Renewable Energy 155 (2020) 761–769

Contents lists available at ScienceDirect

Renewable Energy

journal homepage: www.elsevier.com/locate/renene

Identification and assessment of sites for solar farms development using GIS and density based clustering technique- A case of Pakistan

Fahd Amjad , Liaqat Ali Shah *

Centre of Excellence for CPEC, Pakistan Institute of Development Economics, Pakistan

A R T I C L E I N F O

Article history:
Received 10 March 2019
Received in revised form
8 January 2020
Accepted 15 March 2020
Available online 20 March 2020

Keywords:
Solar irradiance
Sites selection
Density-based clustering
Geographic information systems

A B S T R A C T

Site selection plays a vital role in the entire life cycle of solar farm and merits further consideration. Current studies for the site selection have several limitations. First, application of criteria related to infrastructure in the initial phase can screen out prospective solar farms or make them less desirable. Second, infrastructure criteria are less significant in circumstances where new cities and town are to be planned. Lastly, clustering is often ignored in site selection problem to classify farms and determine their sizes. To overcome these limitations, this study proposes a methodology that focuses primarily on areas that has maximum energy potential and excludes the infrastructure requirements initially from the analysis. The methodology uses Geographical Information System (GIS) for data acquisition and mapping, whereas a novel density-based clustering approach is employed to identify and group sites with high solar potential. The methodology is applied to the geographic boundaries of Pakistan. However, the methodology can be applied to any spatial context subject to the availability of similar data. The paper concludes with recommendations to energy policy makers by providing a list of potential clusters, with their sizes ranging between $10 - 289\ km^2$ located in the Baluchistan province of Pakistan.

1. Introduction

Climate change is one of the most serious threats to mankind. Approximately 80% of the emissions from greenhouse gases (GHGs) are due to the production and consumption of energy [1]; fossil fuels – coal, oil, and natural gas – contribute to about 90% of the total global carbon dioxide (CO_2) emissions alone. The ambiguous relationship between the rise in GHG emissions and climate change, and how those changes affect the environment, have made this a challenging issue to tackle [2]. In the preceding 250 years, avant-garde wide-scale agriculture, changes in uses of land, and the increase in fossil fuel consumption has resulted in increasing GHGs concentrations [3]. Due to the heavy reliance on nonrenewable energy sources, especially in developing countries, GHG emissions have continued to accelerate at an alarming rate. In fact, based on ice core records, the existing levels of CO_2 and methane are a lot more than they have ever been in the last 650,000 years. Even though fossil fuels are a relatively cheaper source of nonrenewable energy, the hazard of global warming has led scientists and researchers to find alternative renewable energy sources to generate power [4,5].

Moreover, sustainability paradigm has further pushed the renewable energy agenda. But the challenge remains there on how to create clean energy using renewable energy resources. Wind, solar, geothermal, and biomass are the available renewable resources that can be utilized for power generation [6]. Of these four, solar energy is the most abundant, inexhaustible and clean of all the renewable energy resources to date [7] and therefore the most promising renewable energy option for large-scale power generation [8]. That is why the application of solar energy for power generation and thermal utilities is on the rise. The renewable global status report also confirms the global energy transition towards renewable energy in general and solar energy in particular [9].

In addition to the abundant availability of solar energy, the steady progress in renewable energy technologies, particularly photovoltaic technology in terms of power generation capabilities and its rapidly falling prices, is opening up new opportunities to harness solar energy more effectively and efficiently [10]. However, efficient solar technology at lower cost is not a sufficient condition to develop solar farms for solar energy harvesting. Other dimensions such as technical, economic and environmental are usually considered to check the viability of a solar farm. These

* Corresponding author.
E-mail addresses: fahd.amjad@gmail.com (F. Amjad), shah.liaqat@hotmail.com (L.A. Shah).

https://doi.org/10.1016/j.renene.2020.03.083
0960-1481/

dimensions are influenced by geographical location, biophysical attributes and the socioeconomic infrastructure of the region [11]. Nevertheless, relying on all the dimensions may bias the site selection more towards developed infrastructure in order to reduce the cost of solar farm development, or the decision can be skewed towards socio-economic concerns which subsequently can influence the site selection process. It is desirable to first identify viable locations for solar farm development and then in the second stage, if required, apply criteria covering economic and environmental issues. However, the existing methods for site selection employ criteria as tools for ranking desirable locations and eliminate the suitable options from the decision equation.

