THE WORLD BANK
IBRD • IDA | WORLD BANK GROUP

ESMAP
Energy Sector Management Assistance Program

Solar Resource Mapping in Tanzania

SITE IDENTIFICATION REPORT

March 2015

This report was prepared by CENER, under contract to The World Bank.

It is one of several outputs from the solar resource mapping component of the activity "Renewable Energy Resource Mapping and Geospatial Planning – Tanzania" [Project ID: P145287]. This activity is funded and supported by the Energy Sector Management Assistance Program (ESMAP), a multi-donor trust fund administered by The World Bank, under a global initiative on Renewable Energy Resource Mapping. Further details on the initiative can be obtained from the ESMAP website.

This document is an **interim output** from the above-mentioned project. Users are strongly advised to exercise caution when utilizing the information and data contained, as this has not been subject to full peer review. The final, validated, peer reviewed output from this project will be the Tanzania Solar Atlas, which will be published once the project is completed.

Candidate Site Identification report

March, 2015

Table of contents

1 BACKGROUND

This report forms part of the Solar Resource Mapping Project for Tanzania (WB Selection #1139235). The project comprises three phases. The first phase involves project inception, preliminary modelling and implementation planning. The key outcome of phase 1 was an un-validated solar atlas based on the synergistic combination of satellite and Numerical Weather Prediction Model (NWPM) derived solar data for Tanzania. In phase 2, ground-based data collection will be undertaken through a measurement campaign at sites selected from areas defined according to the results of phase 1. Finally, in phase 3, a resource atlas with reduced solar resource uncertainty compared to in phase 1 will be generated from post-processing satellite and NWPM solar radiation outputs, validated by ground-based solar data collected during the measurement campaign of phase 2.

It is worth highlighting that the solar resource models used in the Interim Solar Modelling report (February 2015) are not validated with ground-based coincident solar data measured by well-maintained, high accuracy measuring equipment. Therefore, an unknown deviation between the actual solar resource in Tanzania and the modelled data may exist. The solar renewable energy community depends on radiometric measurements to develop and validate solar radiation models (Gueymard & Myers, 2009). The correct planning and execution of a high quality measurement campaign is thus essential for minimizing uncertainties in solar resource assessment.

2 INTRODUCTION

The objective of this report is to propose a long-list of possible site locations for a solar monitoring network in Tanzania, as well as the type of measurement stations to be installed. These sites are selected according to a ranking methodology based on the results of by the interim solar modelling outputs.

Any solar energy project needs precise information on the available solar resource and its variability over time as a first step in development. It is also important to consider variability of the solar resource in space (how it varies over distance) in order to characterize the regional distribution of solar resource. This information makes it possible to determine the most appropriate siting of solar energy systems and also facilitates decisions over which technologies to be used, as well as appropriate policies and investments.

Tanzania has a tropical climate with regional variations due to topography, where the coastal regions are warm and humid and the highland regions are more temperate. A relatively narrow belt of heavy precipitation and very low pressure that forms near the earth's equator is the main driver of seasonal rainfall in Tanzania. This phenomenon migrates southwards through the country between October and December, reaching the southern regions in January and February, and returning northwards from March to May. This behaviour creates two wet periods in the northern and eastern regions, and one wet session in the central, western and southern region. The diversity of climate zones present in Tanzania in conjunction with its varied topography requires a detailed analysis for the determination of solar monitoring network locations.

In this report, section 3 contains the methodology for area selection according to solar resource (these areas are ~25 km^2, the spatial resolution of solar resource assessment), as well as the methodology for the selection of particular sites within the broad areas, according to on-the-ground limitations. In section 4, the results of the site selection are presented and briefly discussed, showing a long-list of possible site locations for a solar monitoring network in Tanzania. Section 5 details the type of measurement stations proposed in each site defined in previous section. Finally, a conclusions section outlines the core findings of the report.

3 METHODOLOGY FOR SITE IDENTIFICATION

The portion of incoming solar radiation that reaches the Earth's surface exhibits large geographical and temporal variability due to its strong dependence on the atmospheric conditions and meteorology. The ideal way to analyse solar resource is by using ground measurements taken from meteorological stations; however this information is often scarce in many parts of the world, and is only specific to the particular location at which it was measured. As an alternative to ground-collected data, the solar resource can be modelled by satellite-based or NWPM data. Meteorological stations are good at providing high frequency and accurate data (given well-maintained, high accuracy measuring equipment) for a given site. On the other hand, models provide data with a lower frequency of measurement, but characterizing a long history over wide territories. Both kinds of data are needed, since the solar renewable energy community depends on radiometric measurements to develop and validate solar radiation models (Gueymard & Myers, 2009).

The main aim of setting up the monitoring network for the ESMAP project is to minimize uncertainties in solar resource assessments made through post processing of satellite data and NWPM solar radiation data using the validated ground-based solar measurements. The combination of these makes it possible to generate a validated solar resource atlas.

