

---

# **Test Documentation**

*Release 1.1a7.dev57+g7e46d87*

**Kostas**

**Nov 10, 2023**



# CONTENTS

<b>1</b>	<b>Electricity for all</b>	<b>3</b>
<b>2</b>	<b>Contents</b>	<b>5</b>
2.1	Introduction to OnSSET	5
2.1.1	Universal access to electricity	5
2.1.2	What is OnSSET?	5
2.1.3	Scope and Objective	6
2.2	The OnSSET Model	6
2.2.1	Electricity demand	6
2.2.2	Electricity supply	8
2.2.3	Methodology Overview	14
2.3	Software installation	15
2.3.1	Requirements	15
2.3.2	Software installation and setup	16
2.3.3	Additional Info	17
2.4	GIS data acquisition	17
2.4.1	Geographic Information Systems	17
2.4.2	GIS data sources	18
2.4.3	GIS data in OnSSET	20
2.5	GIS data preparation	28
2.5.1	<b>Step 1. Proper data types and coordinate system</b>	28
2.5.2	<b>Step 2. Layer projection</b>	29
2.5.3	<b>Step 3. Generate population clusters</b>	29
2.5.4	<b>Step 4. Generate the OnSSET input file</b>	30
2.5.5	GIS country file	30
2.6	Techno-economic and demographic parameter	31
2.6.1	Input model parameters	31
2.6.2	Discount rate	32
2.6.3	Off-grid technology costs	33
2.6.4	Diesel pump price	35
2.6.5	Household size	35
2.7	Running the analysis	35
2.7.1	Setting up PyCharm - Afghanistan Demo	35
2.7.2	Installing/updating modules and packages	38
2.7.3	Running the code	39
2.8	Analysis and visualizations	41
2.8.1	Output parameters	41
2.8.2	Summaries output	42
2.8.3	Maps output	43
2.9	Online interfaces & Applications	45

2.9.1	United Nations app . . . . .	45
2.9.2	World Bank app . . . . .	46
2.10	References . . . . .	47
2.10.1	Publications . . . . .	47
2.10.2	Contribution . . . . .	48
2.10.3	Partners . . . . .	48
2.11	Contact . . . . .	49
2.11.1	Feedback . . . . .	50
2.11.2	Funding . . . . .	50
2.11.3	Admininstration . . . . .	50
2.12	FAQ . . . . .	50
2.13	License . . . . .	51
2.14	onsset . . . . .	51
2.14.1	onsset package . . . . .	51
<b>3</b>	<b>Indices and tables</b>	<b>59</b>
	<b>Python Module Index</b>	<b>61</b>
	<b>Index</b>	<b>63</b>

# OnSSET

Open Source Spatial Electrification Tool

Fig. 1: KTH dES - <http://www.onsset.org>



## ELECTRICITY FOR ALL

In September 2015, the United Nations General Assembly adopted Agenda 2030, which comprises a set of 17 Sustainable Development Goals (SDGs) defined by 169 targets. “Ensuring access to affordable, reliable, sustainable and modern energy for all by 2030” is the seventh goal (SDG7). According to the Tracking SDG 7: The Energy Progress Report of 2020 approximately 789 million people lacked access to electricity in 2018, and many more rely on poor quality electricity services. The majority of those without access (85%) reside in rural areas. Access to electricity services is a prerequisite for sustainable development and a powerful factor in poverty alleviation and well-being.





## CONTENTS

### 2.1 Introduction to OnSSET

#### 2.1.1 Universal access to electricity

The 2030 Agenda for Sustainable Development has set the goal of universal access to electricity by 2030 (SDG7). The challenge is significant. It involves reaching population with limited income, often living in sparsely populated areas, mostly in developing and least developed countries. The choice of which technology to use for increasing electricity access, depends on a number of parameters, from social to techno-economic including e.g. target level of energy access, local population density, distance to the national grid and local resource availability. These parameters are spatial in nature, making geospatial information very useful for their evaluation on both regional, national and sub-national scale.



#### 2.1.2 What is OnSSET?

Over the past few years, the division of Energy Systems at KTH (formerly division of Energy Systems Analysis) has been embracing the advancements in the geospatial field, by developing together with partners an open source, geospatial electrification toolkit – the Open Source Spatial Electrification Tool (OnSSET). OnSSET is a GIS based tool developed to identify the least-cost electrification option(s) between seven alternative configurations:

- Grid connection/extension
- Mini grid systems (Solar PV, Wind Turbines, Diesel gensets, Small scale hydropower)
- Stand-alone systems (Solar PV, Diesel gensets)

The results indicate the technology mix, capacity and investment requirements for achieving universal access in the modelled country, under a certain time period (usually until 2030). The findings can be presented in various formats such as interactive maps, graphs, images, tables etc.

### 2.1.3 Scope and Objective

OnSSET output, can be used to support policy development for electrification by taking into consideration specific guidelines and targets for each country. This, could help answer questions like:

- What would it cost to provide universal electricity access in the country? How does this cost change based on the targeted consumption level per capita?
- In case of limited resources, which customers/areas shall be prioritized? What key consumers should electrification activities focus on?
- What part of population lives within 50 km from the existing grid infrastructure and what part in remote (last-mile) communities?
- What is the most cost effective way to electrify populations in the last mile, what technologies and costs would that imply?
- How do diesel generators compare with off grid renewables (solar PV, wind, hydro) in terms of overall economic performance? What are the implications of an imminent carbon pricing scheme?
- What is the importance/impact of diesel price in the generation mix for electrification?
- How does fuel cost for diesel compares with high upfront cost of renewables in terms of subsidy schemes (continuous vs one-time subsidy)?

To the current state, OnSSET has supported electrification efforts in many countries around the globe including Afghanistan, Nigeria, Ethiopia, Kenya, India, Tanzania, Madagascar and Benin as part of joint collaboration with the World Bank, the United Nations and SNV. These organizations, have warmly embraced the tool in their fight against energy poverty. In addition, OnSSET has featured in several peer reviewed publications including the World Energy Outlook in 2014, 2015, 2017 and 2019.

## 2.2 The OnSSET Model

OnSSET is a bottom-up medium to long term optimization model. Its principle is simple. Using population settlements together with different geospatial characteristics and socio- and tecno-economic data the tool tries to identify the least-cost supply option in order to reach universal electrification in each location.

The least-cost technology mix and investments required to fully electrify a country within a defined timeframe (e.g. 2030) can vary significantly depending on demand and supply characteristics, such as:

- 1) The level of electricity access to be achieved in a country (kWh/person/year).
- 2) The current status and plans of the national electricity grid along with its associated costs.
- 3) Local energy resource availability and off-grid power systems costs.

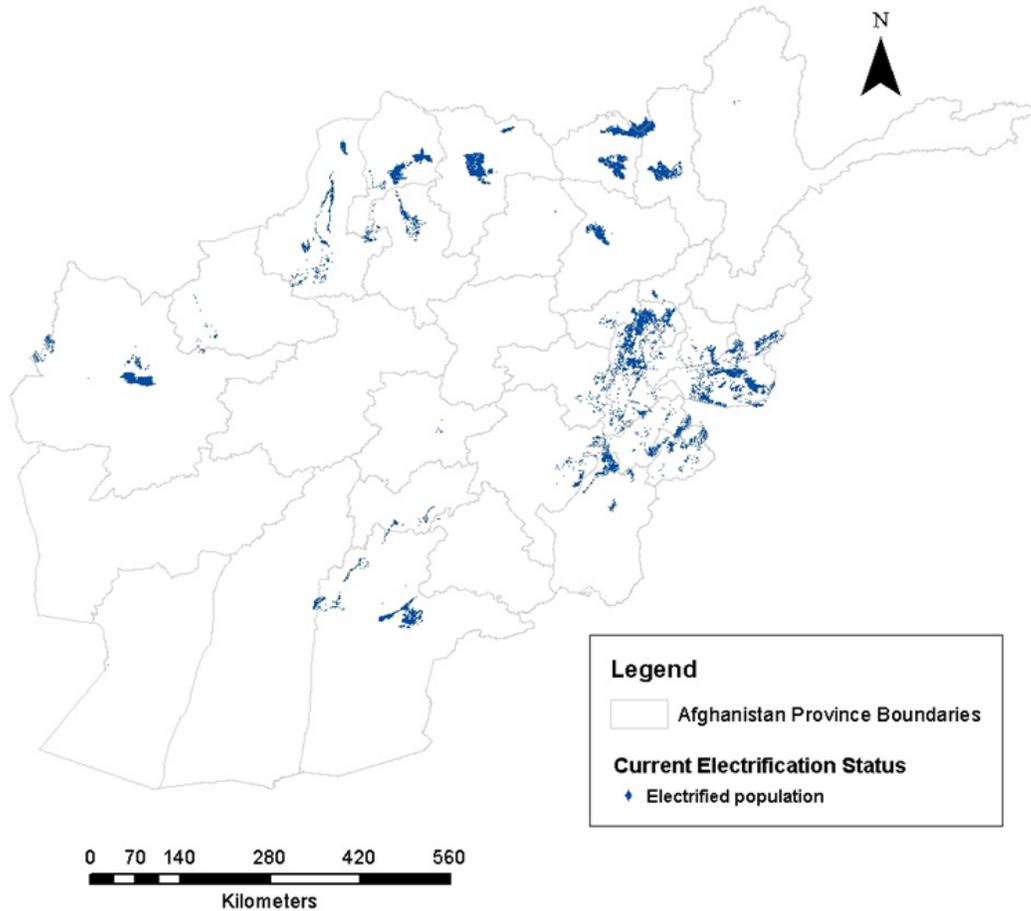
The following paragraphs describe in brief how an electrification analysis with OnSSET works.

### 2.2.1 Electricity demand

The goal of OnSSET is clear, to identify the mix of technologies that will help to fully electrify the population of a country within a certain timeframe. Currently, OnSSET only accounts for residential electricity demand.

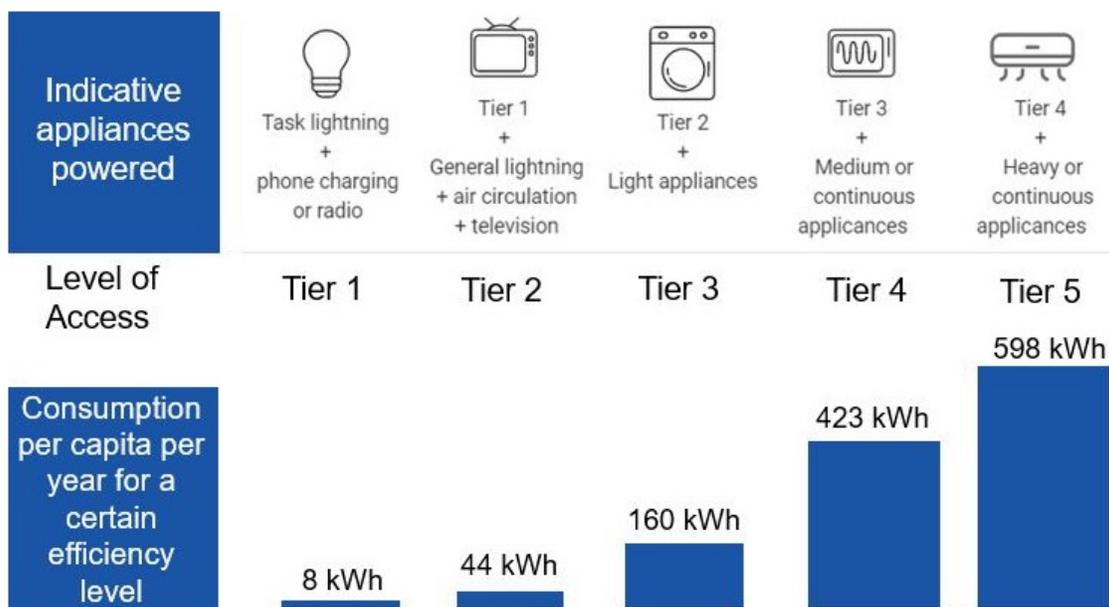
## Geo-location of population and current electrification status

Using population density maps, OnSSET can identify where population is located within the study area. Depending on the number of people living in a settlement, OnSSET categorizes a settlement as either urban or rural. Other geospatial information such as the distance between a settlement and its closest electricity grid, its distance from closest transformer, population density and its brightness index at night (from satellite imagery), helps OnSSET identify the current electrification status per settlement. Each settlement is characterized as electrified or un-electrified.



## Residential Electricity Demand

Once the electrified and un-electrified population has been spatially identified, OnSSET uses population growth rates per type of settlement (urban/rural) to estimate the projected population at the end year of the analysis (e.g. 2030). This is one of the two parameters we need to know to quantify and locate the future electricity demand. The second parameter is the target level of electricity; OnSSET adopts the consumption levels suggested by the Global Tracking Framework (2015).



Consumption levels start from 8 kWh/person/year, enough to support minor daily activities (few hours of lighting, phone charging, radio etc.) and reach up to 598 kWh/person/year, supporting the use of heavier or continuous appliances in a household like refrigerator, washing machine, oven etc. These values can easily be changed in the OnSSET model in order to better fit the study area of your choice. The combination of projected population and targeted consumption level, indicates the electricity demand for residential purposes per location.

Note, that OnSSET provides the option of determining the demand either through a top-down approach or a bottom-up one. In the top-down approach the user selects which tier the urban and rural settlements reach respectively and the demand in each settlement is determined as:

$$\text{Electricity Demand} = \text{Population in 2030} \times \text{Selected Access Tier}$$

In the bottom-up approach the user can use a combination of different GIS-datasets (e.g. GDP and poverty) in order to generate a custom demand estimate for each settlement. This approach requires the user to generate a demand layer and gives more disaggregated measure of demand than the urban-rural division in the top-down approach.

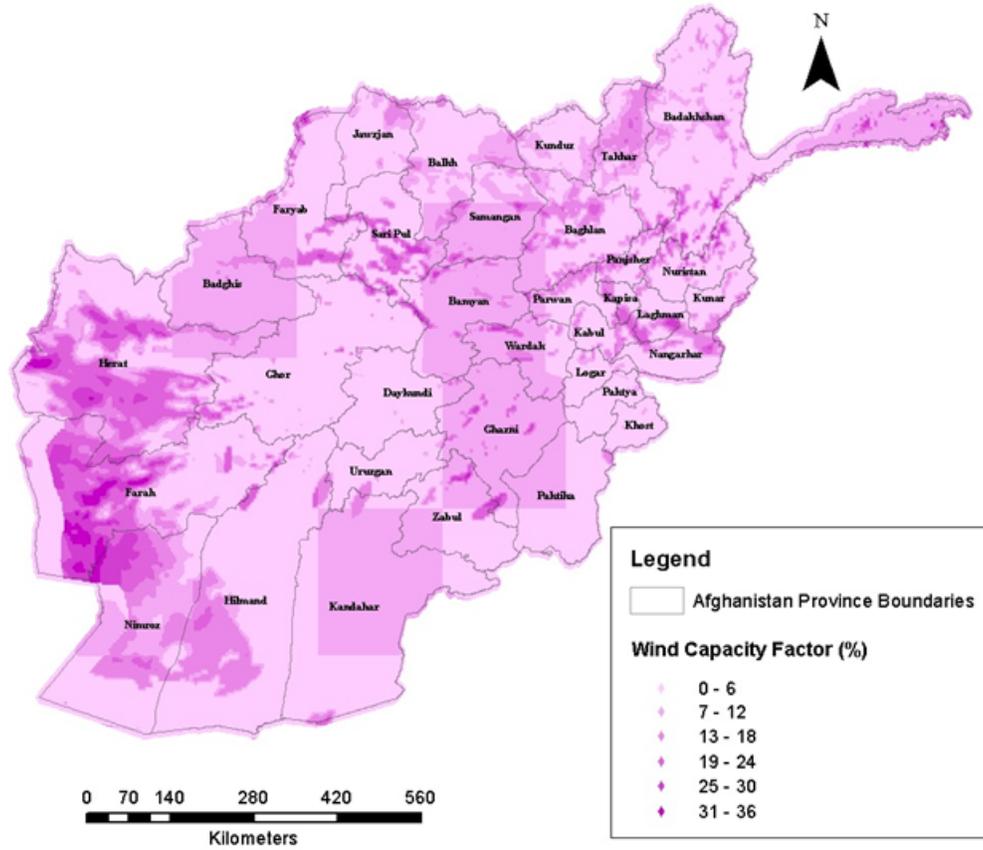
## 2.2.2 Electricity supply

OnSSET uses GIS to seize location specific information of power infrastructure. It also relies on GIS data in order to grasp the spatial and temporal fluctuation of energy resources. All these data, are later translated into techno-economic inputs and are very useful in assessing the performance of various technologies.