2. Methods for site selection

Multi-criteria decision making methods (MCDM) are often employed in renewable energy planning, especially for site selection [12–16]. In MCDM, set of pre-determined alternatives are evaluated against set of criteria and the former are then ranked and/or classified using the latter employing different valuation techniques. Commonly used techniques include Analytical Hierarchy Process (AHP), Technique for Order Preference by Similarity to Ideal Solution (TOPSIS), Elimination and Choice Expressing Reality (ELECTRE) and Fuzzy based Methods [17]. One or combination of these methods are employed to identify, classify or rank locations for solar farm development. For example [18], use the AHP method to identify optimum solar site among three locations in Turkey. Fuzzy based TOPSIS method is used in the study for the purpose [19]. In the study [20], a farm location for hybrid wind/solar has been identified through ELECTRE Method [21]. use AHP, TOPSIS and ELECTRE methods to determine the optimum location for solar farm development in Murcia in the southeast of Spain. The authors employ AHP to obtain weights of the criteria whereas suitable locations in the area are evaluated and classified using TOPSIS and ELETRE. A plethora of research on MCDM application for site selection of renewable energy can be found in the literature [17].

Besides, practitioners and researchers are using GIS in combination with MCDM for the spatial planning of solar farm sites [11,22–25]. The use of MCDM with GIS for spatial planning, particularly site selection, offers several advantages [26]. Prominent among the advantages is the development of a GIS-based decision support system for the selection of solar sites that could support new infrastructure development near those locations to promote the utilization of clean energy [12]. Other advantages include but are not limited to improved efficiency of the solar farm by ensuring a high level of solar irradiance and moderate air temperatures, proximity to consumption points and reduced environmental and social impacts of the project.

Irrespective of the method(s), the commonly used criteria employed to evaluate the site suitability are solar irradiance, slope, distance from road, distance from transmission lines, land use/area and noise issue [17]. However, there are a few disadvantages and implications of using criteria in the initial phase to evaluate solar sites. First, lack of or poor infrastructure that is, roads, transmission lines and grid station can turn optimal locations in terms of energy production into suboptimal. It is likely that the energy generation potential of solar farm outperforms the infrastructure investment. Second, infrastructure criteria are less significant in circumstances where new cities and town are to be planned. In the case of countries like Pakistan where developing new cities and towns become unavoidable to create space for its swiftly urbanizing populace and lessen pressure on its existing urban centers, infrastructure criteria for site selection become less significant. Lastly, clustering is often ignored in site selection problem to classify farms and determine their sizes.

In view of the foregoing, it is important to approach the problem of solar site selection by progressively employing the screening criteria and/or restrictions to filter out dense solar regions with relatively less undulating terrain. Potential sites are then subjected to clustering mechanism to identify high density adjacent solar regions of required size. For clustering, K-means [27] is often used. However, the methodology has several limitations. Notable among them is the lack of clustering of irregular shapes, which is often the case in solar farm development.

The rest of the paper is organized as follows. In section 3, the paper briefly discusses the data types and sources. The following section explains the proposed methodology. Section 5 presents the energy scenario of Pakistan and explains the application of methodology to the geographical boundaries of the country. Results are presented in the same section, followed by conclusion and policy recommendations.

3. Data sources and selection of solar irradiance variables

To carry out the study, three types of data are required. To delimit the zone for solar analysis, administrative boundary data is required and can be obtained from sources [28]. Other two types of data for analysis are solar irradiance dataset and elevation dataset for the area of interest and are obtained from Energy System Management Assistance Program (ESMAP) [29]. Both of the data layers are provided in a geographic spatial reference (EPSG: 4326). The solar irradiance dataset is in raster format, while the administrative boundary dataset is in vector format.

