank the staff of IIASA and the Austrian Academy of Sciences for the funding of the YSSP. The authors would like to gratefully acknowledge the Kenyan Bureau of Statistics for granting permission to use the data from the KIHBS, the KPLC for valuable information on cost data and Thomas Huld from EC-JRC for the solar irradiation data. Further thanks go to the BOKU University Center for International Relations (Selis Schmidt) for the funding of the research visit to Kenya. We would like to thank the two anonymous reviewers for their suggestions and comments which greatly improved the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.esd.2015.01.003.

References

Angrist JD, Pischke J-S. The credibility revolution in empirical economics: how better research design is taking the con out of econometrics. J Econ Perspect 2010;24:3–30. http://dx.doi.org/10.1257/jep.24.2.3.

Bardhan PK, Udry C. Development microeconomics. Oxford; New York: Oxford University Press; 1999.

Bazilian M, Onyeji I, Liebreich M, MacGill I, Chase J, Shah J, et al. Re-considering the economics of photovoltaic power. Renew Energy 2013;53:329–38. http://dx.doi.org/10.1016/j.renene.2012.11.029.

Berndt ER. The practice of econometrics: classic and contemporary. Reading (Mass.); Menlo Park (Calif.); New York: Addison wesley publishing company; 1991.

Blundell R, Meghir C. Bivariate alternatives to the Tobit model. J Econ 1987;34:179–200. http://dx.doi.org/10.1016/0304-4076(87)90072-8.

Briand A, Nauges C, Strand J, Travers M. The impact of tap connection on water use: the case of household water consumption in Dakar, Senegal. Environ Dev Econ 2010; 15:107–26. http://dx.doi.org/10.1017/S1355770X09990076.

Carlevaro F, Croissant Y, Hoareau S. mhurdle: Estimation of models with limited dependent variables, R package version 0.1-1; 2011.

Center for Alternative Technologies, n.d. Solar Products [WWW Document]. URL http://www.cat.co.ke/Solar%20panels%20%20Kenya.htm (accessed 9.7.12).

Center for International Earth Science Information Network (CIESIN), Columbia University, International Food Policy Research Institute (IFPRI), The World Bank, Internacional de Agricultura Tropical (CIAT). Global Rural–Urban Mapping Project. New York, United States: Socioeconomic Data and Applications Center (SEDAC); 2004.

Cragg JG. Some statistical models for limited dependent variables with application to the demand for durable goods. Econometrica 1971;39:829–44. http://dx.doi.org/10.2307/1909582.

Davis M. Rural household energy consumption: the effects of access to electricity—evidence from South Africa. Energy Policy 1998;26:207–17. http://dx.doi.org/10.1016/S0301-4215(97)00100-6.

Deaton A. Chapter 33 Data and econometric tools for development analysis. Handbook of development economics. Elsevier; 1995. p. 1785–882.

Deichmann U, Meisner C, Murray S, Wheeler D. The economics of renewable energy expansion in rural Sub-Saharan Africa. Energy Policy 2011;39:215–27. http://dx.doi.org/10.1016/j.enpol.2010.09.034.

Ekholm T, Krey V, Pachauri S, Riahi K. Determinants of household energy consumption in India. Energy Policy 2010;38:5696–707. http://dx.doi.org/10.1016/j.enpol.2010.05.017.

Filippini M, Pachauri S. Elasticities of electricity demand in urban Indian households. Energy Policy 2004;32:429–36. http://dx.doi.org/10.1016/S0301-4215(02)00314-2.

Filippini M, Hrovatin N, Zorič J. Efficiency and regulation of the Slovenian electricity distribution companies. Energy Policy 2004;32:335–44. http://dx.doi.org/10.1016/S0301-4215(02)00295-1.

Frank Knight. The wealth report 2014; 2014.