### Energy resources assessment

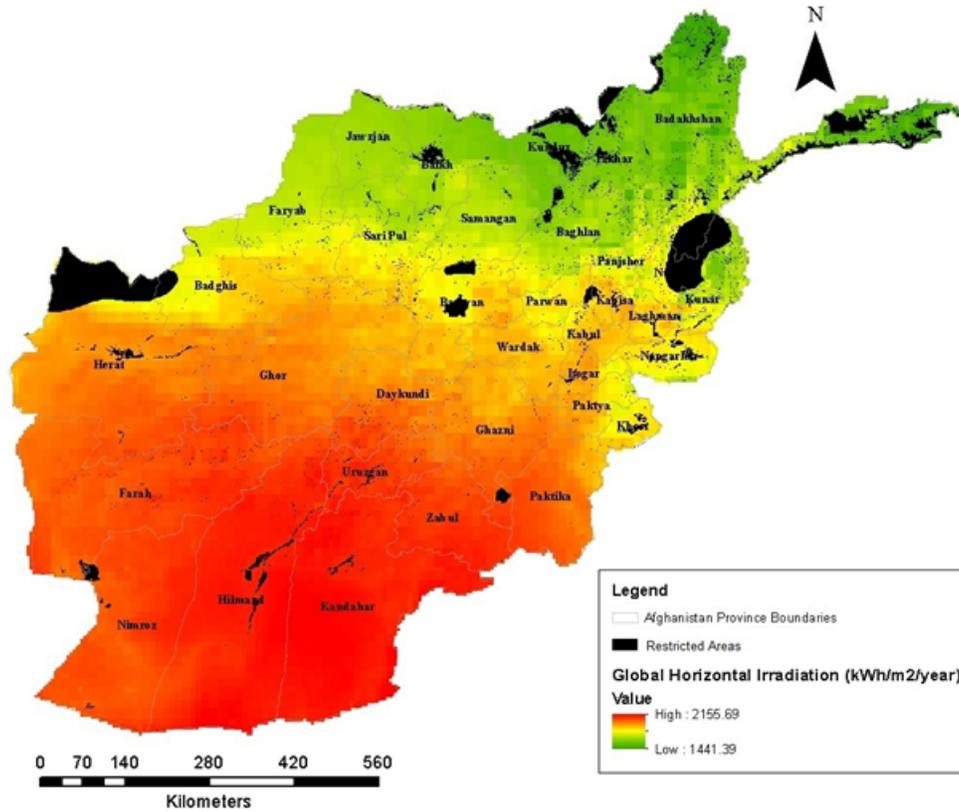
#### Wind energy potential and capacity factors

Wind speed data extracted into a GIS environment are used to calculate wind turbine capacity factors per location. The latter is defined as the ratio of the yearly expected wind energy production to the energy production if the wind turbine were to operate at its rated power throughout the year. The capacity factor reflects the potential wind power at a given site and it can be used for comparing different sites before the installation of wind power plants. Additional documentation on wind power assessment is available [here](#).



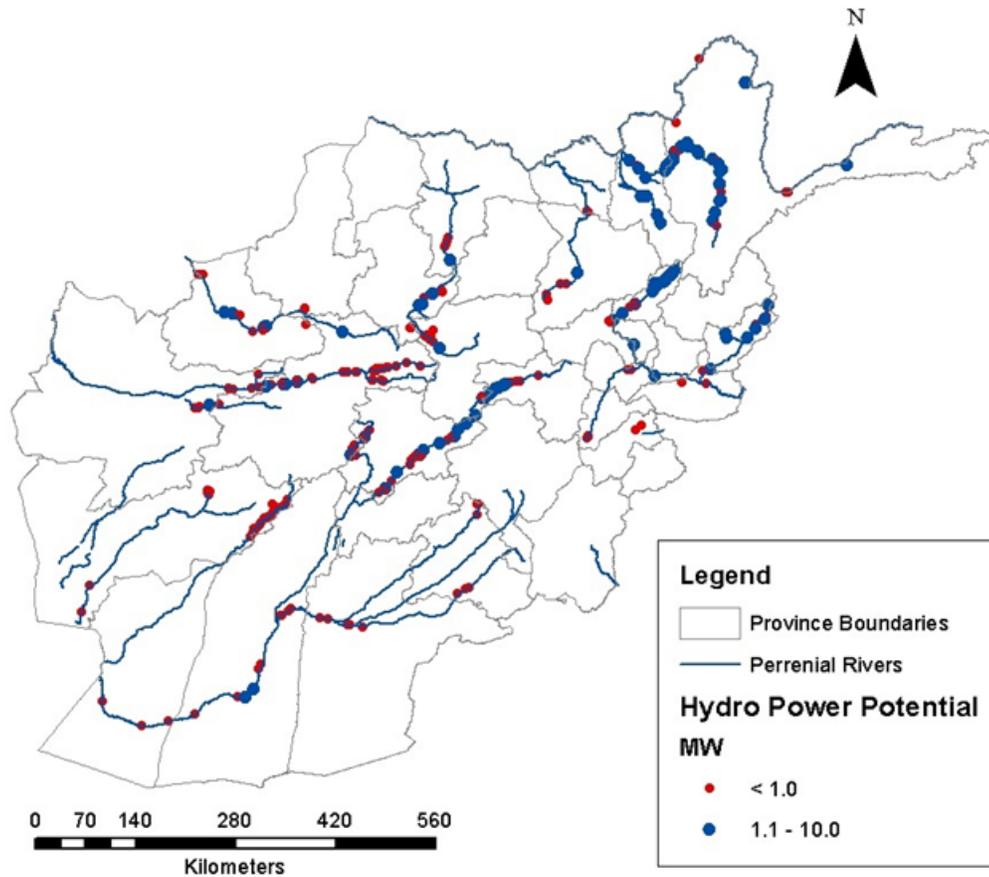
### Solar energy potential

GIS data for Global Horizontal Irradiation (GHI - kWh/m<sup>2</sup>/time) are used to indicate the available solar energy that Photovoltaic (PV) systems can utilize per location. Documentation on solar power assessment is available [here](#).



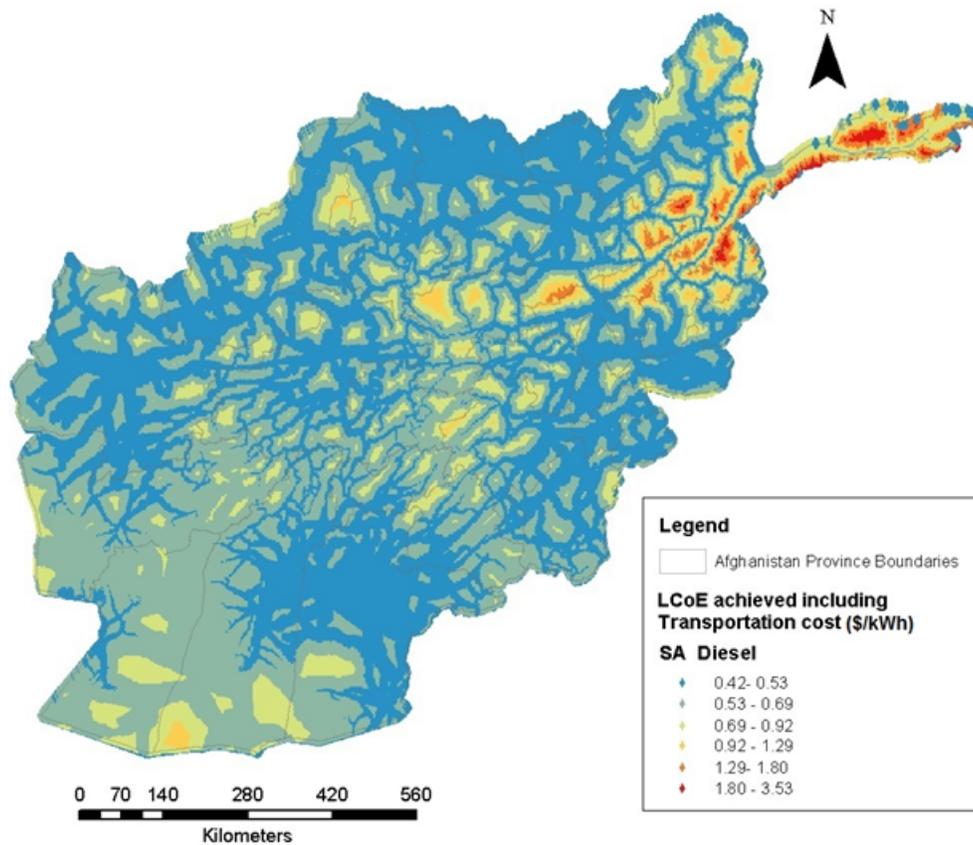
### Hydropower potential

A number of GIS datasets (runoff, accumulation, elevation, river network) have been utilized in a novel methodology developed by KTH dES in order to spatially identify potential sites for small scale hydropower deployment. Documentation on hydropower assessment together with a GIS based assessment tool is available [here](#).



### Transportation Cost for Diesel

Diesel is an important energy carrier, especially in remote areas of many developing countries and is therefore included in the OnSSET analysis. Transportation of diesel incurs costs, which may lead to high costs of electricity for isolated and low populated communities. OnSSET uses GIS datasets that indicate travel time and distance from main urban hubs, in order to calculate and assign transportation costs for diesel in each location. These costs are then included as fuel costs in the calculation of LCoE for diesel gensets. Documentation on diesel cost assessment is available [here](#).



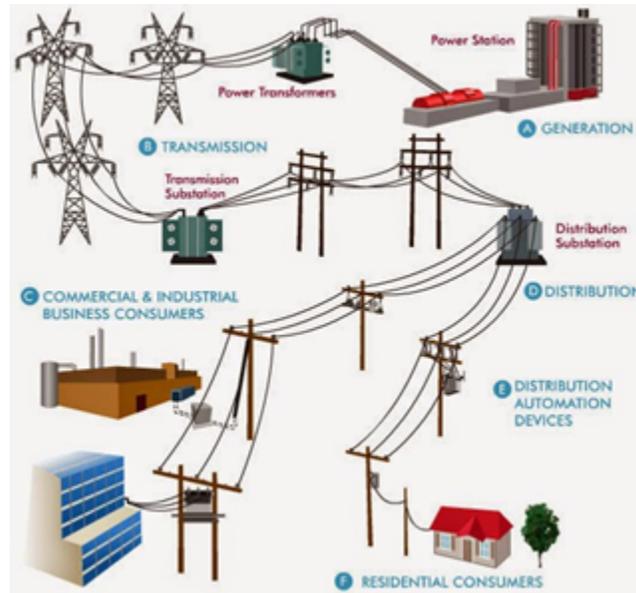
## Electrification options

Over the last few decades, access to electricity in many parts of the world has been primarily achieved through the extension of the national electricity grid. Leveraging on economies of scale, centralized grids offer electricity at low costs. Power grid expansion is however a capital intensive and lengthy process. In contrary, decentralized power system configurations i.e. mini-grids or stand-alone systems may be a better fit for electrification of remote, sparsely populated areas. Technological advances over the past decades along with innovative business models, have made such systems rather cost competitive. In addition, if based on local renewable sources they increase supply reliability while decreasing reliance on imported fuels.

OnSSET estimates and compares the performance of seven technology configurations that fall within three main categories: grid-extension, mini-grids and stand-alone systems.

## Grid extension

Central grids can offer low generating costs. However, grid extension might not be economically or socially feasible if the purpose is to meet a relatively small electricity demand or reaching remotely located areas.



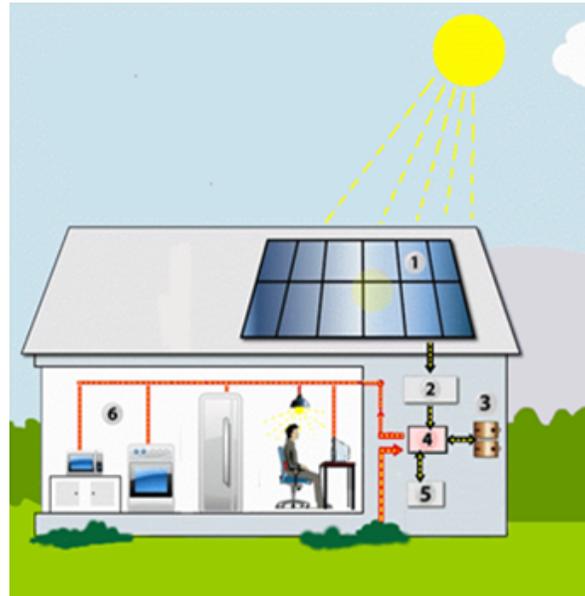
## Mini-grids - Wind Turbines, Solar PVs, Mini/Small Hydro, Diesel generators

Mini-grids usually provide electricity from small power plants with generating capacity of a few MW. They tap locally available energy resources such as solar, hydro, wind, or can use commonly available fuels such as diesel. Overall, they can provide affordable electricity to rural and remote areas with low to medium electricity consumption habits. Cost-wise, if based on renewable sources, they usually have moderate to high upfront investment costs but small operational monetary costs and no fuel costs. On the other hand, diesel generator sets (gensets) are a mature technology with low upfront investment cost but subjected to operational costs depending on diesel pump price and transport costs fluctuations.

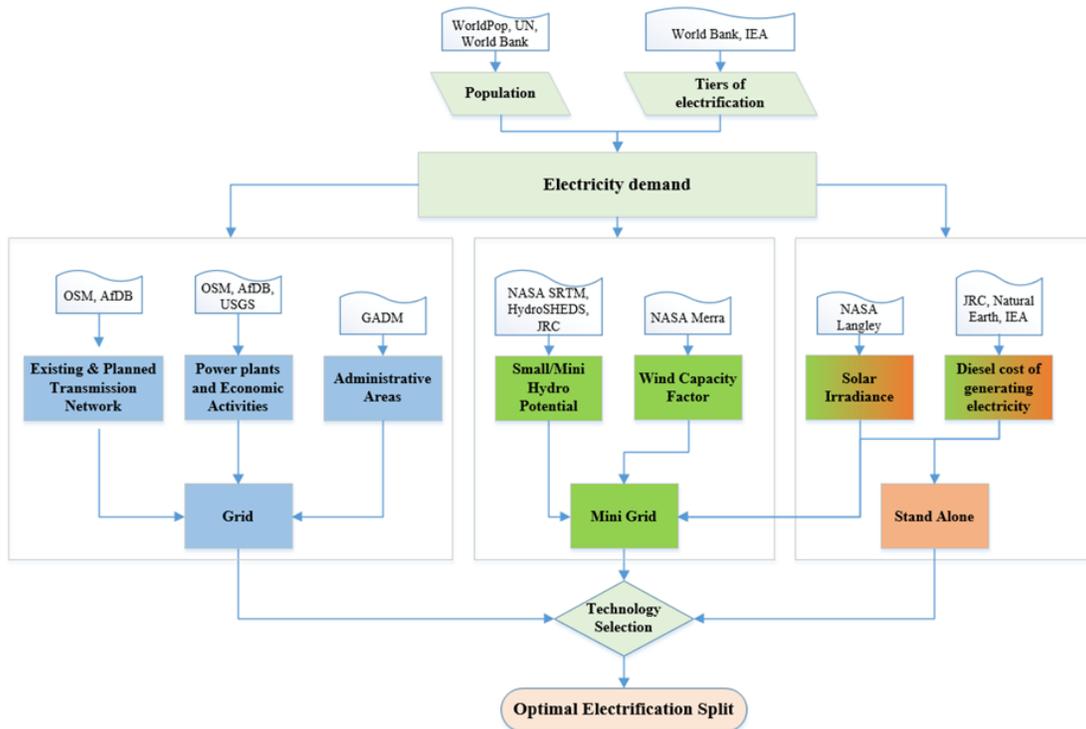


### Stand-alone systems (SA) - Solar PVs, Diesel generators

Like mini-grids, these systems are usually based on local energy resources but the difference is that these can produce only few kWh per day, suitable to cover the electricity demand of a single household or a small business, but no more. Stand-alone systems do not require a T&D network nor construction investments. The capital cost of these systems is not high and depends mainly on size. Batteries, allowing for electricity when dark, may increase the upfront cost for PV systems.



### 2.2.3 Methodology Overview



### Brief description of the electrification algorithm

The electrification algorithm procedure is based on two separate, yet complementary processes. On the one hand, a GIS analysis is required to obtain a settlement table referencing each settlement's position –i.e., its x and y coordinates – and information related to demand, resource availability, infrastructure and economic activities. Night-time light datasets are used in combination with population density and electricity related infrastructure (transformers, transmission lines and distribution lines) in order to identify the presently electrified population. The initial electrification status is listed as either 1 (electrified) or 0 (non-electrified).

The algorithm calculates the cost of generating electricity at each cell for different electrification configurations based on the local specificities and cost related parameters. Depending on the electricity demand, transmission and distribution network requirements, energy resource availability etc. the LCoE for each of the seven technology configurations is calculated in each settlement. The LCoE of a specific technology configuration represents the final cost of electricity required for the overall system to break even over the project lifetime.