To ground-sample complex irradiance spatial distributions, ideally one seeks to deploy as many sensors as possible. However, to minimize costs, optimizing the number and placement of monitoring equipment is critical (Yang & Reindl, 2015). In particular, we seek to establish a network that represents:

- dispersed high-solar irradiation zones with potentially commercially exploitable solar resources
- different local climate and/or topographical conditions

The methodology defined for such a site selection consists of two steps:

1. – Area selection

First, zones with similar solar irradiance characteristics are grouped together in regions, so that a single sensor can represent the whole region. In these regions, several areas of ~25 km^2 are preselected for site selection.

2. - Specific site selection

Once the regions of interest according to solar resource have been defined, several ground-based limitations, constraints and all-purpose criteria must be taken into account in order to select the most appropriate specific location in each region.

3.1 Methodology for area selection

In this section, a number of distinct spatial regions defined by solar irradiance characteristics is carried out, so that at least a single sensor can be installed in each region to be able of characterize the solar whole resource

3.1.1 Clustering method for region classification

The simplest monitoring network consists of a regular grid, determined by only one design parameter: the inter-station spacing. This arrangement is inefficient since a large number of stations would provide redundant information; in order to maximize efficiency, it is common to group spatial regions with similar solar irradiance characteristics. In this section, a methodology is presented for organizing a collection of solar irradiance patterns into clusters based on similarity criteria.

Such criteria group clusters where characteristics of the solar resource are similar to each (Jain, Murty, & Flynn, 1999). The proposed monitoring sites will be placed one in each cluster.

Clustering is undertaken with powerful and commonly-used methods of pattern-analysis and grouping, which are commonly used for applications such as explanatory data mining, machine learning or pattern classification. The cluster analysis carried out in this study is conducted on a dataset of Monthly GHI values from 2003 to 2013 (132 values at each point of Tanzania), which were extracted for a terrestrial domain over Tanzania (lat 1ºS – 12ºS, lon 28.5ºE – 40.5ºE) with high spatial resolution (0.05º x 0.05º, which approximately corresponds to 5 km x 5 km)[1]. Therefore, matrices of 241 x 221 elements were used.

One potential method for the cluster analysis would is *Partitioning Around Medoids* (PAM). This method starts from an initial set of medoids (i.e., choosing random data points as the centre of a cluster instead mean values), and iteratively replaces the position of the medoid to search for an optimum total distances of the resulting clustering. This approach works effectively for small data sets, but unfortunately it does not scale well for a large data set such as in this case. Instead then, Kaufman and Rousseeuw (Kaufman & Rousseeuw, 1990) proposed a similar algorithm for tackling large applications: CLARA (Clustering for Large Applications). This algorithm works by clustering a sample from the dataset and then assigning all objects in the dataset to these clusters. CLARA then applies the PAM algorithm to generate an optimal set of medoids for the sample, instead of finding medoids for the entire data set. The quality of resulting medoids is measured by the average dissimilarity between every object in the entire data set D and the medoid of its cluster, defined as the following cost function:

$$Cost(M,D) = \frac{\sum_{i=1}^{n} dissimilarity(O_i, rep\{M, O_i\})}{n}$$

Where n is the number of objects, M is a set of selected medoids, $dissimilarity(O_i, O_j)$ is the dissimilarity between objects O_i and O_j, and $rep(M, O_i)$ returns a medoid in M which is closest to O_i. To alleviate sampling bias, CLARA repeats the sampling and clustering process a pre-defined number of times and subsequently selects as the final clustering result the set of medoids with the minimal cost.

The CLARA algorithm was employed to classify the dataset of the solar irradiance series into groups. Fig 1 shows maps for partitioning into 2, 4, 6, 7, 8 and 11 clusters. The “elbow criterion” is used to determine the optimal number of clusters (Goutte, Toft, Rostrup, Nielsen, & Hansen, 1999); the Cost values are shown in Fig 2, as a visual analysis of the cluster output. These present a balance between the distinctiveness of the groups and their disaggregation. 7 clusters has been chosen as the best representation of the different regions that capture the principal patterns of temporal variability and their related spatial situation. A greater number of clusters provide regions with higher levels of disaggregation but less distinctiveness. It is observed that 7 regions presents stabilization in the interpretability of clusters provided by the algorithm.

[1] For more details, see the Interim Solar Modelling Report (P145287, February 2015) within the frame of the Solar Resource Mapping Project for Tanzania (WB Selection #1139235)