Fuso Nerini F, Howells M, Bazilian M, Gomez MF. Rural electrification options in the Brazilian Amazon: a multi-criteria analysis. Energy Sustain Dev 2014;20:36–48. http://dx.doi.org/10.1016/j.esd.2014.02.005.

GAMS Development Corporation. General Algebraic Modeling System (GAMS); 2009.

IEA-ETSAP, IRENA. Electricity Storage Technology Policy Brief E18; 2012.

IEA-ETSAP, IRENA. Solar Photovoltaics Technology Brief E11; 2013.

K.C., S., Barakat B, Goujon A, Vegard S, Warren S, Lutz W. Projection of populations by level of educational attainment, age, and sex for 120 countries for 2005–2050. Laxenburg, Austria: International Institute for Applied Systems Analysis; 2010.

Kaundinya DP, Balachandra P, Ravindranath NH. Grid-connected versus stand-alone energy systems for decentralized power—a review of literature. Renew Sustain Energy Rev 2009;13:2041–50. http://dx.doi.org/10.1016/j.rser.2009.02.002.

Kemausuor F, Adkins E, Adu-Poku I, Brew-Hammond A, Modi V. Electrification planning using Network Planner tool: the case of Ghana. Energy Sustain Dev 2014;19: 92–101. http://dx.doi.org/10.1016/j.esd.2013.12.009.

Kenya Electricity Generating Company. About us: Our Values [WWW Document]. URL http://www.kengen.co.ke/index.php?page=aboutus&subpage=values, 2014. [accessed 5.6.14].

Kenya National Bureau of Statistics. Kenyan Integrated Household Budget Survey; 2005 [Nairobi].

Kenya National Bureau of Statistics. Basic report: Kenya integrated household budget survey (KIHBS), 2005/06. Nairobi, Kenya: Kenya National Bureau of Statistics; 2007.

Kenya National Bureau of Statistics. Electricity installed capacity by type of power, 2004–2012; 2014.

Kenya's Ministry of East African Affairs, Commerce and Tourism, Department of Tourism. Hotel bed occupancy by zone; 2010.

Kenyan Bureau of Statistics. Kenya Facts and Figs. 2013 [WWW Document]. URL http://www.knbs.or.ke/index.php?option=com_phocadownload&view=category&id=20&Itemid=1107, 2013.

Kenyan Power and Lighting Company. Transmission line costs; 2010.

Kocaman AS, Huh WT, Modi V. Initial layout of power distribution systems for rural electrification: a heuristic algorithm for multilevel network design. Appl Energy 2012;96: 302–15. http://dx.doi.org/10.1016/j.apenergy.2012.02.029.

Lay J, Ondraczek J, Stoever J. Renewables in the energy transition: evidence on solar home systems and lighting fuel choice in Kenya. Energy Econ 2013;40:350–9. http://dx.doi.org/10.1016/j.eneco.2013.07.024.

Levin T, Thomas VM. Least-cost network evaluation of centralized and decentralized contributions to global electrification. Energy Policy 2012;41:286–302. http://dx.doi.org/10.1016/j.enpol.2011.10.048.

Long JS. Regression models for categorical and limited dependent variables. Thousand Oaks: Sage Publications; 1997.

Louw K, Conradie B, Howells M, Dekenah M. Determinants of electricity demand for newly electrified low-income African households. Energy Policy 2008;36:2812–8. http://dx.doi.org/10.1016/j.enpol.2008.02.032.

Masera OR, Saatkamp BD, Kammen DM. From linear fuel switching to multiple cooking strategies: a critique and alternative to the energy ladder model. World Dev 2000; 28:2083–103. http://dx.doi.org/10.1016/S0305-750X(00)00076-0.

Ministry of Energy, Rural Electrification Authority, Rural Electrification Autrhority, Re-Engineering Africa Consortium. Rural Electrification Master Plan; 2009 [Nairobi].