---

**Note:** The LCoE calculations for the mini-grid and standalone electrification options reflect the total system costs while the LCoE for the grid option is the sum of the average LCoE of the national grid plus the marginal LCoE of transmitting and distributing electricity from the national grid to the demand location.

---

Once the LCoEs for all the off-grid technology configurations have been calculated the grid extension algorithm is executed. For each cell electrified by the national grid the algorithm iterates through all non-electrified cells to test if the conditions for their connection to the electrified settlement are fulfilled. These conditions include: a) lower cost of generating, transmitting and distributing electricity as compared to the off-grid technologies and b) not causing the total additional MV grid length to exceed 50 km if it is connected.

If these conditions are fulfilled, the settlement status is set to electrified (by the national grid). At the same time, the algorithm stores the length of the additional MV lines that have been built thus far by the model to connect this new settlement. This is required to ensure all newly electrified cells comply with the 50 km limit for the length of MV lines. Further, this is also used to consider cost increases for each additional MV extension, due to the requirement to strengthen the previously built grid line. This process is repeated with the newly electrified cells until no additional cells are being electrified, and thus until all settlements to which the grid can be economically extended are reached. Settlements that are not connected to the grid will get access to electricity through mini grid or stand-alone systems. This decision is based on a cost comparison process where the off-grid technology which can meet the electricity demand at the lowest LCoE selected for each cell.

### Penalty cost assignment to electricity grid expansion processes

The expansion of the transmission network to areas lacking access is a capital intensive process. The investment costs are influenced by several factors such as the capacity and the length of the lines as well as by the topology of the subjected area. In this analysis, a number of geospatial factors that affect the investment costs of the transmission network are identified and considered in order to assign an incremental capital cost in locations that indicate specific topological features. Particularly, investment cost is influenced by elevation, the road network, land cover type, slope gradient and distance from substations.

## 2.3 Software installation

### 2.3.1 Requirements

#### QGIS

OnSSET is a spatial electrification tool and as such highly relies on the usage of Geographic Information Systems (GIS). A GIS environment is therefore necessary for two main reasons:

- Extract trivial characteristics for the electrification analysis from GIS layers and combine them all together in a format easy to read in python (a .csv file with all the attributes per population point).
- Visualize the final results in maps.

Historically OnSSET has relied on ArcGIS, but recently the process has been moved to open source GIS environments. OnSSET in it's current form relies on QGIS (together with in QGIS included tools from SAGA, GDAL and GRASS)

---

**Note:** In order to assure that the QGIS section of OnSSET functions correctly make sure to donwload the latest long-term release of QGIS (version 3.10 as of the 7th of August 2020)

---

Download QGIS for free from the official [QGIS website](#).

### Python - Anaconda package

OnSSET is written in python, an open source programming language used widely in many applications. Python is a necessary requirement for the OnSSET tool to work. Programming in python usually relies on the usage of pre-defined functions that can be found in the so called modules. In order to work with OnSSET, certain modules need to be installed/updated. The easiest way to do so is by installing Anaconda, a package that contains a wide range of Python packages in one bundle. Anaconda includes all the Python packages required to run OnSSET successfully. Python itself can be downloaded and installed for free from the official [Python website](#).

### Python Interfaces - Integrated Development Environment (IDEs)

#### PyCharm

Integrated Development Environments are used in order to ease the programming process when multiple or long scripts are required. There are plenty of IDEs developed for Python, KTH dESA has been using PyCharm as the standard IDE to run OnSSET.

#### Jupyter notebook (via Anaconda)

Jupyter notebook is a console-based, interactive computing approach providing a web-based application suitable for capturing the whole computation process: developing, documenting, and executing code, as well as communicating the results. Jupyter notebook is used for the online OnSSET interface, recommended for small analyses and exploring code and results.

#### GitHub

GitHub is a web-based Git repository hosting service. It provides access control and several collaboration features such as bug tracking, feature requests, task management, and wikis for every project. OnSSET is an open source tool therefore the code behind it is open and freely accessible to any user. The code behind the OnSSET tool is called "PyOnsset" and is available in KTH dESA's Github space. A GitHub account will allow you to propose changes, modifications and upgrades to the existing code. Access the repository on [Github](#).

## 2.3.2 Software installation and setup

1. Download **Anaconda** [here](#) and install.
  - Please make sure that you download the version that is compatible with your operating system (Windows/MacOS/Linux - In case you run Windows open the *Windows Control Panel*, go to *System and Security System* and check e.g. Windows 32-bit or 64-bit).
  - Following the installation process make sure that you click on the option "Add Python X.X to PATH". Also by choosing to customize the installation, you can specify the directory of your preference (suggest something convenient e.g. C:/Python35/..).
  - After the installation you can use the Anaconda command line (search for "Anaconda Prompt") to run python. It should work by simply writing "python" and pressing enter, since the path has already been included in the

system variables. In case this doesn't work, you can either navigate to the specified directory and write "python" there, or add the directory to the PATH by editing the [environment variables](#).

2. Download **PyCharm** [here](#) and install.

Please make sure that you download the version that is compatible with your operating system. Select the "Community", open-source version.

3. Download the code from **GitHub** [here](#).

### 2.3.3 Additional Info

- [Basic navigating commands for DOS \(cmd\)](#).
- [Modules and packages installation documentation from python.org](#).

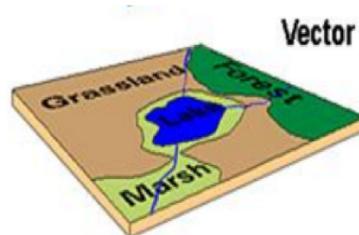
## 2.4 GIS data acquisition

### 2.4.1 Geographic Information Systems

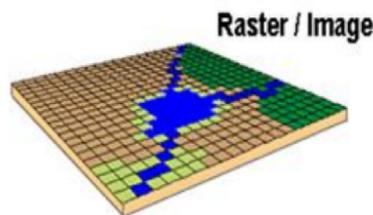
A Geographic Information System (GIS) is an integrated set of hardware and software tools, designed to capture, store, manipulate, analyse, manage, and digitally present spatial (or geographic) data and related attribute information. GIS can relate information from different sources, using two key index variables space (or location) and time. Common GIS data types (models) include:

**Spatial Data:** Describe the absolute and relative location of geographic features.

- Vectors
  - Arcs (Polylines): Line segments forming individual linear features
  - Polygons: Areas enclosed by arcs
  - Points: Single coordinate pairs



- Rasters
  - Grid-Cells: single column/row positions
  - Cell size: Resolution or else the accuracy of the data



**Attribute data:** Describe characteristics of the spatial features. These characteristics can be quantitative and/or qualitative in nature. Attribute data is often referred to as tabular data.

The selection of a particular data model, vector or raster, is dependent on the source and type of data, as well as the intended use of the data. Certain analytical procedures require raster data while others are better suited to vector data.

## 2.4.2 GIS data sources

### EnergyData.info

Every day governments, private sector and development aid organizations collect data to inform, prepare and implement policies and investments. Yet, while elaborate reports are made public, the data underpinning the analysis remain locked in a computer out of reach. Because of this, the tremendous value they could bring to public and private actors in data-poor environments is too often lost.

Energydata.info is an open data platform launched recently by The World Bank Group and several partners, trying to change energy data paucity. It has been developed as a public good available to governments, development organizations, non-governmental organizations, academia, civil society and individuals to share data and analytics that can help achieving universal access to modern energy services. The database considers a variety of open, geospatial datasets of various context and granularity. **KTH Division of Energy Systems (KTH-dES), formerly known as KTH division of Energy Systems Analysis (KTH-dESA)**, contributes on a continuous basis by providing relevant datasets for electrification planning.

Open data and analytics for a sustainable energy future.

Search datasets

275 DATASETS | 15 ORGANIZATIONS | 101 COUNTRIES | 7 APPS

### Featured Apps

**Africa Electricity Grids Explorer**  
BY THE WORLD BANK GROUP

**Global Solar Atlas**  
BY THE WORLD BANK GROUP

[View all apps >](#)

## Indicative open libraries of GIS data

Over the past few years, KTH dES has been actively involved in the field of geospatial analysis. The following table presents a list of libraries and directories that provide access to open GIS data.

Source	Type	Link
Penn	World per region	<a href="http://guides.library.upenn.edu/content.php?pid=324392&amp;sid=2655131">http://guides.library.upenn.edu/content.php?pid=324392&amp;sid=2655131</a>
MIT	World per region	<a href="http://libguides.mit.edu/c.php?g=176295&amp;p=1161383">http://libguides.mit.edu/c.php?g=176295&amp;p=1161383</a>
EDEnextdata	World per region	<a href="https://www.edenextdata.com/?q=content/global-gis-datasets-links-0#Population%20Infrastructure%20Topography%20and%20Administration%20Data">https://www.edenextdata.com/?q=content/global-gis-datasets-links-0#Population%20Infrastructure%20Topography%20and%20Administration%20Data</a>
Stanford	World per region	<a href="https://library.stanford.edu/research/stanford-geospatial-center/data">https://library.stanford.edu/research/stanford-geospatial-center/data</a>
GIS Lounge	Finding GIS data	<a href="http://www.gislounge.com/data-and-gis-resources/">http://www.gislounge.com/data-and-gis-resources/</a>
dragons8mycat	Different countries	<a href="https://dragons8mycat.wordpress.com/gis-data-sources/">https://dragons8mycat.wordpress.com/gis-data-sources/</a>
rtwilson	Different types	<a href="http://freegisdata.rtwilson.com/">http://freegisdata.rtwilson.com/</a>
Planet OSM	Different types	<a href="http://planet.osm.org/">http://planet.osm.org/</a>
Berkeley	Different types	<a href="http://gif.berkeley.edu/resources/data_subject.html">http://gif.berkeley.edu/resources/data_subject.html</a>
Kings College	Different types	<a href="http://www.policysupport.org/waterworld">http://www.policysupport.org/waterworld</a>
CSRC	Different types	<a href="http://rslab.sr.unh.edu/gdatalinks.html">http://rslab.sr.unh.edu/gdatalinks.html</a>
Data Discovery Center	Different types	<a href="http://ddc.unh.edu/">http://ddc.unh.edu/</a>
Spatial Hydrology	Different types	<a href="http://www.spatialhydrology.com/datawarehouse.html">http://www.spatialhydrology.com/datawarehouse.html</a>
Africa Information Highway	Different types	<a href="http://dataportal.opendataforafrica.org/">http://dataportal.opendataforafrica.org/</a>
The Humanitarian Data Exchange	Different types	<a href="https://data.humdata.org/">https://data.humdata.org/</a>

## Country specific databases

With geospatial analysis gaining momentum in many research areas, many countries have set up their own geo-databases in an effort to facilitate interdisciplinary research activities under a geospatial context. Here are few examples:

Country	Source
Bolivia	<a href="http://geo.gob.bo/#viewer">http://geo.gob.bo/#viewer</a>
Brazil	<a href="http://www.ibge.gov.br/english/geociencias/default_prod.shtm#REC_NAT">http://www.ibge.gov.br/english/geociencias/default_prod.shtm#REC_NAT</a>
East Timor	<a href="http://goleaddog.com/gis-map/asia/timor-leste/">http://goleaddog.com/gis-map/asia/timor-leste/</a>
Malawi	<a href="http://www.masdap.mw/">http://www.masdap.mw/</a>
Namibia	<a href="http://www.uni-koeln.de/sfb389/e/e1/download/atlas_namibia/main_namibia_atlas.html">http://www.uni-koeln.de/sfb389/e/e1/download/atlas_namibia/main_namibia_atlas.html</a>
Nepal	<a href="http://geoportal.icimod.org/">http://geoportal.icimod.org/</a>
Russia	<a href="http://gis-lab.info/qa/vmap0-eng.html">http://gis-lab.info/qa/vmap0-eng.html</a>

### 2.4.3 GIS data in OnSSET

OnSSET is a GIS-based tool and therefore requires data in a geographical format. In the context of the power sector, necessary data includes those on current and planned infrastructure (electric grid networks, road networks, power plants, industry, public facilities), population characteristics (distribution, location), economic and industrial activity, and local renewable energy flows. The table below lists all layers required for an OnSSET analysis.

#	Dataset	Type	Description
1	Population density & distribution	Raster	Spatial identification and quantification of the current (base year) population. This dataset sets the basis of the ONSSET analysis as it is directly connected with the electricity demand and the assignment of energy access goals.
2	Administrative boundaries	Polygon	Delineates the boundaries of the analysis.
3	Existing HV network (Optional)	Line shape-file	Used to identify and spatially calibrate the currently electrified/non-electrified population. This is layer is optional.
4	Power Substations (Optional)	Point shape-file	Current Substation infrastructure used to identify and spatially calibrate the currently electrified/non-electrified population. It is also used in order to specify grid extension suitability. This is layer is optional.
5	Roads (Optional)	Line shape-file	Current Road infrastructure used to,identify and spatially calibrate the currently electrified/non-electrified population. It is also used in order to specify grid extension suitability. This is layer is optional.
6	Planned HV network (Optional)	Point shape-file	Represents the future plans for the extension of the national electric grid. It also includes extension to current/future substations, power plants, mines and queries. This is layer is optional.
7	Existing MV network (Optional)	Line shape-file	Used to identify and spatially calibrate the currently electrified/non-electrified population. This is layer is optional.
8	Planned MV network (Optional)	Point shape-file	Represents the future plans for the extension of the national electric grid. This is layer is optional.
9	Night-time lights	Raster	Dataset used to,identify and spatially calibrate the currently electrified/non-electrified population.
10	GHI	Raster	Provide information about the Global Horizontal Irradiation (kWh/m2/year) over an area. This is later used to identify the availability/suitability of Photovoltaic systems.
11	Wind speed	Raster	Provide information about the wind velocity (m/sec) over an area. This is later used to identify the availability/suitability of wind power (using Capacity factors).
12	Hydro power potential (Optional)	Point shape-file	Points showing potential mini/small hydropower potential. Dataset developed by KTH dESA including environmental, social and topological restrictions and provides power availability in each identified point. Other sources can be used but should also provide such information to reassure the proper model function. This is layer is optional.
13	Travel time	Raster	Visualizes spatially the travel time required to reach from any individual cell to the closest town with population more than 50,000 people.
14	Elevation Map	Raster	Filled DEM maps are use in a number of processes in the analysis (Energy potentials, restriction zones, grid extension suitability map etc.).
15	Land Cover	Raster	Land cover maps are use in a number of processes in the analysis (Energy potentials, restriction zones, grid extension suitability map etc.).
16	Service transformers	Point shape-file	Current Transformer infrastructure used to identify and spatially calibrate the currently electrified/non-electrified population. This is layer is optional.

**Note:**

- Before a model can be built, one must acquire the layers of data outlined above.
- You are recommended to use all the layers listed in the table above, but some of the are optional and can be omitted (see table above)

More often than not, each layer must be acquired on its own. The final outcome is a .csv-file conveying all the information necessary to initiate an OnSSET electrification analysis.