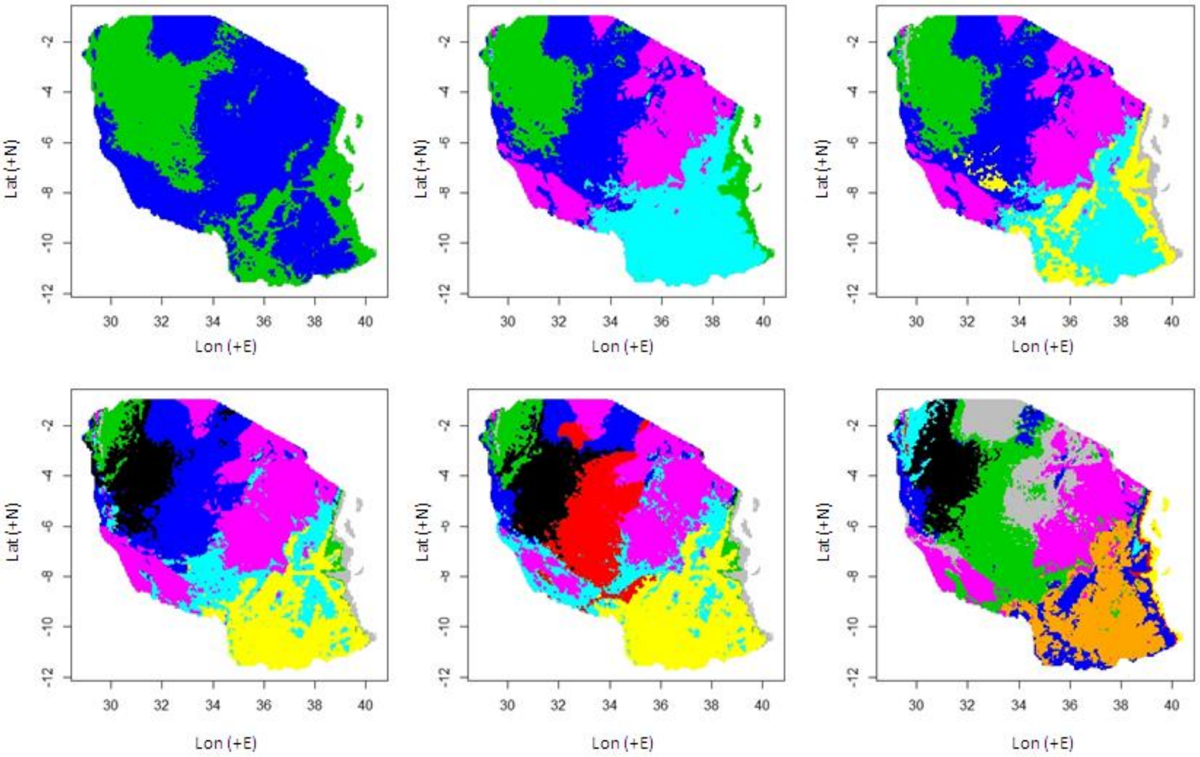

Fig 1 Clustering of solar regimes into different numbers of spatial regions for the time-period 2003–2013. Upper row 2, 4 and 6 clusters; Lower row 7, 8 and 11 clusters.

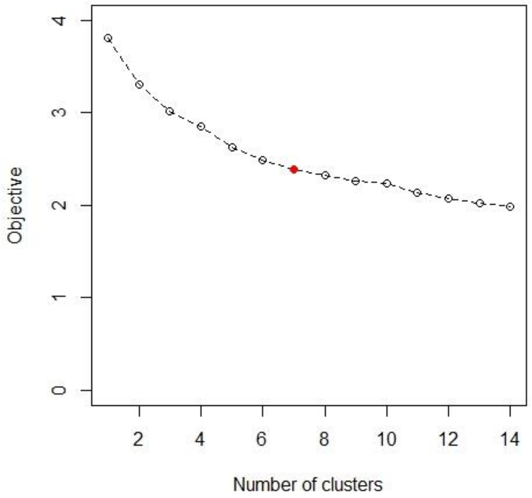

Fig 2 Cost function and number of clusters. Elbow criterion.

3.1.2 Selection of areas

Each area (0.05º x 0.05º) of the domain is ranked according to the following criteria:

1. solar resource in the area, GHI component
2. size of the region (in which the area is located)
3. resource range (spatial variability annual value) within the region (in which the area is located),
4. nearest distance to a city

The ranking method takes into account different features relevant to a monitoring station's utility, such as representativeness (2, 3), logistical considerations (4) and potential for future solar plant development in that area (1, 4). The ranking method used equal weights for each criterion defined.

A second area in each region is selected as a secondary choice via the above criteria, with the requirement that it is sited at a distance equal or higher than 100km from the first ranked area. This distance is considered as general criteria for regional representativeness (McArthur, 2005): successive areas selected in each region should be 100 km or more away from better-ranked areas.

3.2 Specific site selection

This section presents a methodology for identifying a longlist of possible site locations within each of the areas defined in section 3.1. In this analysis, the WB's Safeguards policies[2] and possible limitations on site selection have been taken into account.

3.2.1 Limitations on site selection

Whilst the area selection method in section 3.1 take into account proximity to cities as a logistical constraint, this section considers a broad range of other constrains, primarily using GIS tools.

Limitations taken into account include the following:

- Protected and environmentally sensitive areas (which as far as possible shall be addressed through a preliminary screening process)
- Ease of access for commissioning and maintenance purposes (e.g. road networks)

In addition to these limitations for selecting specific sites, the following criteria have been applied:

- Bodies of water: rivers and lakes
- Terrain slope

These constraints can be described as either 'avoidance' or 'proximity' criteria. Fig 3 maps avoidance criteria, showing optimal areas as green, unfeasible areas as grey, and intermediate areas as orange.