Narula K, Nagai Y, Pachauri S. The role of decentralized distributed generation in achieving universal rural electrification in South Asia by 2030. Energy Policy 2012;47: 345–57. http://dx.doi.org/10.1016/j.enpol.2012.04.075.

Ondraczek J. Are we there yet? Improving solar PV economics and power planning in developing countries: the case of Kenya. Renew Sustain Energy Rev 2014;30:604–15. http://dx.doi.org/10.1016/j.rser.2013.10.010.

Organisation for Economic Co-operation and Development, International Energy Agency. World energy outlook 2013. Paris: OECD/IEA; 2013.

Pachauri S. An analysis of cross-sectional variations in total household energy requirements in India using micro survey data. Energy Policy 2004;32:1723–35. http://dx.doi.org/10.1016/S0301-4215(03)00162-9.

Parshall L, Pillai D, Mohan S, Sanoh A, Modi V. National electricity planning in settings with low pre-existing grid coverage: development of a spatial model and case study of Kenya. Energy Policy 2009;37:2395–410. http://dx.doi.org/10.1016/j.enpol.2009.01.021.

Pasek J. Weights: weighting and weighted statisticsR package version 0.70. ed.; 2011.

Powell JL. Least absolute deviations estimation for the censored regression model. J Econ 1984;25:303–25. http://dx.doi.org/10.1016/0304-4076(84)90004-6.

R Development Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2012.

Riahi K, Dentener F, Gielen D, Grubler A, Jewell J, Klimont Z, et al. Chapter 17 — energy pathways for sustainable development. Global energy assessment — toward a sustainable future. Laxenburg, Austria: Cambridge University Press, Cambridge, UK and New York, NY, USA and the International Institute for Applied Systems Analysis; 2012. p. 1203–306.

Rigollier C, Lefèvre M, Wald L. The method Heliosat-2 for deriving shortwave solar radiation from satellite images. Sol Energy 2004;77:159–69. http://dx.doi.org/10.1016/j.solener.2004.04.017.

Rosenbaum PR. Design of observational studies. New York, NY: Springer; 2010.

Rosnes O, Vennemo H. The cost of providing electricity to Africa. Energy Econ 2012;34: 1318–28. http://dx.doi.org/10.1016/j.eneco.2012.06.008.

Solarlink. Solar Panels [WWW Document]. URL http://www.solarlink.de/solarmodulegb.htm. [accessed 9.7.12].

Sollatek Electronics Kenya Limited. Solar Power [WWW Document]. URL http://www.sollatek.co.ke/category/solar-power/. [accessed 9.7.12].

Szabó S, Bódis K, Huld T, Moner-Girona M. Energy solutions in rural Africa: mapping electrification costs of distributed solar and diesel generation versus grid extension. Environ Res Lett 2011;6:034002. http://dx.doi.org/10.1088/1748-9326/6/3/034002.

Tatiétsé TT, Villeneuve P, Ngundam J, Kenfack F. Contribution to the analysis of urban residential electrical energy demand in developing countries. Energy 2002;27:591–606. http://dx.doi.org/10.1016/S0360-5442(02)00006-3.

The World Bank. Kenya Economics Update; 2014.

Tobin J. Estimation of relationships for limited dependent variables. Econometrica 1958; 26:24. http://dx.doi.org/10.2307/1907382.

United Nations General Assembly. Resolution adopted by the General Assembly [on the report of the Second Committee (A/65/436 and Corr.1)] 65/151. International year of sustainable energy for all; 2011.

Wholesale Solar. Compare Solar Panels Before You Buy [WWW Document]. URL http://www.wholesalesolar.com/solar-panels.html. [accessed 9.7.12].

Wooldridge JM. Econometric analysis of cross section and panel data. Cambridge, Mass: MIT Press; 2010.

Zvoleff A, Kocaman AS, Huh WT, Modi V. The impact of geography on energy infrastructure costs. Energy Policy 2009;37:4066–78. http://dx.doi.org/10.1016/j.enpol.2009.05.006.