---

## GIS basic datasets

### Administrative boundaries

Cov- erage	Type	Resolution	Year	Source	Link
World	shape- file	Counties,provinces, departments, bibhag, bundeslander, daerah istimewa, fivondronana,,krong, landsvæðun, opština, sous-préfectures, counties & thana	2011	GADM	<a href="https://gadm.org/">https://gadm.org/</a>
World,( & shape- per coun- try)	shape- file	Countries	2011	DIVA- GIS	<a href="http://www.diva-gis.org/Data">http://www.diva-gis.org/Data</a>
Eu- rope	geo- database/ shapefile	Countries, provinces	2013	Eu- ro- stat	<a href="http://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/administrative-units-statistical-units">http://ec.europa.eu/eurostat/web/gisco/geodata/reference-data/administrative-units-statistical-units</a>

**Population data**

Coverage	Type	Resolution	Year	Source	Link
World	Various	1 arc-second	(depending on country)	HDX	<a href="https://data.humdata.org/organization/facebook">https://data.humdata.org/organization/facebook</a>
World	raster	250 meter and 1 km	1975, 1990, 2000, 2015	Global Human Settlement Layer	<a href="https://ghsl.jrc.ec.europa.eu/">https://ghsl.jrc.ec.europa.eu/</a>
Africa, Asia, America	Raster	100 m grid cells	(depending on country)	Worldpop	<a href="https://www.worldpop.org/geodata/listing?id=29">https://www.worldpop.org/geodata/listing?id=29</a>
World	grid	2.5 arc-minute grid cells	90/95/00	SEDAC	<a href="http://sedac.ciesin.columbia.edu/data/set/gpw-v3-population-density/data-download">http://sedac.ciesin.columbia.edu/data/set/gpw-v3-population-density/data-download</a>
World	shape-file, raster (grid)	2.5 arc-minute grid cells	2000	UNEP	<a href="http://geodata.grid.unep.ch/results.php">http://geodata.grid.unep.ch/results.php</a>
Ghana, Haiti, Malawi, South Africa, Sri Lanka	raster (grid)	1 arc-second	2015	CIESIN	<a href="https://ciesin.columbia.edu/data/hrsl/">https://ciesin.columbia.edu/data/hrsl/</a>
World	Various	Various	2016	dhsprogram	<a href="http://spatialdata.dhsprogram.com/home/">http://spatialdata.dhsprogram.com/home/</a>

**Transmission lines data**

UK	shape-file	Power transmission lines, underground cables, stations etc.	na	National Grid	<a href="http://www2.nationalgrid.com/uk/services/land-and-development/planning-authority/shape-files/">http://www2.nationalgrid.com/uk/services/land-and-development/planning-authority/shape-files/</a>
US	raster	100 m grid cells	2015	ArcGIS online	<a href="http://www.arcgis.com/home/item.html?id=918e6d9b1cc84d15ba13e911d18a0c5e">http://www.arcgis.com/home/item.html?id=918e6d9b1cc84d15ba13e911d18a0c5e</a>
World	OSM potential	points or polylines	2015	OSM of various mirrors	
World	From Vmap level 0	Power lines and utilities	na	Can be downloaded from:	<a href="http://gis-lab.info/qa/vmap0-eng.html">http://gis-lab.info/qa/vmap0-eng.html</a>

**Power plants location data**

Coverage	Type	Resolution	Year	Source	Link
World	shapefile (4 levels)	Generators, substations, masts	2009	Vmap level 0	<a href="http://gis-lab.info/qa/vmap0-eng.html">http://gis-lab.info/qa/vmap0-eng.html</a>

**Elevation**

Coverage	Type	Resolution	Year	Source	Link
World	geoTIFF	30 m spatial resolution	2009	METI Japan, NASA	<a href="http://www.jspacesystems.or.jp/ersdac/GDEM/E/2.html">http://www.jspacesystems.or.jp/ersdac/GDEM/E/2.html</a>
World	geoTIFF	30 m posting, 1x1 degree tiles	2009, 2011	METI Japan, NASA	<a href="https://asterweb.jpl.nasa.gov/gdem.asp">https://asterweb.jpl.nasa.gov/gdem.asp</a>
World	.bil and/or .tif	15 arcseconds/30arcseconds	various	ISCGM	<a href="https://globalmaps.github.io/">https://globalmaps.github.io/</a>
World	GeoTIFF	16 arcseconds/30arcseconds	various	NOOA	<a href="http://www.ngdc.noaa.gov/mgg/topo/gltiles.html">http://www.ngdc.noaa.gov/mgg/topo/gltiles.html</a>
World	GeoTIFF	17 arcseconds/30arcseconds	various	DGADV	<a href="http://www.dgadv.com/dowdem/">http://www.dgadv.com/dowdem/</a>
World + Arctic areas	GeoTIFF	30 arcseconds	various	WebGIS	<a href="http://www.webgis.com/terr_world.html">http://www.webgis.com/terr_world.html</a>

**Travel time to major cities**

Coverage	Type	Resolution	Year	Source	Link
World	ESRI grid	30 arc sec	2008 (data from 2000)	Joint Research Center EU	<a href="http://forobs.jrc.ec.europa.eu/products/gam/download.php">http://forobs.jrc.ec.europa.eu/products/gam/download.php</a>
Africa (sub-Saharan)	csv, ASCII raster, GeoTIFF	5 arc sec	2010	Harvest Choice	<a href="https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/YKDWJD">https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/YKDWJD</a>
World	Raster, GeoTIFF	5 arc sec	2015	Univeristy of Oxford	<a href="https://map.ox.ac.uk/explorer/#/explorer">https://map.ox.ac.uk/explorer/#/explorer</a>

## Mining and Quarrying

Coverage	Type	Resolution	Year	Source	Link
USA	Shapefile, csv, KML, KMZ	Active mines and mineral plants in the US	2003	USGS	<a href="http://mrddata.usgs.gov/mineplant/">http://mrddata.usgs.gov/mineplant/</a>
World	Shapefile, dBase, HTML, Tab text, csv, Google earth	points	2012-2013		<a href="http://minerals.usgs.gov/minerals/pubs/country/?utm_source=feedburner&amp;utm_medium=feed&amp;utm_campaign=Feed%3A+usgs_mpubs+%28USGS+Minerals+Periodicals%29">http://minerals.usgs.gov/minerals/pubs/country/?utm_source=feedburner&amp;utm_medium=feed&amp;utm_campaign=Feed%3A+usgs_mpubs+%28USGS+Minerals+Periodicals%29</a>
					<a href="http://mrddata.usgs.gov/mineral-resources/minfac.html">http://mrddata.usgs.gov/mineral-resources/minfac.html</a>
					<a href="http://mrddata.usgs.gov/mineral-operations/">http://mrddata.usgs.gov/mineral-operations/</a>

## Solar

Coverage	Type	Resolution	Year	Source	Link
World	ESRI ASCII GRID, GeoTIFF	250 m	2017	SolarGIS	<a href="https://globalsolaratlas.info/">https://globalsolaratlas.info/</a>

## Wind

Coverage	Type	Resolution	Year	Source	Link
World	Geo-TIFF	250m	2018	Technological University of Denmark	<a href="https://globalwindatlas.info/">https://globalwindatlas.info/</a>

Land cover

Coverage	Type	Resolution	Year	Source	Link
World	HDF-EOS	500 m	2001-2018	NASA-MODIS	<a href="https://lpdaac.usgs.gov/products/mcd12q1v006/">https://lpdaac.usgs.gov/products/mcd12q1v006/</a>
World	CI Land cover - raster	300 m	time series from 1992 to 2015	ESA	<a href="http://maps.elie.ucl.ac.be/CCI/viewer/">http://maps.elie.ucl.ac.be/CCI/viewer/</a>
World	GeoTiff, Google earth, jpeg,png	1-0.1 degrees	2001-2010	NASA-NEO	<a href="http://neo.sci.gsfc.nasa.gov/view.php?datasetId=MCD12C1_T1">http://neo.sci.gsfc.nasa.gov/view.php?datasetId=MCD12C1_T1</a>
World	Raster, csv	0.0028 - 0.0083 degrees	2000, 2005, 2010	ESA-ENVISAT	<a href="http://maps.elie.ucl.ac.be/CCI/viewer/index.php">http://maps.elie.ucl.ac.be/CCI/viewer/index.php</a>
World/Protected areas	Shapefile, KML, csv	na	2014	Protected planet	<a href="http://www.protectedplanet.net/">http://www.protectedplanet.net/</a>
World	various	various	2015	Global Land Cover Facility	<a href="http://landcover.org/data/">http://landcover.org/data/</a>
World	Rasters for: Costal areas, Cultivated areas, Forests, Mountains, Islands, Inland waters etc.	0.00833 degrees	2000	SEDAC	<a href="http://sedac.ciesin.columbia.edu/data/set/ma-ecosystems/data-download">http://sedac.ciesin.columbia.edu/data/set/ma-ecosystems/data-download</a>
World	Raster for croplands	0.0833 degrees	2000	SEDAC	<a href="http://sedac.ciesin.columbia.edu/data/set/aglands-croplands-2000/data-download">http://sedac.ciesin.columbia.edu/data/set/aglands-croplands-2000/data-download</a>
World	Various Rasters on Land Use	various	1990-2010	Nelson Institute	<a href="http://nelson.wisc.edu/sage/data-and-models/datasets.php">http://nelson.wisc.edu/sage/data-and-models/datasets.php</a>
World	Soil type	various	na	Worldmap.Harvard	<a href="http://worldmap.harvard.edu/data/geonode:DSMW_RdY">http://worldmap.harvard.edu/data/geonode:DSMW_RdY</a>
World	Various Rasters on Land Use	various	1980-2014	EarthStat	<a href="http://www.earthstat.org/data-download/">http://www.earthstat.org/data-download/</a>

The model classifies the land cover in order to calculate the grid extension penalties. The default classification values are based on the MODIS dataset found [here](#), where the legend ranges from 1-17 with the values and corresponding land cover type can be seen below. If land cover data is retrieved from other data sources with different classification values they should be reclassified in GIS (using the Reclassify tool in ArcGIS or r.reclass in QGIS) to match those below. Alternatively changes can be made in the Python code instead. If this reclassification is not performed it may lead to an incorrect grid penalty factor or, if the highest values are above 17, an error message while running the code.

Value	Label
1	Evergreen Needleleaf forest
2	Evergreen Broadleaf forest
3	Deciduous Needleleaf forest
4	Deciduous Broadleaf forest
5	Mixed forest
6	Closed shrublands
7	Open shrublands
8	Woody savannas
9	Savannas
10	Grasslands
11	Permanent wetlands
12	Croplands
13	Urban and built-up
14	Cropland/Natural vegetation mosaic
15	Permanent snow and ice
16	Barren
17	Water bodies

### Others

Coverage	Type	Resolution	Year	Source	Link
World	Coast Lines, oceans	Physical vectors, ESRI shapefiles, GeoTIFF (1:10, 1:50 and 1:110 m)	2015	Natural Earth	<a href="http://www.naturalearthdata.com/downloads/">http://www.naturalearthdata.com/downloads/</a>
World	Climate data	30 arc seconds and 2.5/5/10 arc minutes	na	WorldClim	<a href="http://www.worldclim.org/">http://www.worldclim.org/</a>
World/USA	Climate change scenarios	various	na	na	<a href="https://gisclimatechange.ucar.edu/">https://gisclimatechange.ucar.edu/</a>
World/Australia	Water and Landscape Dynamics	0.05 to 1 degrees	1979-2012	Australian National University	<a href="http://www.wenfo.org/wald/data-software/">http://www.wenfo.org/wald/data-software/</a>
Open Street Map (OSM) - Osmosis	osm.pbf	depending on mirror source	up to date	NOAA	<a href="http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html">http://ngdc.noaa.gov/eog/dmsp/downloadV4composites.html</a>
Nighttime lights	Raster file	0.0042 degrees	2012-2020	na	<a href="https://eogdata.mines.edu/download_dnb_composites.html">https://eogdata.mines.edu/download_dnb_composites.html</a>
Africa information Highway	various	vectors	various	AfDB	<a href="http://dataportal.opendataforafrica.org/">http://dataportal.opendataforafrica.org/</a>
World	Climate data	various	various	Oregon State University	<a href="http://globalclimatedata.org/">http://globalclimatedata.org/</a>

## Methodology for Open Street Map data and Osmosis

---

### Note:

- Open Street Map (OSM) is a collaborative project that intends to provide free and open access data used in mapping the world. This document aims at describing in brief the methodology used in order to obtain OSM data and transform them in compatible and useful information with the use of Osmosis and QGIS.
  - To begin with, bulk download of updated OSM data can be performed through the Planet OSM: <http://planet.osm.org/>.
  - The files can be downloaded as .xml and .pbf format. However, due to the large volume of data there are various mirrors/extracts that provide access to masked data for different regions of the planet. More information can be found here: <http://wiki.openstreetmap.org/wiki/Planet.osm#Downloading>. In previous cases Geofabrik.de where used successfully (<https://download.geofabrik.de/africa.html>).
  - From Geofabrik, data can be downloaded per region in .pbf format. In the latest version of QGIS it is possible to insert this data directly by simply dragging the file onto the QGIS window. However, since the files are usually very large it is recommended to transform the .pbf into a spatialite database.
  - To do this transformation open up the OSGeo shell following with your installation, navigate to the folder in which you have your .pbf file (by typing `cd [folder path]`) and enter the following line: **ogr2ogr -f SQLite X.sqlite Y.pbf** (**Note!** change **X** to the name you want to use for your spatialite database and **Y** to the name of your downloaded .pbf file)
  - Once This transformation is finished (it may take some time) drag the resulting .sqlite file into QGIS and work with it instead of the .pbf file.
  - OSM data provide access to a tremendous amount of information of various types. Feel free to explore the potential and share the results with an enthusiastic community.
- 

## 2.5 GIS data preparation

Once all necessary layers have been successfully acquired, the user would need to prepare the datasets for their input into the OnSSET model. This requires the creation of a .csv file. There are four steps that need to be undertaken to process the GIS data so that it can be used for an OnSSET analysis.

### 2.5.1 Step 1. Proper data types and coordinate system

In this first step the user would need to secure all the datasets. Before starting the analysis make sure that all datasets have the same coordinate system (preferably **World Geodetic Datum 1984 (WGS84)**) You can check the coordinate system of your layers by importing them into QGIS and then right-clicking on them and open the **Properties** window. In the Properties window go to the **Information** tab, here the coordinate system used is listed under *CRS* for both rasters and vectors.

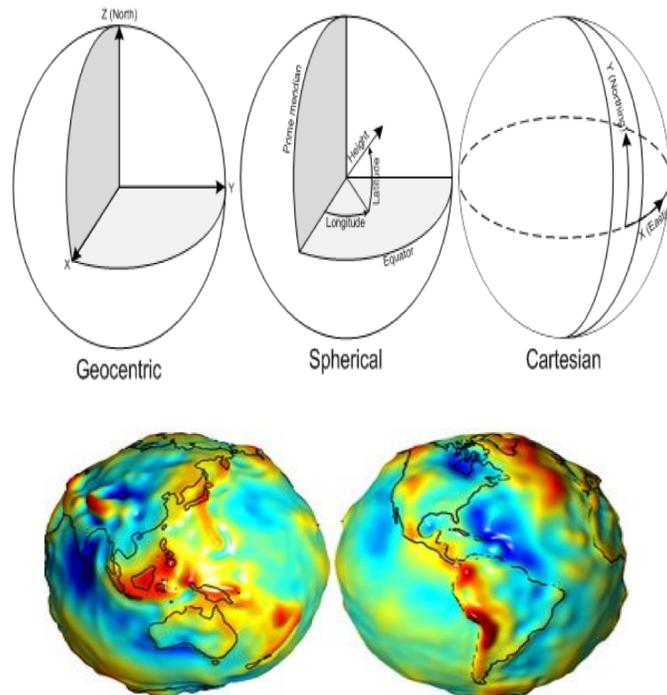
## 2.5.2 Step 2. Layer projection

In this step the user would need to determine the projection system he/she wish to use. Projection systems always distort the datasets and the system chosen should be one that minimizes this distortion. **Do not** manually project the datasets yourself (the extraction plugin presented below does this for you) it is however good to have an idea of which system to use before starting to work with the datasets. Here follows a few important key aspects:

**Projection** is the systematic transformation of the latitude and longitude of a location into a pair of two dimensional coordinates or else the position of this location on a plane (flat) surface. A projection is necessary every time a map is created and all map projections distort the surface in some fashion.

**Coordinate System:** Simply put, it is a way of describing a spatial property relative to a center.

**Datum:** The center and orientation of the ellipsoid



## 2.5.3 Step 3. Generate population clusters

Once the previous steps are finished we can generate our population clusters. These clusters represents settlements and will set the base of your OnSSET analysis. To generate the clusters you need the three following dataset: Administrative boundaries (polygon vector), Population (raster) and night-time lights (raster). In order to facilitate the cluster generation KTH-dES has developed QGIS based plugins available [here](#)

In the repository go to the **Plugin** folder and select the option that you wish to use. Currently there are four options to choose from:

- Option 1. Plugin developed for QGIS 3.10 (the latest stable version) and works with HRSL
- Option 2. Plugin developed for QGIS 3.4 and works with HRSL
- Option 3. Plugin developed for QGIS 3.2 and works with HRSL
- Option 4. Plugin developed for QGIS 3.2 and works with GHS

Once you have selected the option you wish to use, download the file named hrsl\_clustering.zip. In order to install and use the plugin please refer to the instructions published in its [repository](#).

## 2.5.4 Step 4. Generate the OnSSET input file

Once the clusters are generated, we can combine all layers together into a single table. The principle is simple. We will use the clusters from **Step 3** in order to create a base table. Every row in this table represents a population settlement. Then, we will adhere one by one all the layers into this table so that every row (settlement) acquires its specific characteristics based on its location. One can perform the process manually by identifying the tools in the GIS environment of his/her preference.

In order to facilitate the process KTH-dES has prepared a set of QGIS plugins that can be used. The plugin is available in the **Plugin** folder in the following [repository](#).

There are three options:

- Option 1. Works for QGIS 3.10 (latest LTR)
- Option 2. Works for QGIS 3.4
- Option 3. Works for QGIS 3.2

Choose the one that suits you, download the corresponding .zip-file named gep\_onsset and follow the instructions in its [repository](#) in order to install and run the plugin.

## 2.5.5 GIS country file

The table below shows all the parameters that should be sampled and put into the csv file representing the study area.