Rivers and lakes **Conservation areas** **Slope**

[2] www.worldbank.org/safeguards

Fig 3 Mapping of Avoidance criteria

With regards to proximity criteria, the following have been taken into consideration:

- Transmission system
- Cities/towns
- Road networks
- Airports

Fig 4 maps proximity criteria, showing optimal areas as green, unfeasibly areas as red, and intermediate areas as orange.

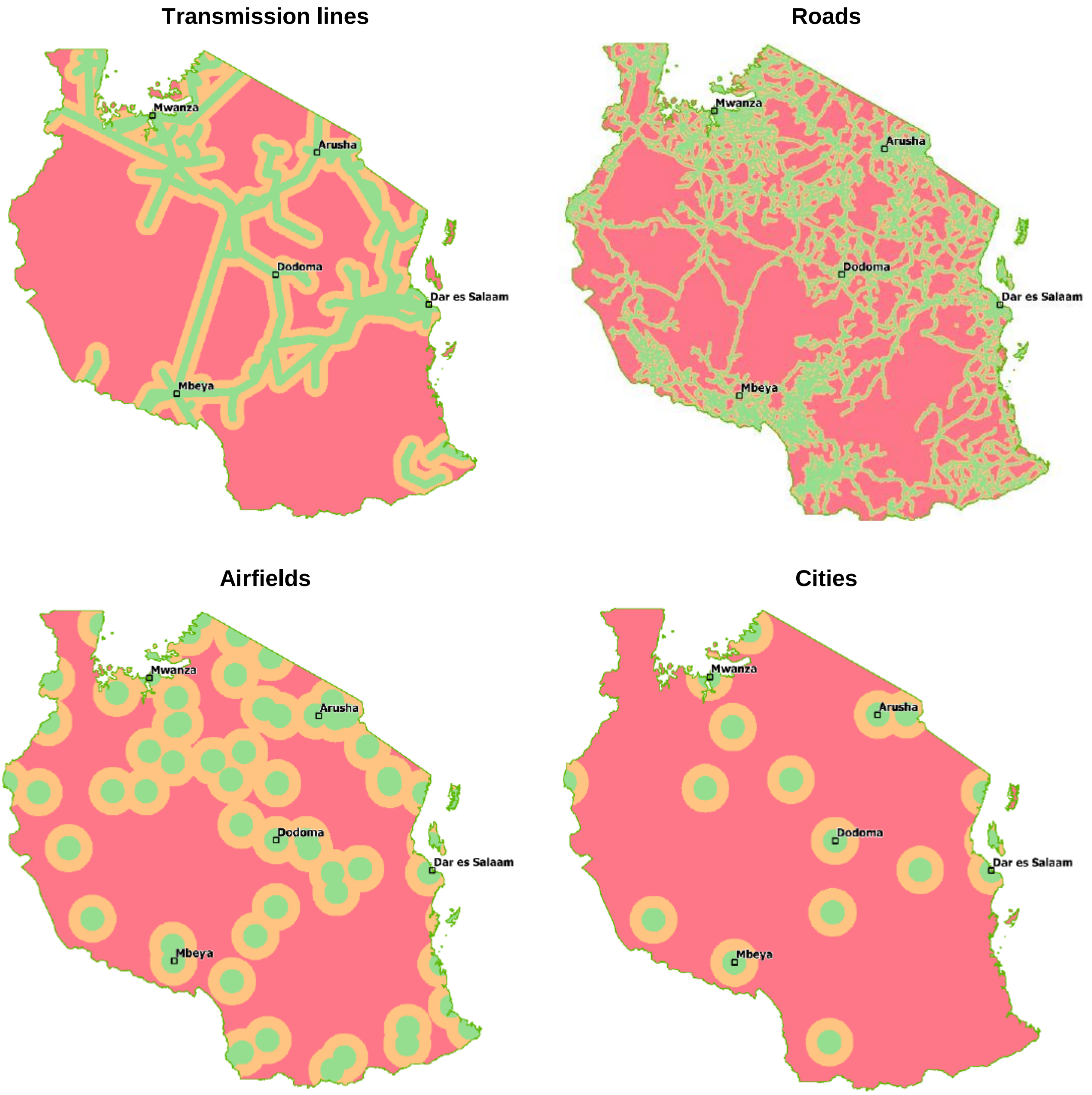

Fig 4 Mapping of proximity criteria

4 LONGLIST OF POSSIBLE SITE LOCATIONS

This section presents the results of the methodology for site identification presented in section 3: a longlist of possible site locations that meet all of the site selection requirements. The results are ranked in line with the overall methodology: firstly the regions defined by means of the clusters analysis are shown along with the selected areas (~25 km^2) within these regions; secondly the specific sites selected within these regions are shown.

4.1 Results of the methodology

4.1.1 Regions and areas selected

This section shows the solar resource regions resulting from the clustering analysis of section 3.1.1. The analysis carried out provides 7 different regions, plotted in Fig 5 as colour lines, where darker tones indicate higher solar resource values. Black lines correspond to country limits.