Two variables of solar irradiance, the Direct Normal Irradiance (DNI) and the Diffuse Horizontal Irradiance (DHI), are often considered important for the decisions of solar site location. In a clear and cloudless atmosphere, the radiation that reaches the ground directly is called DNI (hypotenuse in Fig. 1(b)). On the other hand, of the radiation that passes through the atmosphere, part of it is scattered. The scattered part of the radiation is called DHI (represented by the horizontal line in Fig. 1(b).

Global Horizontal Irradiance (GHI) is the amount of radiation received by a surface horizontal to the ground. In other words, it is the amount of direct radiation (DNI) and diffuse radiation (DHI) and is calculated mathematically as follows:

$$GHI = DNI.\cos(Z) + DHI$$

where Z is zenith angle, as shown in Fig. 1(b).

Note: $GHI,\ DNI,\ DHI$ *are measured in* kWh/m^2.

For the current study, however, only yearly DNI data is considered. If more detailed monthly production capacity is required, then the monthly average DNI can be used, which is available from the same date source.

4. Methodology

The proposed methodology is depicted in Fig. 2. The methodology is divided into three parts: solar mapping, zone evaluation and clustering of the regions to compute solar farm sizes.

The solar mapping and zone evaluation part of the methodology are carried out using QGIS and R. For clustering, the DBSCAN Package in R software is used. For the visualization of clusters, the FACTOREXTRA package in R is employed. Each phase of the methodology is briefly explained in the following subsections.

4.1. Solar mapping

For the solar mapping, the three types of data: vector data for administrative boundaries of the area of interest, raster data for the

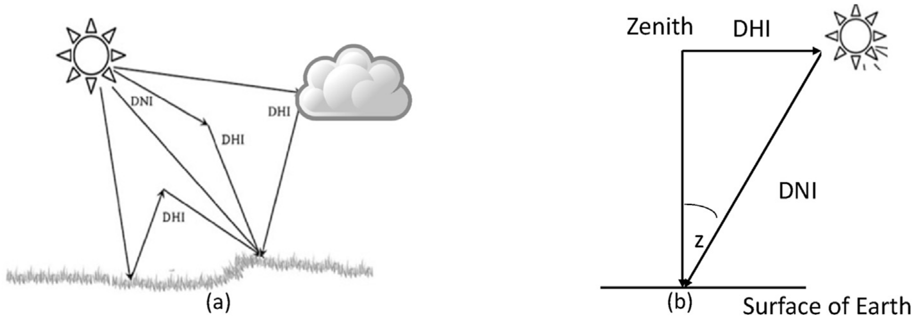


Fig. 1. (a) Components of solar irradiation intercepted by the earth's surface, (b) solar irradiance triangle.