Parameter	Description
Country	Name of the country
Nighthlights	Maximum light intensity observed in cluster
Pop	Population of cluster
id	Id of cluster, important when generating maps
GridCellArea	Area of each cluster (km <sup>2</sup> )
ElecPop	Population that lives in areas with visible night-time lights
WindVel	Wind speed (m/s)
GHI	Global Horizontal Irradiation (kWh/m <sup>2</sup> /year)
TravelHours	Distance to the nearest town (hours)
ResidentialDemandTierCustom	Indicative residential electricity demand target
Elevation	Elevation from sea level (m)
Slope	Ground surface slope gradient (degrees)
LandCover	Type of land cover as defined by the source data
CurrentHVLineDist	Distance to the closest existing HV line (km)
CurrentMVLineDist	Distance to the closest existing MV line (km)
PlannedHVLineDist	Distance to the closest planned HV line (km)
PlannedMVLineDist	Distance to the closest planned MV line (km)
TransformerDist	Distance from closest existing transformers (km)
SubstationDist	Distance from the existing sub-stations (km)
RoadDist	Distance from the existing road network (km)
HydropowerDist	Distance from closest identified hydropower potential (km)
Hydropower	Closest hydropower technical potential identified
HydropowerFID	ID of the nearest hydropower potential

continues on next page

Table 1 – continued from previous page

Parameter	Description
X_deg	Longitude
Y_deg	Latitude
IsUrban	All 0 after extraction, urban/rural split gets assigned in the algorithm
PerCapitaDemand	Indicative residential electricity demand target
HealthDemand	Indicative electricity demand for health
EducationDemand	Indicative electricity demand for educational facilities
AgriDemand	Indicative electricity demand for agricultural processes
ElectrificationOrder	Indicates order of electrification; retrieved by grid extension algorithm; default =0
Conflict	Indicator of level of conflict (default =0; otherwise option 1-4)
CommercialDemand	Indicative electricity demand target for commercial activity
ResidentialDemandTier1	Indicative residential electricity demand target equal to Tier 1
ResidentialDemandTier2	Indicative residential electricity demand target equal to Tier 2
ResidentialDemandTier3	Indicative residential electricity demand target equal to Tier 3
ResidentialDemandTier4	Indicative residential electricity demand target equal to Tier 4
ResidentialDemandTier5	Indicative residential electricity demand target equal to Tier 5

**Note:** It is very important that the columns in the csv-file are named exactly as they are named in the **Parameter**-column in the table above.

## 2.6 Techno-economic and demographic parameter

### 2.6.1 Input model parameters

In order to run the OnSSET analysis certain input parameters have to be used. Some of these are specified directly in the Python code while others are specified in an excel specs file.

#### Python input parameters

The values for the parameters in the table below are already filled in the OnSSET Python files. However, some of these may need to be updated with new or country specific values.

Parameter	Description	Unit
BASE_YEAR	The base year of the analysis. Note that this parameter is highly related to the input GIS data, which should, if possible, be calibrated towards this year (e.g. the population distribution map should represent the base year values).	•
END_YEAR	The final year of the analysis.	•
Scenario	The input value here represents the annual electricity consumption per household that is expected to be achieved, by the end year.	kWh/hh/year
sa_diesel_capital_cost	The capital cost (per capacity unit) of a stand- alone diesel generator.	\$/kW
sa_pv_capital_cost	The capital cost (per capacity unit) of a stand- alone PV module.	\$/kW
mg_diesel_capital_cost	The capital cost (per capacity unit) of a mini grid diesel generator.	\$/kW
mg_pv_capital_cost	The capital cost (per capacity unit) of a mini grid PV system.	\$/kW
mg_wind_capital_cost	The capital cost (per capacity unit) of a mini grid wind powered system.	\$/kW
mg_hydro_capital_cost	The capital cost (per capacity unit) of a mini grid hydropower system.	\$/kW
existing_grid_cost_ratio	Incremental cost increase for extension of the grid from an electrified settlement to an un-electrified one. Default value set at 10%.	ratio
discount_rate	The discount rate applied to different technology configuration choices throughout the period of analysis	ratio
grid_cell_area	The size of each grid cell used in the study area. This value typically ranges between 1 and 10 km <sup>2</sup>	km <sup>2</sup>

## 2.6.2 Discount rate

The discount rate is a rather important factor when doing modelling and planning for energy systems. This has to do with the fact that the discount rate is aimed at the variable cost. In other words a high discount rate tends to favour technologies with high variable costs (e.g. technologies that use non-renewable fuels), while lower discount rates favours technologies with high capital investment (e.g. renewables). This makes it a very good variable for running sensitivity analysis. For realistic values of the discount rate study the literature regarding your study area.

### 2.6.3 Off-grid technology costs

When calculating the total cost for the off-grid technologies (all the stand-alone and mini-grid systems) there are several different costs that need to be taken into account. The technology is a fixed cost that need to be taken into account, but the variable costs necessary for generating electricity also have to be taken into account (e.g. fuel costs). The cost of installing the system also need to be taken into account. The connection to the grid is however excluded and hence not taken into consideration.

#### Specs-file input parameters

The inputs in the table below should be entered in an .xlsx file named CountrySpecs.xlsx (Replace Country with the name of the area studied)

Parameter   Description	Unit
Country	Enter the name of the study area •
Pop2015	The population of the selected area in the base year. People
UrbanRatio2015	The ratio of urban population in the selected area in the base year. Ratio
Pop2030	The projected population People
UrbanRatio2030	The ratio of urban population in the selected area in the end year. Ratio
NumPeoplePerHHRural	Number of people per household in rural areas People
NumPeoplePerHHUrban	Number of people per household in urban areas People
DieselPriceLow	The low price of diesel. Used in the calculation of the operational costs for diesel generators (in both mini grid and stand-alone systems). \$/liter
DieselPriceHigh	The high price of diesel. Used in the calculation of the operational costs for diesel generators (in both mini grid and stand-alone systems). \$/liter
GridPrice	The cost of which the national grid generates electricity. This value is based on the mix of technologies used in the country. \$/kWh
GridCapacityInvestmentCost	The investment required per unit of additional capacity for the national grid. This is an average value based on, the mix of technologies used in the country. \$/kWh
GridLosses	This value represents the country's average technical losses on transmission and distribution. Ratio
BaseToPeak	This value represents the ratio between the base and peak load in the selected country. It is used for sizing the necessary capacity to be installed per settlement in order to cover the respective demand. Ratio
ExistingGridCostRatio	Incremental cost increase for extension of the grid from an electrified settlement to an un-electrified one. Default value set at 10%. Ratio
MaxGridExtensionDist	The input parameter sets the maximum distance for which the grid can be extended in order to electrify a settlement due to techno-economic considerations. The default value in the model is 50km. km
ElecActual	The electrification rate in the selected area in the base year. What ratio of population is electrified. Ratio
MinNightLights	Night time light is an indicator [0-63]
34	retrieved by NASA's satellite imageries and it represents light intensity over a geographical area. The value ranges from 0 to 63 with higher value representing higher <b>Chapter 2. Contents</b>

The variables *MinNightLights*, *MaxGridDist*, *MaxRoadDist*, *PopCutOffRoundOne*, *PopCutOffRoundTwo* and *Urban-CutOff* are not found in the literature but rather defined by the user. When running the code these values are altered to find a combination that satisfies that the cells are assigned electrified and urban status matching the values entered at *ElecActual* and *UrbanRatio2015*.

All the values with **ratio** as the unit should take a value between 0 and 1 in the excel file.

## 2.6.4 Diesel pump price

For the diesel price the best option is to find the diesel price directly from the literature regarding your studyarea. In cases for which this is not possible you should try to find the crude oil price in your country. 1.15 is an empirical value used in order to estimate the diesel price (\$/liter) from the crude oil price(\$/liter). It is assumed that the price of diesel is 15% higher than the crude oil price. If you have difficulties finding the crude oil price in your area of interest you can refer to e.g. [http://www.oecd-ilibrary.org/energy/world-energy-outlook\\_20725302](http://www.oecd-ilibrary.org/energy/world-energy-outlook_20725302).

## 2.6.5 Household size

The household size is an important parameter in the electrification planning analysis as it affects the connection costs per household. These are calculated based on:

- a) the projected mean national household size values
- b) the existing and projected national, urban and rural populations
- c) the urban to rural household size ratio given in demographics and health country surveys.

---

**Note:** It is very important that the columns in the Specs-file are named exactly as they are named in the **Parameter**-column in the table above.

---

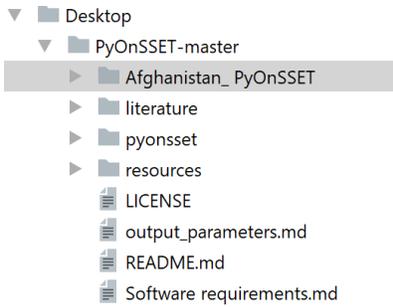
## 2.7 Running the analysis

### 2.7.1 Setting up PyCharm - Afghanistan Demo

1. Download the code from **GitHub** [here](#).

Select “Download ZIP” from the “Clone or Download” tab and save the zipped folder on your Desktop. Unzip it and open the “Afghanistan\_PyOnSSET” folder. If not already there, create a folder and name it “run”. This folder will contain the final output with the results.

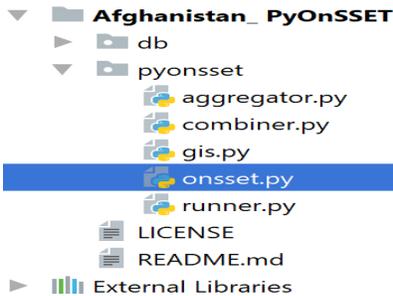
2. Download the “Afghanistan.csv” dataset from [here](#) (to be updated) and save it in “~/PyOnSSET-master/Afghanistan\_PyOnSSET/db”.
3. Once you have successfully completed all the previous steps, open **PyCharm** or any other Python IDE framework. Select “Open” and then select “~/Desktop/PyOnSSET-master/Afghanistan\_PyOnSSET” as your working directory as shown below:



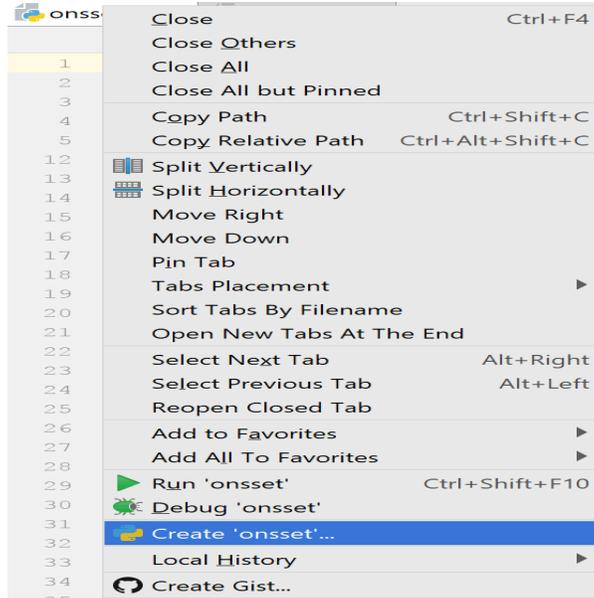
4. Go to *File*->*\*Default Settings\**->*\*Project Interpreter\** and from the drop-down list select the Anaconda directory which we installed in the previous steps. Once all the packages load, click Apply and OK.



5. Once you have set your working directory, select "pyonsset", and from the drop-down menu double-click on a) *onsset.py* and b) *runner.py* in that order, in order to open them.

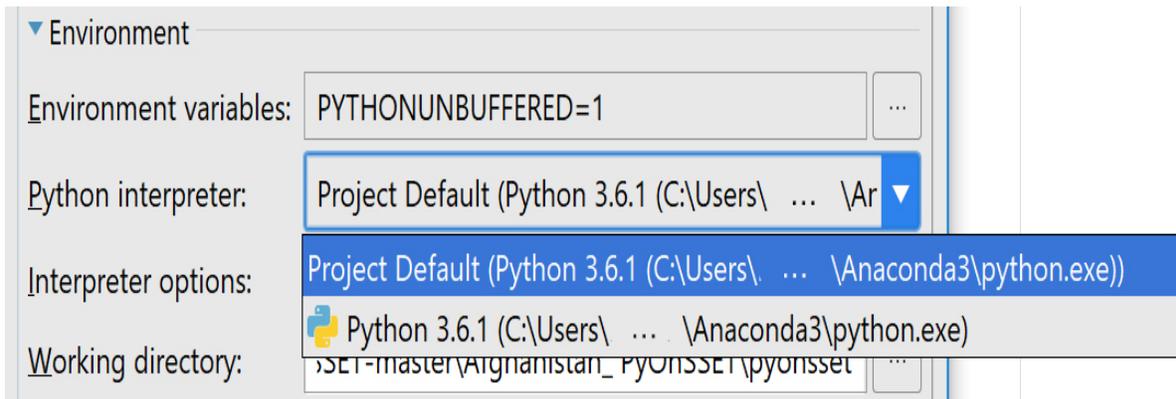


6. Now that you have opened "onsset.py", right-click on its tab and click on "Create/Select 'onsset'...".



**Note:**

- In the configuration window, make sure that the Anaconda interpreter is selected in the “Python interpreter” option before you click OK.



- You are now ready to run “onsset.py” by clicking the Run button at the upper-right corner of your window.



- Repeat the same steps explained above for “runner.py” only if you get a “Process finished with exit code 0” message in the built-in console. Otherwise, move to step 7 before you run “runner.py”.

7. After running a python script, the output is shown in the console built in PyCharm.

- If you get a “Process finished with exit code 0” message, that means that everything worked OK.
- Otherwise, the console shows messages with errors or warnings. In case you get a “Module Not Found Error” message, that means that a module or package is missing.

## 2.7.2 Installing/updating modules and packages

### 1. Using Anaconda

Search for “Anaconda Prompt” in your Windows Search. In the command line, type `conda install <package name>` and the installation/update will go through automatically. (Note: In the Proceed ([y]/[n])? prompt, type “y”).

e.g. installation of module “pyproj”

```

Anaconda Prompt
(C:\Users\... \Anaconda3) C:\Users\... >conda install pyproj
Fetching package metadata .....
Solving package specifications: .
Package plan for installation in environment C:\Users\... \Anaconda3:
The following NEW packages will be INSTALLED:
  pyproj: 1.9.5.1-py36_0
The following packages will be UPDATED:
  conda: 4.3.21-py36_0 --> 4.3.22-py36_0
Proceed ([y]/n)? y
pyproj-1.9.5.1 100% |#####| Time: 0:00:00 10.50 MB/s
conda-4.3.22-p 100% |#####| Time: 0:00:00 11.55 MB/s
    
```

### 2. pip

**pip** is a package management system used to install and manage software packages written in Python. Search for “Command Prompt” in your Windows Search. Using the command line, navigate to your Python directory and then type “pip” and enter. This will show you all the possible choices you have within pip.

- In order to update pip itself you can just type `pip install --upgrade pip`.

```

Administrator: Command Prompt
C:\Users\... \Local\Programs\Python\Python35-32\Scripts>pip install --upgrade pip
    
```

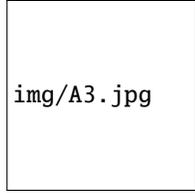
- In order to install any additional module required you can type `pip install <package name>`.  
e.g. installation of module “matplotlib” `pip install matplotlib`

```

Administrator: Command Prompt
C:\Users\... \Local\Programs\Python\Python35-32\Scripts>pip install matplotlib
    
```

### 3. Directly in PyCharm

Go to the project interpreter (*File>Settings>Project Interpreter*). Click on the green plus-sign in the upper right corner.



Search for the required package and click *Install Package*.

### 2.7.3 Running the code

There are two Python codes that are required to run the analysis; **onsset.py** and **runner.py**.

**onsset.py** defines most of the functions and calculations behind the analysis. Running this code however does not perform the analysis. **runner.py** is the code that actually executes the analysis. When running this code the user will be prompted to enter several inputs, explained below, and runner.py will then retrieve the functions from onsset.py.

#### Note:

##### Please make sure to:

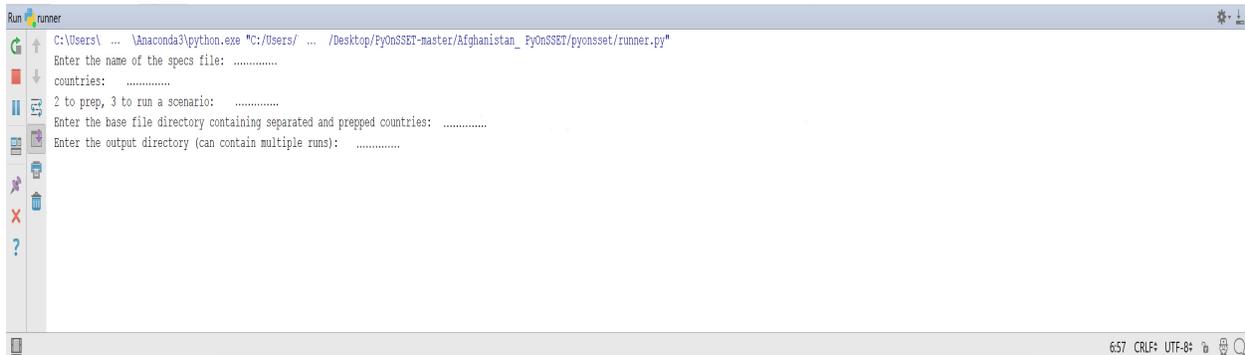
- Not edit or modify in any way your input/output .csv files unless required.
- Close any open .csv files before running the programs.