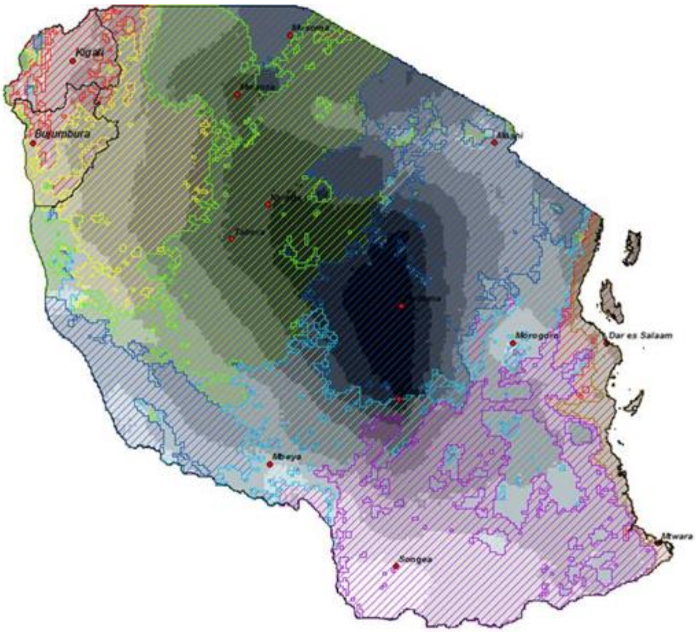

Fig 5 Colour-coded solar regimes (darker greys indicate higher solar resource values)

The regions found have the following characteristics (Table 1):

Table 1 Characteristics of solar regimes found

Region #	Colour code	Location	Land covered (~km^2)	Mean GHI (kWh/m^{2}d)	GHI Variability (%)
1	Green	Central west; north	37765	5.32	30
2	Red	North west; east	7900	4.71	17
3	Yellow	North west	17680	4.95	21
4	Light blue	Central south; east; west	24685	4.93	40
5	Dark blue	East; north; south west	41590	5.30	34
6	Brown	East	3490	4.71	18
7	Pink	South	29585	4.60	36

Each area (~25 km^2) has been ranked within its cluster region in accordance with the procedure detailed in section 3.1.2. Fig 6 shows the locations of the areas selected: red text indicates the first option for a given region, orange text indicates the second, and yellow text is used for subsequent options (sites for Rwanda and Burundi are included). The sites shown in Fig 6 are the 16th best ranked places, and they will be reduced to the final longlist of sites (section 4.2) according to the rank of each region.

Distance Criteria **Avoidance Criteria**

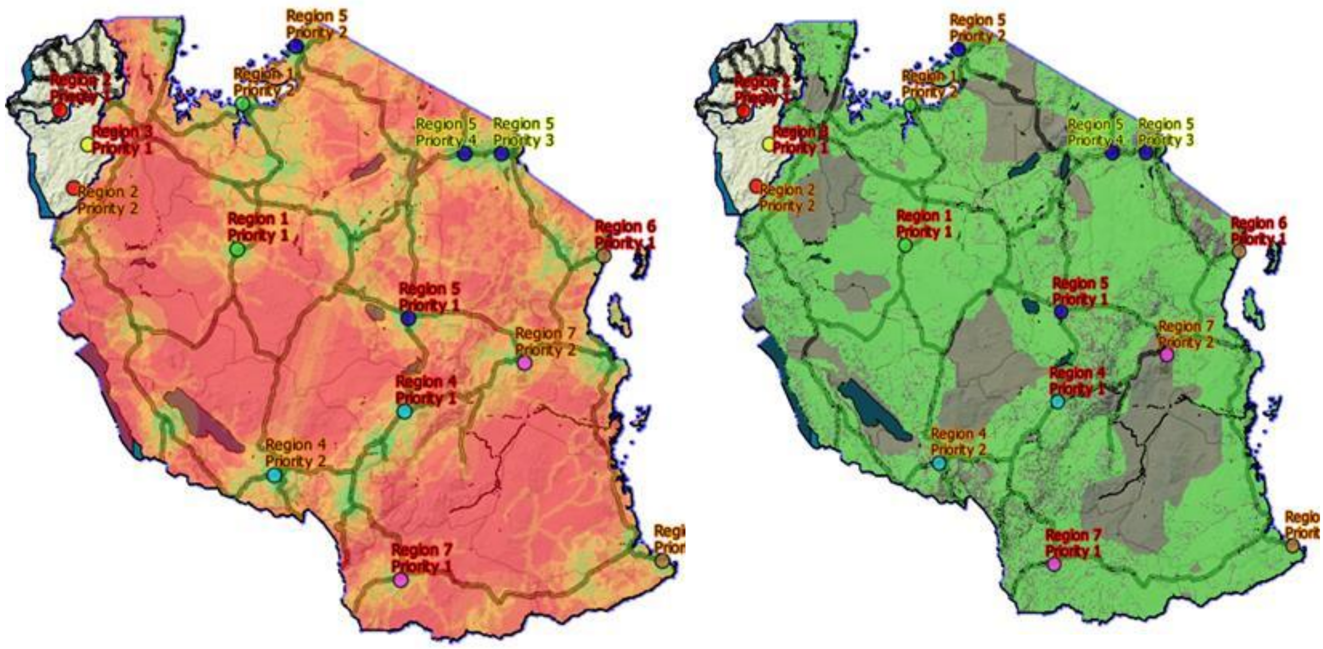

Fig 6 Locations of the selected areas

4.2 Longlist of specific sites selected

This section shows possible site locations for the solar radiation monitoring network as described above.