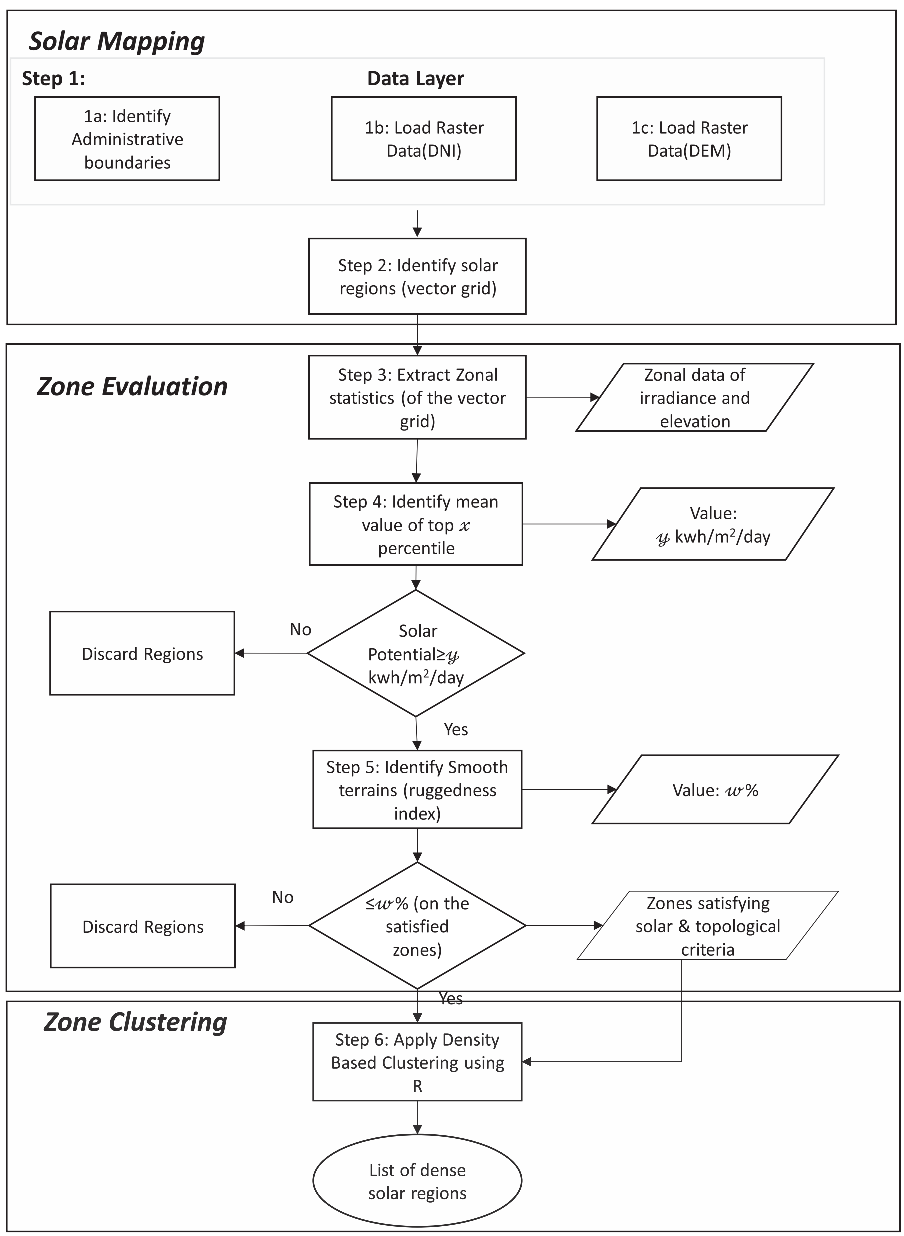


Fig. 2. Methodology for solar site selection.

DNI and raster data for the elevation are loaded into the QGIS (Step 1). However, to localize data within the raster layer regarding solar irradiance, a vector grid of 1×1 km^2 can be used (Step 2). This localization using meshing is to divide the region of interest into sizeable areas in order to extract raster statistics. Subsequently, raster statistics of irradiance and elevation are extracted (Step 3) using the zonal statistics function of the Quantum GIS software version 2.18 [30].

4.2. Zone evaluation

In the zone evaluation phase of the methodology, several decisions are made. Restrictions on the data are decided. For instance, application of restriction to include or exclude data such as of solar irradiance above certain threshold is decided at this stage. For this purpose, multiple scenarios can be planned, that is top 5%, 10% or any other percentile can be decided for analysis (Step 4). Same is the case with other parameters such as terrain ruggedness (Step 5). It is decided to what extent the land quality is permissible and scenarios are built accordingly. Undulating land with relatively less ruggedness is preferable for large solar farm development.

4.3. Zone Clustering

Once the solar mapping and zone evaluation are performed, the next phase is to identify clusters within the dataset[1] (Step 6). The purpose is to compute specific minimum area size for solar farm development.

Different clustering techniques, such as k-means [31], and k-medoid – PAM algorithm [32], can be applied. The two approaches are useful if the dataset is concentrated around a central point. However, there are often places where dense regions are irregularly and longitudinally shaped and, in this scenario, the two clustering approaches do not yield satisfactory results because they do not provide the exact location but rather provide solar strategic zones in the area of interest.

However, the density-based clustering algorithm does not have the same limitations as the other two previously discussed clustering algorithms. This algorithm distinguishes the regions with high and low densities and can be used to identify the clusters of any shape, including circular clusters. It does not require the prior identification of the number of clusters, and it is also sensitive to the presence of outliers in the dataset. For a more detailed discussion on the clustering algorithm and their validation process, readers are referred to Ref. [33].