Two changes in the Python code may be required to be able to run the analysis.

Line 8 in runner.py may be changed from `from pyonsset.onsset import *` to `from onsset import *` if runner.py and onsset.py are saved in the same folder.

In runner.py line 11, change directory from db to the one where the .csv file containing the prepared GIS data and the specs file are stored. The line should then look like e.g. `os.chdir('C:/Desktop/Afghanistan_PyOnSSET/db')`.

OnSSET is then ready to be executed. Press run and enter the specified information when prompted. The runner will appear at the bottom of the window and look like:



```

Run runner
C:\Users\... \Anaconda3\python.exe "C:/Users/... /Desktop/PyOnSSET-master/Afghanistan_PyOnSSET/pyonsset/runner.py"
Enter the name of the specs file: .....
countries: .....
2 to prep, 3 to run a scenario: .....
Enter the base file directory containing separated and prepped countries: .....
Enter the output directory (can contain multiple runs): .....
657 CRLF: UTF-8:

```

First the program will ask for the name of the specs file. Enter the name of the specs file and press enter, and remember to include the file extension. The specs file should be an excel file, and the entry should therefore be e.g. `Specs.xlsx`.

Secondly the countries of the study are required. Enter the name of the country and press enter. If more than one country are included in the study enter the names of all countries separated by a blank space (Example: `Afghanistan Pakistan`)

**Note:** Python is sensitive to upper and lower case letters. Always make sure that names of files, directories and countries are matching exactly when running the code and entering information.

---

The name of the specs file and the country names need to be entered every time. The rest of the code has been divided into three parts with different functionalities. Next you will be asked to 'Enter 1 to split, 2 to prep, 3 to run a scenario: '. The function of the three options are to:

1. Split a csv file containing GIS data for multiple countries into csv files that each has the data for one of the countries. If the original .csv file is for one country only this step can be skipped.
2. Calculate some initial data needed for the electrification analysis, such as projected population, wind capacity factors etc. This step needs to be run only once to use OnSSET for one area.
3. Perform the electrification analysis. Once the previous steps have been run the third part can be run multiple times to examine scenarios with e.g. different electricity access targets or fuel costs.

**Step 1: Split** When prompted, enter the name of the file containing data for multiple countries and press enter. Example: `Countries.csv`. Again remember to include the file extension.

Next enter the directory where the new files should be saved. This can be the same directory as the file containing all the countries or a new one with separated countries only e.g. `C:/Desktop/Afghanistan_PyOnSSET/db/countryfiles`. A new directory is recommended to prevent overwriting of the original .csv file if it is named after a country.

OnSSET then creates new .csv files with data for each country. The files will be stored with the name of the country followed by the .csv extension, e.g. `Afghanistan.csv`.

**Step 2: Prep** The preparation step starts by prompting the user to enter the directory where the separated country files are saved. If selection 1 was executed the entry should be the one specified for output in selection 1 (`C:/Desktop/Afghanistan_PyOnSSET/db/countryfiles`), otherwise it is the directory where the original .csv file is stored (e.g. `C:/Desktop/Afghanistan_PyOnSSET/db`). If selection 1 has not been executed the csv file must be named `country.csv` (e.g. `Afghanistan.csv`) or OnSSET will not be able to find it.

For each country chosen initially the prepping will be executed and the .csv file with GIS data will be overwritten in the same folder with a new file with the same name (e.g. `Afghanistan.csv`) containing additional information. The specsfile is overwritten as well with updated data for *MaxGridDist*, *MinNightLights* and *MaxRoadDist* as well as *ElecModelled* which is the calculated value for the electrified population in the base year. After running step 2, open the specs file to check if *ElecModelled* is sufficiently close to *ElecActual*. If not, change the values of *MaxGridDist*, *MinNightLights*, *MaxRoadDist*, *PopCutOffRoundOne* and/or *PopCutOffRoundTwo* and run step 2 again until *ElecModelled* is sufficiently close to *ElecActual*.

**Step 3: Run scenario(s)** In the third selection the electrification scenarios are run. This part of the code need to be run once for every scenario. The user is first prompted to specify where the prepped files from selection 2 are saved. Enter the same as was entered in selection 2. Secondly the user is prompted to enter where the result files should be saved. The result files will save as *Scenario\_Tier\_Dieselchoice.csv* and *Scenario\_Tier\_Dieselchoice.csv* (Example: `Afghanistan_1_High.csv` for a Tier 1 analysis with high diesel price).

The following step is to enter the electricity target simply by typing the tier number between 1-5 and pressing enter for urban and rural areas respectively. The annual household electricity consumption is displayed in the terminal:

```

World Bank Tiers of Electricity Access
  1: 7.738 kWh/person/year
  2: 43.8 kWh/person/year
  3: 160.6 kWh/person/year
  4: 423.4 kWh/person/year
  5: 598.6 kWh/person/year
Enter the tier number for urban: 1
Enter the tier number for rural: 1
    
```

Next step is to choose between the high and low diesel price from the specs file. For the question Use high diesel value? <y/n> `` enter ``y to use the high diesel cost or n to use the low diesel cost.

The final input required before running the analysis is if a result file should be created for each country in the analysis or if a combined result file should be created: Combine countries into a single file? <y/n>. Enter y to combine or n to generate separate files. Following this step the electrification analysis is executed.

---

**Note:** Sometimes it may be of interest to run scenarios with the same energy tier and diesel price while changing other variables such as the discount rate. In a case with two runs with the same energy tier and diesel price the first result files will be overwritten. Remember to move or rename them before initiating the second run.

---

## 2.8 Analysis and visualizations

The output of the analysis are two csv files. One large file containing data for each geolocated cell and one smaller file with summarized data.

### 2.8.1 Output parameters

The following table displays the parameters included in the excel file that includes all the grid cells. The first 18 parameters come from the input csv file while the remaining stem from running the Python code.

Parameter	Description
Country	Name of the country
Pop	Population in base year
X	Longitude
Y	Latitude
GHI	Global Horizontal Irradiation (kWh/m2/year)
SolarRestriction	Defines if an areas is restricted to solar PV deployment (1: restricted, 0: non restricted)
WindVel	Wind speed (m/s)
TravelHours	Distance to the nearest town (hours)
NightLights	Nighttime light intensity (0-63)
Elevation	Elevation from sea level (m)
Slope	Ground surface slope gradient (degrees)
LandCover	Type of land cover as defined by the source data
GridDistCurrent	Distance from the existing electricity grid network (km)
GridDistPlan	Distance from the planned electricity grid network (km)
SubstationDist	Distance from the existing sub-stations (km)
RoadDist	Distance from the existing and planned road network (km)
HydropowerDist	Distance from identified hydropower stations (km)

Table 2 – continued from previous page

Parameter	Description
Hydropower	Closest hydropower technical potential identified
X_deg	Longitude in degrees
Y_deg	Latitude in degrees
RoadDistClassified	Classification of the distance from road network based on a suitability index ranging from 1 to 5 (1 implying the lowest suitability)
SubstationDistClassified	Classification of the distance from sub-stations based on a suitability index ranging from 1 to 5 (1 implying the lowest suitability)
LandCoverClassified	Classification of the land cover based on a suitability index ranging from 1 to 5 (1 implying the lowest suitability)
ElevationClassified	Classification of the elevation value based on a suitability index ranging from 1 to 5 (1 implying the lowest suitability)
SlopeClassified	Classification of the slope gradient based on a suitability index ranging from 1 to 5 (1 implying the lowest suitability)
GridClassification	Classification of the distance to the grid network based on a suitability index ranging from 1 to 5 (1 implying the lowest suitability)
GridPenalty	Multiplier of the grid expansion cost based on a suitability index deriving from a weighted aggregation of other factors
WindCF	Identified capacity factor of wind turbines (%)
Pop2015Act	Population in 2015
IsUrban	Percentage of urban population
Pop2030	Population in 2030
Elec2015	Electrification status of population in 2015 (0 un-electrified - 1 electrified)
NewConnections	Number of un-electrified population that is expected to get access to electricity by the end year
mg_hydro	Levelized cost of electricity achieved by mini grid hydropower
mg_pv	Levelized cost of electricity achieved by mini grid PV
mg_wind	Levelized cost of electricity achieved by mini grid wind turbines
mg_diesel	Levelized cost of electricity achieved by mini grid diesel
sa_diesel	Levelized cost of electricity achieved by stand-alone diesel
sa_pv	Levelized cost of electricity achieved by stand-alone PV
minimum_tech	Off grid technology that provides the lowest levelized cost of electricity
minimum_tech_lcoe	The lowest levelized cost of electricity selected between off grid technologies
Elec2030	Grid connection status in 2030 (1 implies grid connection - 0 implies off grid solution)
grid	Levelized cost of electricity achieved by the grid (99 implies grid un-suitability)
MinGridDist	Distance from the existing and planned grid network (km)
minimum_overall	Technology that provides the lowest levelized cost of electricity
minimum_overall_lcoe	The lowest levelized cost of electricity selected
minimum_overall_code	Code of technology selected (1:grid, 2:stand-alone diesel, 3:stand-alone PV, 4:Mini-grid diesel . 5: Mini-grid PV)
minimum_category	Type of optimal supply type (Grid – Mini grid – Stand-alone)
NewCapacity	Capacity requirement (kW)
InvestmentCost	Investment requirement (USD)

## 2.8.2 Summaries output

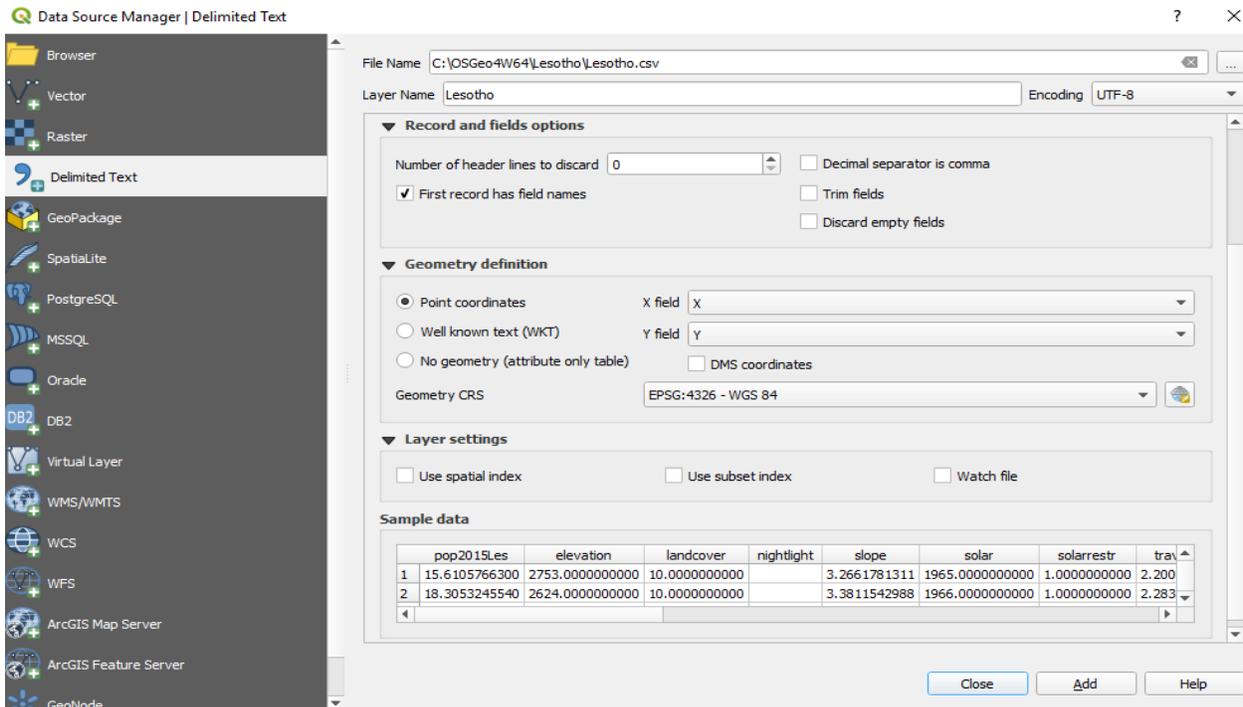
The values in the summaries file provide summaries for the whole country/study area.

Variable	Description	Unit
Population	The population served by each,technology in the end year.	people
New Connections	The number of newly electrified,population by each technology in the end year.	people
Capacity	The additional capacity required by,each technology to fully cover the demand in the end year.	kW
Investment	The investment required by each,technology to reach the electrification target in the end year.	\$
Technology mix	Generation mix as calculated by the,OnSSET analysis.	.
LCoE	Lowest LCoE achieved in each location,as calculated by the OnSSET analysis.	\$/kWh

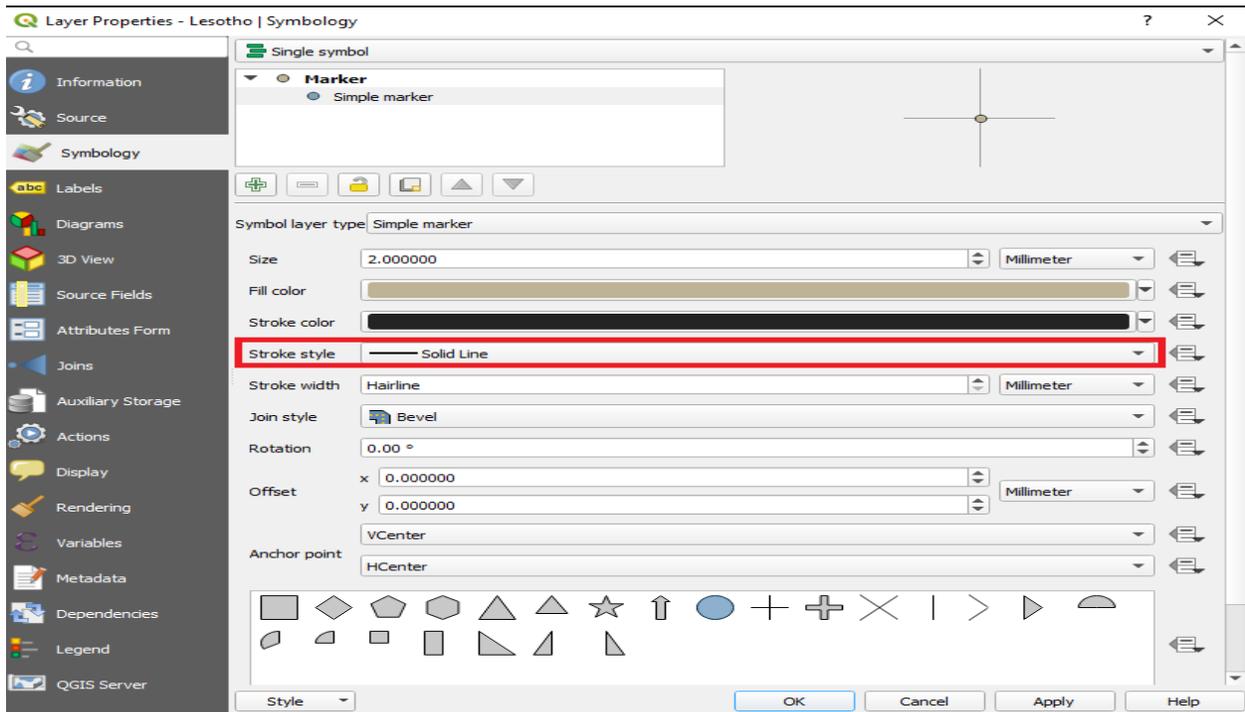
## 2.8.3 Maps output

### Creating maps in ArcGIS or other GIS software

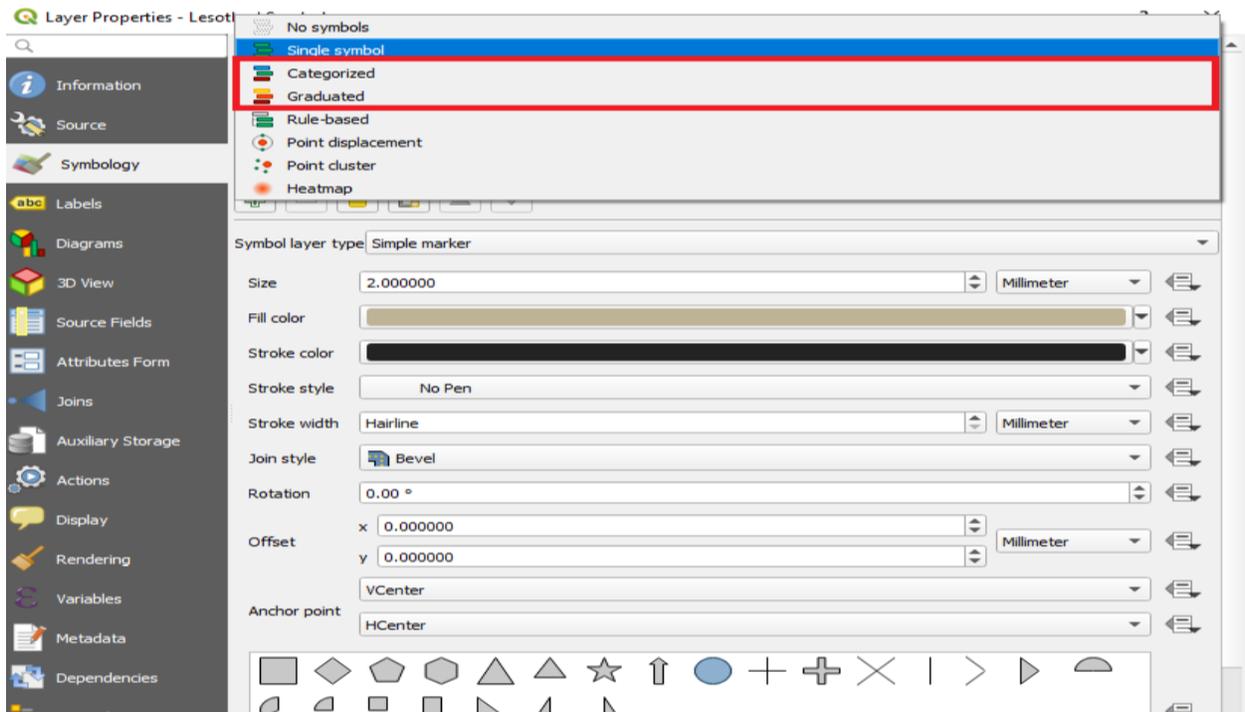
Visualization of the model outputs are an important part of a geospatial electrification analysis. This allows for an easy display of electrification strategies and enhanced understanding of the results. The most flexible way of visualizing the results are to create the maps using a GIS program. In QGIS, go to *Layer>Add Layer>Add Delimited Text Layer* and select the result .csv file containing the scenario results to be displayed. Next, choose the column with the coordinates as well as the projection system to generate a map. For the X and Y Field choose X and Y for coordinates in km or X\_deg and Y\_deg for coordinates in degrees.