According to the Terms of Reference (TOR) of the Solar Resource Mapping Project for Tanzania, high quality solar data from a minimum of six sites over a 24-month period must be obtained. The deployment of only 6 monitoring stations however would not enable one station to be installed within each of the solar characteristic regions indicated in section 4.1.1. It is recommended that one measurement station should be installed in each region to be able of characterize all of Tanania's varied solar resources. Furthermore, detailed analysis of the regions (Table 1) suggests that in several regions more than 1 monitoring station may be required to fully record the region's characteristics.

Recommendations for measurement sites are given for each of the regions defined:

Region #1: The 2nd largest area of the regions. It has a mean solar resource of 5.32 kWh/m^{2}d and a variability rank of 1.59 kWh/m^{2}d.

Recommended number of monitoring sites: **2 sites**. The first site is located in the central region of Tanzania, and the second one in the northern lakes region. Table 2 provides a list of potential locations for the stations within these sites.

Table 2 Proposed specific sites for Region #1

Monitoring sites	Description	Proposed specific emplacements
Site 1	Tabora Town	Tanesco facilities TMA monitoring station Tabora
Site 2	Mwanza city	Dar Institute of technology, Mwanza Campus (Makongoro road) TMA monitoring station Mwanza Mwanza Airport (2.4412556 N, 32.9232250 E) TANESCO – Mwanza office (2.6563767 S, 32.6276565 E) VETA [3]compound, Mwanza

Region #2. This is a relatively small region, with the 6th largest area of the solar regimes (located at the North West and the East of Tanzania, and in Rwanda). It has a mean solar resource of 4.71 kWh/m^{2}d.

Recommended number of monitoring sites: **2 sites**. - The first site is located Rwanda, and the second one in Burundi. Table 3 provides a list of potential locations for the stations within these sites.

[3] Vocational Education and Training Authority – it is a government agency with training centres across the country

Table 3 Proposed specific sites for Region #2

Monitoring sites	Description	Proposed specific emplacements
Site 3	Kigali city	Airport Rwanda Meteorological Agency
Site 4	Bujumbura city	University of Burundi REGIDESO – electricity offices and substations

Region #3. This is a medium size region covering the 5nd biggest area of the solar regimes (located at the North West of Tanzania). It has a solar resource of 4.95 kWh/m^{2}d and a variability rank of 1.05 kWh/m^{2}d.

Recommended number of monitoring sites: **1 site** - The site is located in Burundi. Table 4 shows several specific sites proposed in this region.

Table 4 Proposed specific sites for Region #3

Monitoring sites	Description	Proposed specific emplacements
Site 5	Nyabikere	REGIDESO – electricity offices and substations Nyabikere Commune headquarters

Region #4. This is a medium size region, with the 4th largest area of the solar regimes (dispersed along different zones of Tanzania: Central south; East; West). It has a mean solar resource of 4.93 kWh/m^{2}d and a variability rank of 1.97 kWh/m^{2}d.

Recommended number of monitoring sites: **2 sites**. – The first site is located in Iringa, and the second one is located in Mbeya. Table 5 provides a list of potential locations for the stations within these sites.

Table 5 Proposed specific sites for Region #4

Monitoring sites	Description	Proposed specific emplacements
Site 6	Iringa town	Ruaha Univertsity College TMA monitoring station Iringa Mkwawa University Education College Airport TANESCO – Iringa office (7.773042 S, 35.700034 E) TANESCO – Ipogolo substation
Site 7	Mbeya city	Airport TMA monitoring station Mbeya Mzumbe university Mwanza University of Science and Technology TANESCO- Mbeya offices (9.110927 S, 32.937296 E and 9.602384 S, 33.870286 E) TANESCO – Iyunga and Mwakibete substations

Region #5. This is the largest region of all of the solar regimes (dispersed along different zones of Tanzania: East; North; West). It has a mean solar resource of 5.30 kWh/m^{2}d and a variability rank of 1.83 kWh/m^{2}d.

Recommended number of monitoring sites: **3 sites**. The first site is located in Dodoma, the second one in Musoma, and the third one in Mochi. Table 6 provides a list of potential locations for the stations within these sites.