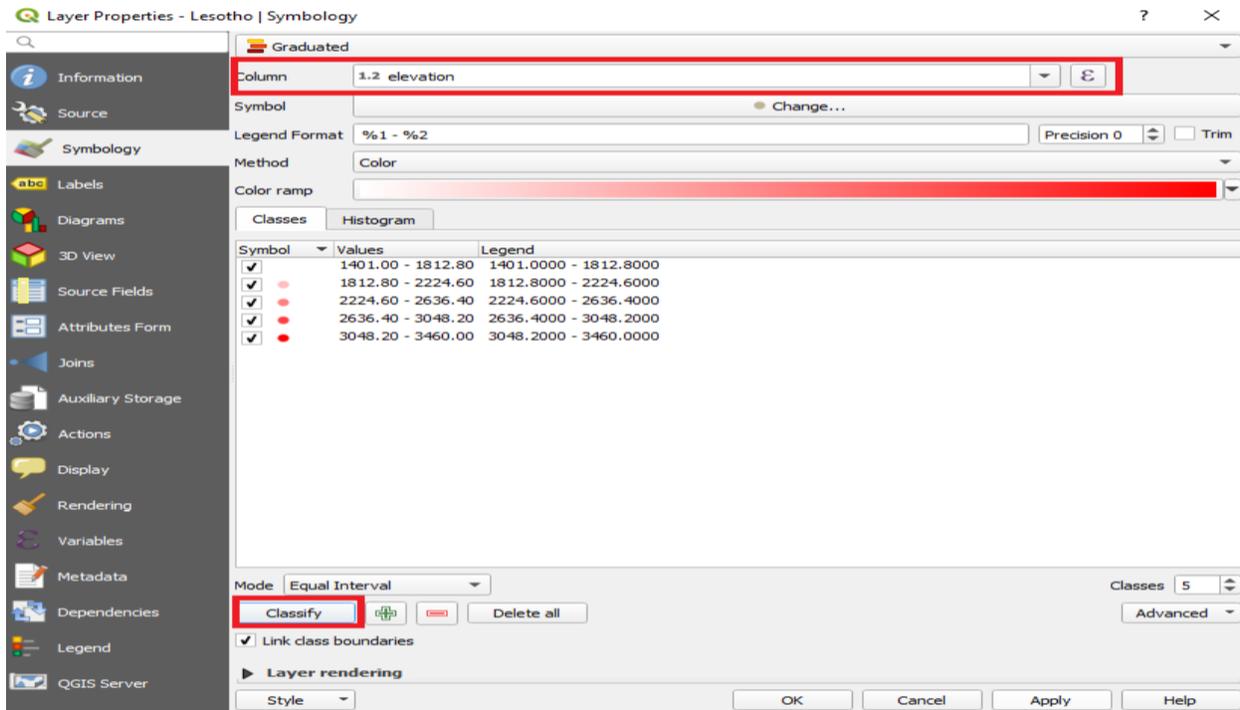


After doing this you get a point layer in your layer panel. Right-click on the layer and go to *Properties>Symbology*. Remove the border of all the circles.



And then you can choose how to visualise your data. Choose **Categorized** for discrete data and **Graduated** for continuous data. In column choose the data you want to visualise and then click on classify.





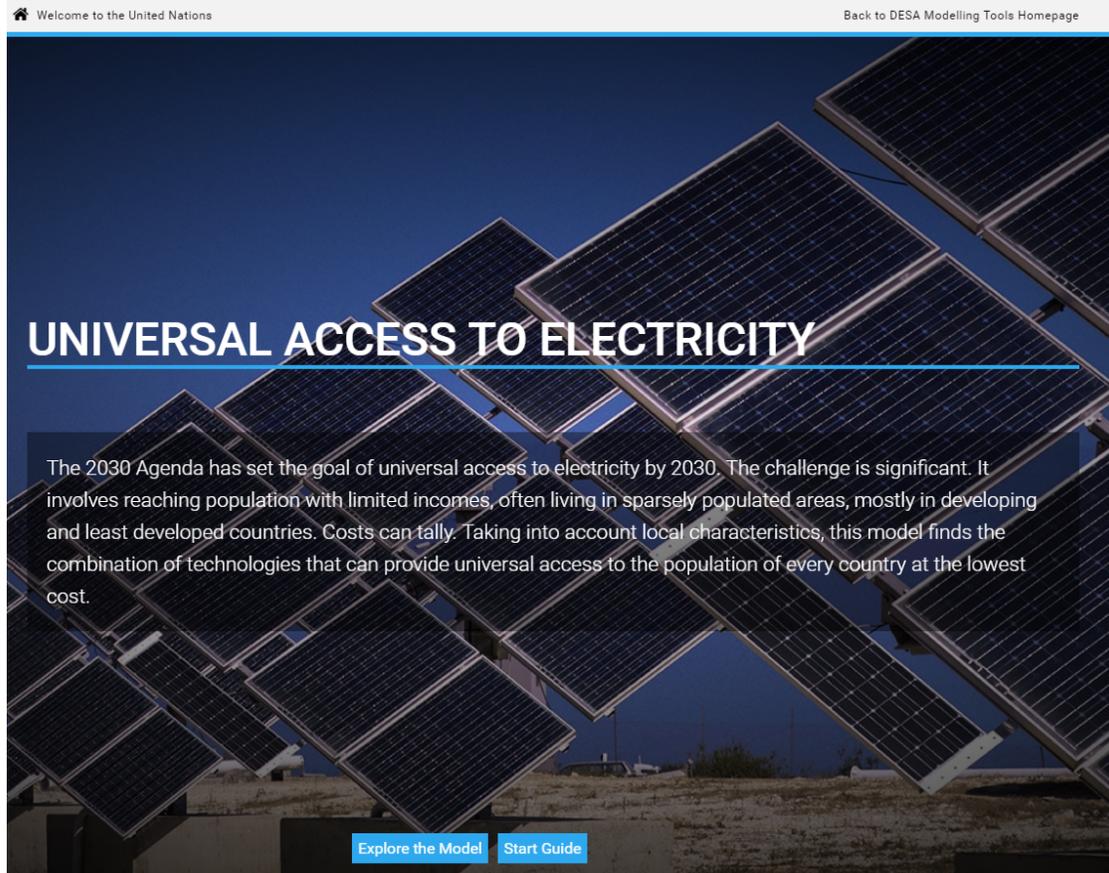
After these steps you can easily create your own maps and include north arrows, legends and scale bars by going to *Project>New Print Layout*

## 2.9 Online interfaces & Applications

For those who would like to just take a taste of an OnSSET electrification analysis, KTH dESA has collaborated international organizations in order to develop interactive platforms that allow a quick display of “pre-cooked” electrification results for a number of countries.

### 2.9.1 United Nations app

The online interface [Universal Electrification Access](#) is part of the modelling tools for sustainable development and is based on OnSSET. This interface is not a model but it provides the user easy access to the methodology behind OnSSET, the datasets used and key results -with regards to electrification planning - obtained for a set of 30 scenarios and model runs for 54 developing countries in Sub-Saharan Africa and Latin America.



### 2.9.2 World Bank app

A [Web-based application](#) developed by the World Bank, Energy Sector Management Assistance Program (ESMAP) and KTH Royal Institute of Technology in Stockholm, shows the dynamic of an OnSSET based electrification analysis. The tool considers a variety of open, geospatial data to determine the least-cost options for bringing electricity to areas as granular as 1 square km in size. So far three countries are included – Tanzania, Zambia, and Nigeria – and will be adapted for more.

Open data and analytics for a sustainable energy future.

Search datasets

275 DATASETS | 15 ORGANIZATIONS | 101 COUNTRIES | 7 APPS

### Featured Apps



**Africa Electricity Grids Explorer**  
BY THE WORLD BANK GROUP



**Global Solar Atlas**  
BY THE WORLD BANK GROUP

[View all apps >](#)

## 2.10 References

Since its development, OnSSET has been employed for a high number of studies presented at international conferences and published on peer reviewed journals. Here you find a list of the publications on OnSSET, which is going to be continuously updated.

### 2.10.1 Publications

The Role of Open Access Data in Geospatial Electrification Planning and the Achievement of SDG7. An OnSSET-Based Case Study for Malawi

A GIS-based approach for electrification planning—A case study on Nigeria

The benefits of geospatial planning in energy access – A case study on Ethiopia

A cost comparison of technology approaches for improving access to electricity services

A climate, land-use, energy and water nexus assessment of Bolivia

UN Sustainable development goals from a Climate Land Energy and Water perspective for Kenya

Lighting up the World at NASA-ESA International Workshop on Environment and Alternative Energy

Open Geospatial Data for Energy Access

Sustainable Development Goals: The EU at the global partnership forefront

Taming a Wicked Problem: Energy Access Planning from an Energy-Poor Perspective

Lighting up the World at European Geosciences Union General Assembly (to be updated)

## 2.10.2 Contribution

### energy specialists

In the policymaking arena, electrification planning is often captured by private consultants' analytical infrastructure. The open source nature and the modular structure of the developed tool allow users to build and modify the code according to their requirements and sources and to develop rapid cost effective electrification plans with minimal resource and computational requirements as compared to existing electrification planning tools that are time-, data- and resource- intensive.

### Education

OnSSET can be used by graduate students to enhance its functionalities and develop new applications in order to help address energy access issues. Enabling graduate students to build and iteratively develop formal energy planning models will impart this knowledge base to very wide range of energy market roles and positions.

### Community

OnSSET community welcomes professionals and experts from different levels: decision makers, policy officers, energy planners, developers of new model functionalities, programmers. All your valuable comments and contributions are welcomed here.

## 2.10.3 Partners





**WORLD BANK GROUP**



**Vetenskapsrådet**



## 2.11 Contact

Ask any questions on [OnSSET forum](#).

You will need just a few clicks to sign up to the forum. Find the link at the top-right corner of the [Optimus.community](#) page.

**Contact us:**

E-mail: [media@desa.kth.se](mailto:media@desa.kth.se) (tbu) or meet the team [here](#).



### 2.11.1 Feedback

Please post all comments and feedback you may have regarding this manual [here](#)..

### 2.11.2 Funding

OnSSET is a non-profit project. We rely on external funding and welcome partnerships. In this regard please contact us at [seap@desa.kth.se](mailto:seap@desa.kth.se).

### 2.11.3 Administration

#### dES KTH

##### Postal Address:

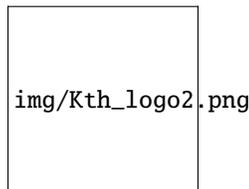
KTH - School of Industrial Engineering and Management

division of Energy Systems

Brinellvägen 68

SE-100 44 STOCKHOLM

<https://www.desa.kth.se>



## 2.12 FAQ

#### [Q] What prior knowledge is required for someone to use OnSSET?

[A] One should be familiar with the basic principles of energy planning and energy economics. Experience in working with Geographic Information Systems is also essential in order to build, analyse and visualize spatial datasets. Python skills are necessary if one would like to modify the code.

#### [Q] What type of renewable sources are used in an OnSSET model? Is it possible to add other sources (biomass etc.)?

[A] Currently, solar photovoltaics, wind turbines and mini and small hydropower plants are considered as off grid renewable energy technologies. The inclusion of bioenergy and hybrid technologies constitutes an ongoing addition.

#### [Q] How long does it take to complete an OnSSET electrification analysis for one country?

[A] This depends on the size of the country, on the resolution of the analysis, on the country specific characteristics of the grid infrastructure and on the electricity consumption level scenarios. Indicatively, to run Nigeria (area: 923,768 km<sup>2</sup>), at 1 km resolution and the highest electricity consumption it takes ca. 4.5 hours (assuming that the necessary GIS layers are developed beforehand).

## 2.13 License

MIT License

Copyright (c) 2020 Andreas Sahlberg, Alexandros Korkovelos, Babak Khavari, Oluchi Monwe, Dimitrios Mentis, Christopher Arderne

Permission is hereby granted, free of charge, to any person obtaining a copy of this software and associated documentation files (the “Software”), to deal in the Software without restriction, including without limitation the rights to use, copy, modify, merge, publish, distribute, sublicense, and/or sell copies of the Software, and to permit persons to whom the Software is furnished to do so, subject to the following conditions:

The above copyright notice and this permission notice shall be included in all copies or substantial portions of the Software.

THE SOFTWARE IS PROVIDED “AS IS”, WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR COPYRIGHT HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.

## 2.14 onsset

### 2.14.1 onsset package

Submodules

**onsset.gui\_runner** module

**onsset.onsset** module

**class** `onsset.onsset.SettlementProcessor`(*path*)

Bases: `object`

Processes the DataFrame and adds all the columns to determine the cheapest option and the final costs and summaries

**apply\_limitations**(*eleclimit, year, time\_step, prioritization, auto\_densification=0*)

**calc\_summaries**(*df\_summary, sumtechs, tech\_codes, year*)

The next section calculates the summaries for technology split, capacity added and total investment cost

**static calc\_wind\_cfs**(*wind\_vel*)

**calculate\_investments\_and\_capacity**(*sa\_diesel\_investment, sa\_diesel\_capacity, sa\_pv\_investment, sa\_pv\_capacity, mg\_diesel\_investment, mg\_diesel\_capacity, mg\_pv\_investment, mg\_pv\_capacity, mg\_wind\_investment, mg\_wind\_capacity, mg\_hydro\_investment, mg\_hydro\_capacity, grid\_investment, grid\_capacity, year*)

**calculate\_new\_connections**(*year, time\_step, start\_year*)

this method defines new connections for grid related purposes

**Parameters**

- **year** (*int*) –
- **time\_step** (*int*) –
- **start\_year** (*int*) –

**calculate\_off\_grid\_lcoes**(*mg\_hydro\_calc, mg\_wind\_calc, mg\_pv\_calc, sa\_pv\_calc, mg\_diesel\_calc, sa\_diesel\_calc, year, end\_year, time\_step, techs, tech\_codes, diesel\_techs=0*)

Calculate the LCOEs for all off-grid technologies

**calculate\_total\_demand\_per\_settlement**(*year*)

this method calculates total demand for each settlement per year

**Parameters**

- year** (*int*) –

**calibrate\_current\_pop\_and\_urban**(*pop\_actual, urban\_current*)

The function calibrates population values and urban/rural split (as estimated from GIS layers) based on actual values provided by the user for the start year.