Table 6 Proposed specific sites for Region #5

Monitoring sites	Description	Proposed specific emplacements
Site 8	Dodoma town	John University of Tanzania TMA monitoring station Dodoma TANESCO – Dodoma office Airport VETA compound Dodoma
Site 9	Musoma town	Airport TMA monitoring station Musoma TANESCO - substation Kemgesi Windpower Cooperative
Site 10	Moshi town	Airport TMA monitoring station Moshi Moshi Technical School TPC Langasani (rainfall station located at 3.47S, 37.33E) TANESCO - Arusha office (3.416995 S, 36.710204 E) TANESCO – substation

Region #6. This small region covers the 7th largest area of the solar regimes (Eastern Tanzaina). It has a mean solar resource of 4.71 kWh/m^{2}d and a variability rank of 0.88 kWh/m^{2}d.

Recommended number of monitoring sites: **2 sites** - The first site is located in Dar es Salaam, and the second one is located in Mtwara. Table 7 provides a list of potential locations for the stations within these sites.

Table 7 Proposed specific sites for Region #6

Monitoring sites	Description	Proposed specific emplacements
Site 11	Dar es Salaam city	Univesity of Dar es Salaam Dar es Salaam Institute of Technology TANESCO offices in Dar es Salaam (Kinondoni North, Ilala, Temeke, Kinondoni South)
Site 12	Mtwara town	Airport TMA monitoring station Mtwara TANESCO – power station

Region #7. This is a wide region located at the south of Tanzania, covering the 3nd biggest area of the solar regimes. It has a mean solar resource of 4.60 kWh/m^{2}d and a variability rank of 1.65 kWh/m^{2}d.

Recommended number of monitoring sites: **2 sites** - The first site is located in Songea, and the second one is located in Morogoro. Table 8 provides a list of potential locations for the stations within these sites.

Table 8 Proposed specific sites for Region #7

Monitoring sites	Description	Proposed specific emplacements
Site 13	Songea town	Airport TMA monitoring station Songea Songea Teachers Training College TANESCO power station
Site 14	Morogoro town	Airport TMA monitoring station Morogoro TANESCO – Morogoro office (6.821518 S, 37.663702 E) TANESCO - Morogoro substation (6.799987 S, 37.661932 E) VETA Compound Morogor

At this stage the sites' locations are provisional; the exact locations of these sites may change depending on the outcome of field visits to the locations. Furthermore possible partnership with the Tanzania Meteorological Agency (TMA) may lead to installation on their pre-existing sites).

5 PROPOSAL OF NUMBER AND TYPE OF MEASUREMENT STATIONS

This section proposes the number and type of measurement stations to be deployed, in order to maximize the cost–benefit ratio of the monitoring network. A Tier 1 station is recommended for installation in the Dar es Salam site, and for the remaining 13 sites, Tier 2 stations are recommended. A brief description of the characteristics of these stations is presented below:

Tier 1 station

Tier 1 is a research quality station that provides the highest accuracy measurements, collecting data according to the protocols of the Baseline Surface Radiation Network (http://www.bsrn.awi.de/). A Tier 1 station's design, as described in the Technical Proposal (Selection #1139235: "Renewable Energy Resource Mapping: Tanzania, Africa Region"), is characterized by:

- Solar tracker and shadow accessories for pyrheliometers and pyranometers. Model Sun Tracker-2000 and Model SA-2000 shade disk kit
- Two thermoelectric pyranometer Secondary Standard. Model GEO-SR20. For measures of GHI and diffuse components
- A thermoelectric Pyrheliometer First Class. Model GEO-DR01 for DNI measurements

The Tier 1 station also includes a weather station with temperature, humidity, barometric pressure, wind speed and direction sensors. The quality and variety of types of data collected by this station makes it a valuable tool for many different kinds scientific research. Such equipment requires trained personnel readily available to operate and maintain the station properly. Dar es Salaam has been selected as the location for this station in order to enhance research capabilities at the University of Dar es Salaam and other local research institutions and private companies in the area.

Tier 2 station (Thermopile Radiometric Station)

- Solar tracker and shadow accessories for pyrheliometers and pyranometers. Model Sun Tracker-2000 and Model SA-2000 shade disk kit
- One thermoelectric Pyranometer First Class. Model GEO-SR11
- One thermoelectric Pyrheliometer First Class. Model GEO-DR01

There are several common elements for both Tier1 and the Tier 2 stations, principally data acquisition systems, remote communication, solar panels and batteries for electricity supply and meteorological sensors. The specifications for such common elements are as follows:

- Remote Automatic Data Acquisition and Transmission Unit. Model METEODATA-3016 (3008 for Tier2 alternatives)
- Communications ADD-INS (GPRS, Ethernet, Satellite internet)
- Lightning protection
- Solar panel 95W peak power
- Meteorological tower model TM-180-09. 10 m of height (3 m for Tier2)
- Wind speed and direction sensors. Model 03002 Wind Sentry
- Relative humidity and air temperature sensors. Model STH-S331

- Barometric pressure sensor. Model 61302V

When required, an alternate Tier2 station setup may be employed compared with the specifications of the Technical proposal of Selection #1139235: “Renewable Energy Resource Mapping: Tanzania, Africa Region”. This configuration offers high quality measurements of GHI and DNI components of solar radiation, allowing diffuse radiation to be determined mathematically. Meanwhile, the other Tier2 configurations obtain direct measurements of GHI and diffuse components, with DNI estimated by a model.