**calibrate\_elec\_current**(*elec\_actual, elec\_actual\_urban, elec\_actual\_rural, start\_year, min\_night\_lights=0, min\_pop=50, max\_transformer\_dist=2, max\_mv\_dist=2, max\_hv\_dist=5, buffer=True*)

Calibrate the current electrification status

**choose\_minimum\_off\_grid\_tech**(*year, mg\_hydro\_calc, techs, tech\_codes*)

Choose minimum LCOE off-grid technology

First step determines the off-grid technology with minimum LCOE Second step determines the value (number) of the selected minimum off-grid technology

**Parameters**

- **year** (*int*) –
- **mg\_hydro\_calc** (*dict*) –

**static classify\_elevation**(*elevation*)

Classify the elevation distance according to bins and labels

**Parameters**

- elevation** (*list*) –

**static classify\_land\_cover**(*column*)

this is a different method employed to classify land cover and create new columns with the classification

**Parameters**

- column** (*series*) –

**Notes**

0, 11 = 1 6, 8 = 2 1, 3, 5, 12, 13, 15 = 3 2, 4 = 4 7, 9, 10, 14, 16 = 5

**static classify\_road\_distance**(*road\_distance*)

Classify the road distance according to bins and labels

**Parameters**

**road\_distance** (*list*) –

**static classify\_slope**(*slope*)

Classify the slope according to bins and labels

**Parameters**

**slope** (*list*) –

**static classify\_substation\_distance**(*substation\_distance*)

Classify the substation distance according to bins and labels

**Parameters**

**substation\_distance** (*list*) –

**closest\_electrified\_settlement**(*new\_electrified, unelectrified, cell\_path\_real, grid\_penalty\_ratio, elecorder*)

**compute\_diesel\_cost**(*dataframe: DataFrame, sa\_diesel\_cost: Dict, mg\_diesel\_cost: Dict, year: int*)

Calculate diesel fuel cost

**Parameters**

- **dataframe** (*pandas.DataFrame*) –
- **sa\_diesel\_cost** (*Dict*) –
- **mg\_diesel\_cost** (*Dict*) –
- **year** (*int*) –

**Return type**

*pandas.DataFrame*

**condition\_df**()

Do any initial data conditioning that may be required.

**current\_mv\_line\_dist**()

**diesel\_cost\_columns**(*sa\_diesel\_cost: Dict[str, float], mg\_diesel\_cost: Dict[str, float], year: int*)

Calculate diesel fuel cost based on TravelHours column

**Parameters**

- **sa\_diesel\_cost** (*Dict*) –
- **mg\_diesel\_cost** (*Dict*) –
- **year** (*int*) –

**static do\_kdtree**(*combined\_x\_y\_arrays, points*)

**elec\_extension**(*grid\_calc, max\_dist, year, start\_year, end\_year, time\_step, grid\_capacity\_limit, grid\_connect\_limit, new\_investment, new\_capacity, auto\_intensification=0, prioritization=0*)

Iterate through all electrified settlements and find which settlements can be economically connected to the grid Repeat with newly electrified settlements until no more are added

**get\_grid\_lcoe**(*dist\_adjusted, elecorder, additional\_transformer, year, time\_step, end\_year, grid\_calc*)

**grid\_penalties**(*data\_frame*)

this method calculates the grid penalties in each settlement

First step classifies the parameters and creates new columns with classification

Second step adds the grid penalty to increase grid cost in areas that higher road distance, higher substation distance, unsuitable land cover, high slope angle or high elevation

**static haversine\_vector**(*lon1, lat1, lon2, lat2*)

Calculate the great circle distance between two points on the earth (specified in decimal degrees)

**limit\_hydro\_usage**(*mg\_hydro\_calc, year*)

**pre\_electrification**(*grid\_price, year, time\_step, end\_year, grid\_calc, grid\_capacity\_limit, grid\_connect\_limit*)

” ...

**prepare\_wtf\_tier\_columns**(*num\_people\_per\_hh\_rural, num\_people\_per\_hh\_urban, tier\_1, tier\_2, tier\_3, tier\_4, tier\_5*)

Prepares the five Residential Demand Tier Targets based customized for each country

**project\_pop\_and\_urban**(*pop\_future, urban\_future, start\_year, years\_of\_analysis*)

This function projects population and urban/rural ratio for the different years of the analysis

**results\_columns**(*techs, tech\_codes, year, time\_step, prio, auto\_intensification*)

Calculate the capacity and investment requirements for each settlement

Once the grid extension algorithm has been run, determine the minimum overall option, and calculate the capacity and investment requirements for each settlement

#### Parameters

**year** (*int*) –

**set\_residential\_demand**(*rural\_tier, urban\_tier, num\_people\_per\_hh\_rural, num\_people\_per\_hh\_urban, productive\_demand*)

this method defines residential demand per tier level for each target year

#### Parameters

- **rural\_tier** (*int*) –
- **urban\_tier** (*int*) –
- **num\_people\_per\_hh\_rural** (*float*) –
- **num\_people\_per\_hh\_urban** (*float*) –
- **productive\_demand** (*int*) –

**set\_scenario\_variables**(*year, num\_people\_per\_hh\_rural, num\_people\_per\_hh\_urban, time\_step, start\_year, urban\_tier, rural\_tier, end\_year\_pop, productive\_demand*)

this method determines some basic parameters required in LCOE calculation it sets the basic scenario parameters that differ based on urban/rural so that they are in the table and can be read directly to calculate LCOEs

#### Parameters

- **year** (*int*) –
- **num\_people\_per\_hh\_rural** (*float*) –
- **num\_people\_per\_hh\_urban** (*float*) –
- **time\_step** (*int*) –
- **start\_year** (*int*) –
- **urban\_tier** (*int*) –
- **rural\_tier** (*int*) –
- **end\_year\_pop** (*int*) –
- **productive\_demand** (*int*) –

**update\_grid\_extension\_info**(*grid\_lcoe, dist, dist\_adjusted, prev\_dist, elecorder, new\_elec\_order, max\_dist, new\_lcoes, grid\_capacity\_limit, grid\_connect\_limit, cell\_path\_real, cell\_path\_adjusted, electrified, year, grid\_calc, grid\_investment, new\_investment, grid\_capacity, new\_capacity*)

**class** `onsset.onsset.Technology`(*tech\_life=0, base\_to\_peak\_load\_ratio=0, distribution\_losses=0, connection\_cost\_per\_hh=0, om\_costs=0.0, capital\_cost={inf: 0}, capacity\_factor=0.9, grid\_penalty\_ratio=1, efficiency=1.0, diesel\_price=0.0, grid\_price=0.0, standalone=False, mini\_grid=False, existing\_grid\_cost\_ratio=0.1, grid\_capacity\_investment=0.0, diesel\_truck\_consumption=0, diesel\_truck\_volume=0, om\_of\_td\_lines=0*)

Bases: `object`

Used to define the parameters for each electricity access technology, and to calculate the LCOE depending on input parameters.

**distribution\_network**(*people, energy\_per\_cell, num\_people\_per\_hh, grid\_cell\_area, productive\_nodes=0*)

This method calculates the required components for the distribution network This includes potentially MV lines, LV lines and service transformers

#### Parameters

- **people** (*float*) – Number of people in settlement
- **energy\_per\_cell** (*float*) – Annual energy demand in settlement (kWh)
- **num\_people\_per\_hh** (*float*) – Number of people per household in settlement
- **grid\_cell\_area** (*float*) – Area of settlement (km<sup>2</sup>)
- **productive\_nodes** (*int*) – Additional connections (schools, health facilities, shops)

#### Notes

Based on: <https://www.mdpi.com/1996-1073/12/7/1395>

**get\_lcoe**(*energy\_per\_cell, people, num\_people\_per\_hh, start\_year, end\_year, new\_connections, total\_energy\_per\_cell, prev\_code, grid\_cell\_area, additional\_mv\_line\_length=0.0, capacity\_factor=0.9, grid\_penalty\_ratio=1, fuel\_cost=0, elec\_loop=0, productive\_nodes=0, additional\_transformer=0, penalty=1*)

Calculates the LCOE depending on the parameters.

#### Parameters

- **people** (*float or pandas.Series*) – Number of people in settlement
- **new\_connections** (*float or pandas.Series*) – Number of new people in settlement to connect
- **prev\_code** (*int or pandas.Series*) – Code representation of previous supply technology in settlement
- **total\_energy\_per\_cell** (*float or pandas.Series*) – Total annual energy demand in cell, including already met demand
- **energy\_per\_cell** (*float or pandas.Series*) – Annual energy demand in cell, excluding already met demand
- **num\_people\_per\_hh** (*float or pandas.Series*) – Number of people per household in settlement
- **grid\_cell\_area** (*float or pandas.Series*) – Area of settlement (km<sup>2</sup>)
- **additional\_mv\_line\_length** (*float or pandas.Series*) – Distance to connect the settlement
- **additional\_transformer** (*int*) – If a transformer is needed on other end to connect to HV line
- **productive\_nodes** (*int or pandas.Series*) – Additional connections (schools, health facilities, shops)
- **elec\_loop** (*int or pandas.Series*) – Round of extension in grid extension algorithm
- **penalty** (*float or pandas.Series*) – Cost penalty factor for T&D network, e.g. <https://www.mdpi.com/2071-1050/12/3/777>
- **start\_year** (*int*) –
- **end\_year** (*int*) –
- **capacity\_factor** (*float or pandas.Series*) –
- **grid\_penalty\_ratio** (*float or pandas.Series*) –
- **fuel\_cost** (*float or pandas.Series*) –

**Return type**

lcoe or discounted investment cost

**classmethod set\_default\_values**(*base\_year, start\_year, end\_year, discount\_rate, hv\_line\_type=69, hv\_line\_cost=53000, mv\_line\_type=33, mv\_line\_amperage\_limit=8.0, mv\_line\_cost=7000, lv\_line\_type=0.24, lv\_line\_cost=4250, lv\_line\_max\_length=0.5, service\_transf\_type=50, service\_transf\_cost=4250, max\_nodes\_per\_serv\_trans=300, mv\_lv\_sub\_station\_type=400, mv\_lv\_sub\_station\_cost=10000, mv\_mv\_sub\_station\_cost=10000, hv\_lv\_sub\_station\_type=1000, hv\_lv\_sub\_station\_cost=25000, hv\_mv\_sub\_station\_cost=25000, power\_factor=0.9, load\_moment=9643*)

Initialises the class with parameter values common to all Technologies

**td\_network\_cost**(*people, new\_connections, prev\_code, total\_energy\_per\_cell, energy\_per\_cell, num\_people\_per\_hh, grid\_cell\_area, additional\_mv\_line\_length=0, additional\_transformer=0, productive\_nodes=0, elec\_loop=0, penalty=1*)

Calculates all the transmission and distribution network components

### Parameters

- **people** (*float*) – Number of people in settlement
- **new\_connections** (*float*) – Number of new people in settlement to connect
- **prev\_code** (*int*) – Code representation of previous supply technology in settlement
- **total\_energy\_per\_cell** (*float*) – Total annual energy demand in cell, including already met demand
- **energy\_per\_cell** (*float*) – Annual energy demand in cell, excluding already met demand
- **num\_people\_per\_hh** (*float*) – Number of people per household in settlement
- **grid\_cell\_area** (*float*) – Area of settlement (km<sup>2</sup>)
- **additional\_mv\_line\_length** (*float*) – Distance to connect the settlement
- **additional\_transformer** (*int*) – If a transformer is needed on other end to connect to HV line
- **productive\_nodes** (*int*) – Additional connections (schools, health facilities, shops)
- **elec\_loop** (*int*) – Round of extension in grid extension algorithm
- **penalty** (*float*) – Cost penalty factor for T&D network, e.g. <https://www.mdpi.com/2071-1050/12/3/777>

**transmission\_network**(*peak\_load*, *additional\_mv\_line\_length=0*, *additional\_transformer=0*, *mv\_distribution=False*)

This method calculates the required components for connecting the settlement Settlements can be connected to grid or a hydropower source This includes potentially HV lines, MV lines and substations

### Parameters

- **peak\_load** (*float*) – Peak load in the settlement (kW)
- **additional\_mv\_line\_length** (*float*) – Distance to connect the settlement
- **additional\_transformer** (*int*) – If a transformer is needed on other end to connect to HV line
- **mv\_distribution** (*bool*) – True if distribution network in settlement contains MV lines

### Notes

Based on: <https://www.mdpi.com/1996-1073/12/7/1395>

## onsset.runner module

**onsset.runner.calibration**(*specs\_path*, *csv\_path*, *specs\_path\_calib*, *calibrated\_csv\_path*)

### Parameters

- **specs\_path** –
- **csv\_path** –
- **specs\_path\_calib** –
- **calibrated\_csv\_path** –

`onsset.runner.scenario(specs_path, calibrated_csv_path, results_folder, summary_folder)`

#### Parameters

- `specs_path` (*str*) –
- `calibrated_csv_path` (*str*) –
- `results_folder` (*str*) –
- `summary_folder` (*str*) –

#### **onsset.specs module**

Columns in the specs file must match these exactly

#### **Module contents**

The `onsset` package contains the following modules:

- `onsset.py` : main functions of the model
- `runner.py` : runner is used to calibrate inputs and specify scenario runs

## INDICES AND TABLES

- genindex
- modindex
- search



## PYTHON MODULE INDEX

### O

onsset, 58  
onsset.onsset, 51  
onsset.runner, 57  
onsset.specs, 58



- A**
- `apply_limitations()` (*onsset.settlementprocessor* method), 51
- C**
- `calc_summaries()` (*onsset.onsset.settlementprocessor* method), 51
- `calc_wind_cfs()` (*onsset.onsset.settlementprocessor* static method), 51
- `calculate_investments_and_capacity()` (*onsset.settlementprocessor* method), 51
- `calculate_new_connections()` (*onsset.settlementprocessor* method), 52
- `calculate_off_grid_lcoes()` (*onsset.settlementprocessor* method), 52
- `calculate_total_demand_per_settlement()` (*onsset.settlementprocessor* method), 52
- `calibrate_current_pop_and_urban()` (*onsset.settlementprocessor* method), 52
- `calibrate_elec_current()` (*onsset.settlementprocessor* method), 52
- `calibration()` (*in module onsset.runner*), 57
- `choose_minimum_off_grid_tech()` (*onsset.settlementprocessor* method), 52
- `classify_elevation()` (*onsset.settlementprocessor* static method), 52
- `classify_land_cover()` (*onsset.settlementprocessor* static method), 52
- `classify_road_distance()` (*onsset.settlementprocessor* static method), 53
- `classify_slope()` (*onsset.onsset.settlementprocessor* static method), 53
- `classify_substation_distance()` (*onsset.settlementprocessor* static method), 53
- `closest_electrified_settlement()` (*onsset.settlementprocessor* method), 53
- `compute_diesel_cost()` (*onsset.settlementprocessor* method), 53
- `condition_df()` (*onsset.onsset.settlementprocessor* method), 53
- `current_mv_line_dist()` (*onsset.settlementprocessor* method), 53
- D**
- `diesel_cost_columns()` (*onsset.settlementprocessor* method), 53
- `distribution_network()` (*onsset.onsset.technology* method), 55
- `do_kdtree()` (*onsset.onsset.settlementprocessor* static method), 53
- E**
- `elec_extension()` (*onsset.onsset.settlementprocessor* method), 53
- G**
- `get_grid_lcoe()` (*onsset.onsset.settlementprocessor* method), 54
- `get_lcoe()` (*onsset.onsset.technology* method), 55
- `grid_penalties()` (*onsset.onsset.settlementprocessor* method), 54
- H**
- `haversine_vector()` (*onsset.settlementprocessor* static method), 54
- L**
- `limit_hydro_usage()` (*onsset.settlementprocessor* method), 54

## M

module  
   onsset, 58  
   onsset.onsset, 51  
   onsset.runner, 57  
   onsset.specs, 58

## O

onsset  
   module, 58  
 onsset.onsset  
   module, 51  
 onsset.runner  
   module, 57  
 onsset.specs  
   module, 58

## P

pre\_electrification() (ons-  
   set.onsset.SettlementProcessor method),  
   54  
 prepare\_wtf\_tier\_columns() (ons-  
   set.onsset.SettlementProcessor method),  
   54  
 project\_pop\_and\_urban() (ons-  
   set.onsset.SettlementProcessor method),  
   54

## R

results\_columns() (ons-  
   set.onsset.SettlementProcessor method),  
   54

## S

scenario() (in module onsset.runner), 57  
 set\_default\_values() (onsset.onsset.Technology  
   class method), 56  
 set\_residential\_demand() (ons-  
   set.onsset.SettlementProcessor method),  
   54  
 set\_scenario\_variables() (ons-  
   set.onsset.SettlementProcessor method),  
   54  
 SettlementProcessor (class in onsset.onsset), 51

## T

td\_network\_cost() (onsset.onsset.Technology  
   method), 56  
 Technology (class in onsset.onsset), 55  
 transmission\_network() (onsset.onsset.Technology  
   method), 57

## U

update\_grid\_extension\_info() (ons-  
   set.onsset.SettlementProcessor method),  
   55