These stations fulfil the requirements of the annex C of TOR as all of them are capable of providing the three components of solar radiation.

6 CONCLUSIONS

Satellite or NWPM based solar radiation data are not capable of reproducing instantaneous solar irradiance values or local patterns to the same accuracy as ground sensors, but they can provide robust aggregated values over a long history and wide territories. Minimizing uncertainties in remote solar resource assessment methods requires a high quality measurement campaign (with well-maintained, high accuracy measuring equipment) for validation purposes.

This report has proposed a longlist of possible site locations for a solar monitoring network in Tanzania, as well as detailed the types of measurement stations to be installed. The longlist was determined by use of a ranking methodology that combined the classification of representative solar characteristic zones with logistical considerations that affect the potential for commercial exploitation of solar resources, as well as represent different climate and topographical conditions.

In order to avoid sites that provide redundant information in the solar monitoring network, it is common to group spatial regions with similar solar irradiance characteristics. The CLARA algorithm has been employed in this work (section 3.1.1) to classify the dataset of the solar irradiance series obtained in the solar modelling report (P145287, February 2015) into groups that represent solar regimes, obtaining 7 clusters (or regions). Within these regions, a methodology for area selection has been applied, taking into account features important in site selection: representativeness, logistical considerations and possible future solar energy facility development (section 3.1.2). These areas are approximately ~5 x 5 km in size. Finally, for specific site selection within selected areas, several limitations and constraints have been taken into consideration (section 3.2), concluding that at least one measurement site for each region is required, and in some regions more than 1 measurement sites seem to be necessary. In particular, 14 sites have been identified (section 4.2). The exact locations of these sites must be analyzed in detail with local partners, and they must be analyzed in conjunction with Tanzania Meteorological sites in the search of synergies (for maintenance procedures). According to the longlist of specific sites selected within the regions, and the different site characteristics, a Tier 1 station is installed in Dar es Salam, and Tier 2 stations are installed in the other 13 sites (section 5). Table 9 summarizes the longlist of possible site locations, highlighting in blue the top priority sites. It is important to note that the list includes sites from Rwanda and Burundi, because the analysis has been made for the whole geographic area.

Table 9 Longlist of possible site locations. Tz, Rw and Br refer to Tanzania, Rwanda and Burundi respectively.

Region #	Location	Land covered (~km2)	Sites (type of station)	Nearest city
1	Central west; north	37765	Site 1 (Tier 2)	Tabora (Tz)
			Site 2 (Tier 2)	Mwanza (Tz)
2	North west; east	7900	Site 3 (Tier 2)	Kigali (Rw)
			Site 4 (Tier 2)	Buyumbura (Br)
3	North west	17680	Site 5 (Tier 2)	Nyabikere (Br)
4	Central south; east; west	24685	Site 6 (Tier 2)	Iringa (Tz)
			Site 7 (Tier 2)	Mbeya (Tz)
5	East; north; south west	41590	Site 8 (Tier 2)	Dodoma (Tz)
			Site 9 (Tier 2)	Musoma (Tz)
			Site 10 (Tier 2)	Moschi (Tz)
6	East	3490	Site 11 (Tier 1)	Dar es Salaam (Tz)
			Site 12 (Tier 2)	Mtwara (Tz)
7	South	29585	Site 13 (Tier 2)	Songea (Tz)
			Site 14 (Tier 2)	Morogoro (Tz)

The ranking methodology applied categorizes the most representative solar regimes according to the monthly solar irradiation series (2003-2013) provided by the interim solar modelling report (P145287, February 2015), which has provided a sufficient understand of the measurement network required to best characterize the solar resource in Tanzania. Therefore, not only have average solar irradiance values been considered, but also inter-annual and intra-annual solar irradiance variations.

7 REFERENCES

Goutte, C., Toft, P., Rostrup, E., Nielsen, F. Å., & Hansen, L. K. (1999). On clustering fMRI time series. NeuroImage, 9(3), 298-310.

Gueymard, C. A., & Myers, D. R. (2009). Evaluation of conventional and high-performance routine solar radiation measurements for improved solar resource, climatological trends, and radiative modeling. Solar Energy, 83(2), 171-185.

Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: review. ACM Computing Surveys, 31(3), 264–323.

Kaufman, L., & Rousseeuw, P. J. (1990). Finding Groups in Data: An Introduction to Cluster Analysis.

McArthur, L. J. B. (2005). BASELINE SURFACE RADIATION NETWORK (BSRN) Operations Manual Version 2.1. WORLD CLIMATE RESEARCH PROGRAMME. WCRP-121.

Yang, D., & Reindl, T. (2015). Solar irradiance monitoring network design using the variance quadtree algorithm. Renewables: Wind, Water, and Solar, 1(1).