The density-based clustering requires two parameters: first is the radius of influence or epsilon value, and second, the conditions of minimum points, which identifies the required density parameter. To make a cluster, the density-based clustering algorithm starts arbitrarily in the identified zone and encircles a region that has an epsilon value (Eps $= 0.02268$) to encompass the minimum points, which in our case are 9. The epsilon value is the Euclidean distance between the centers of the two adjacent 1-km^2 sites, and the minimum points simply means the size of the cluster to be formed. Since the distance between two points is 1 km, the 9-point condition translates into a cluster size of a minimum of 9 km^2. If the required points are encircled, then the algorithm proceeds to the adjacent zones and sees if the required conditions are met; that is, the new circle encloses a minimum of 9 points. If yes, annex it to the developing cluster; if no, then disrupt the cluster formation. The progressive cluster formation using a density-based algorithm is shown in Fig. 3.

In short, the density-based clustering algorithm computes cluster size of a farm from the number of adjacent 1×1 km^2 zones that form the cluster. The location of the cluster is the mean values of latitude and longitude for all of the zones within the identified cluster.

5. Case study: solar potential of Pakistan

The proposed methodology has been applied to the geographical regions of Pakistan. The case of Pakistan is selected because the country is currently conducting power market study under CPEC to identify supply-demand gap till 2030. The focus is to move from fossil fuel to include more renewable energy in the energy mix. As of now, the share of renewable energy in Pakistan is only 1.1% [34].

5.1. Pakistan's potential of solar energy

Pakistan has favorable conditions for exploitation of solar energy, since it receives high solar irradiance in certain parts. The recent solar energy mapping with the help of the World Bank (ESMAP – Energy Sector Management Assistance Program) identifies the solar potential of the country [29]. The country receives solar energy with an average of 5–7 kW h/m^2/day and approximately 1800–2200 kW h/m^2 per year radiation [35].

Despite the favorable conditions in the country for harnessing solar power, as well as the falling prices in solar energy technology, the generation capacity is still non-existent, and the use of solar energy is in its early days. Many reasons can be attributed to the low use of solar energy for power and thermal utilities. This is partly due to the non-availability of the solar potential survey [36], which has been made available recently in 2016. Furthermore, the government over the decades has focused on thermal power generation as well as hydroelectric energy, and less attention was given to other renewable energy. The first policy devised for renewable energy came out in 2006. Even the policy guidelines lack the support to motivate the private sector and individuals to install and run solar energy farms [37]. In China Pakistan Energy Corridor (CPEC) energy project portfolio, a major focus is on coal power plants. Out of 17,000 MW of planned generation, only 1000 MW is to be generated from solar energy [38]. Of the 1000 MW, the 300 MW Quaid-e-Azam solar farm has been developed, and the remaining phases will be developed later. However, due to the sub-optimal location and other environmental and technical factors, further development of the solar farm is under review [39].

In addition to the Quaid-e-Azam solar farm, a few small-scale solar farms were developed and commissioned recently. The first on-grid solar power plant of 178 kW was commissioned in 2010 at the Pakistan Engineering Council building [40]. Later in 2016, the parliament building went green by installing 1 MW of solar panels. The National Electric Power Regulatory Authority (NEPRA) has provided a Feed-in Tariff (FiT) to both of these projects to sell their surplus energy to the national gird. Considering the potential the country has, and the intent of the State to develop solar farms, 28 solar companies with the cumulative capacity of 1556 MW have obtained Letters of Intent (LoI) from the Alternative Energy Development Board [41]. It is, however, important to verify the viability and the optimality of the proposed locations with respect to their solar potential.

5.2. Application of methodology

The methodology is applied to the geographical region of Pakistan. The administrative boundary of Pakistan is taken from Ref. [28], and the solar irradiance dataset and raster data for

[1] The dataset contains 1×1 km^2 areas with their corresponding coordinates and with their solar irradiation attributes.

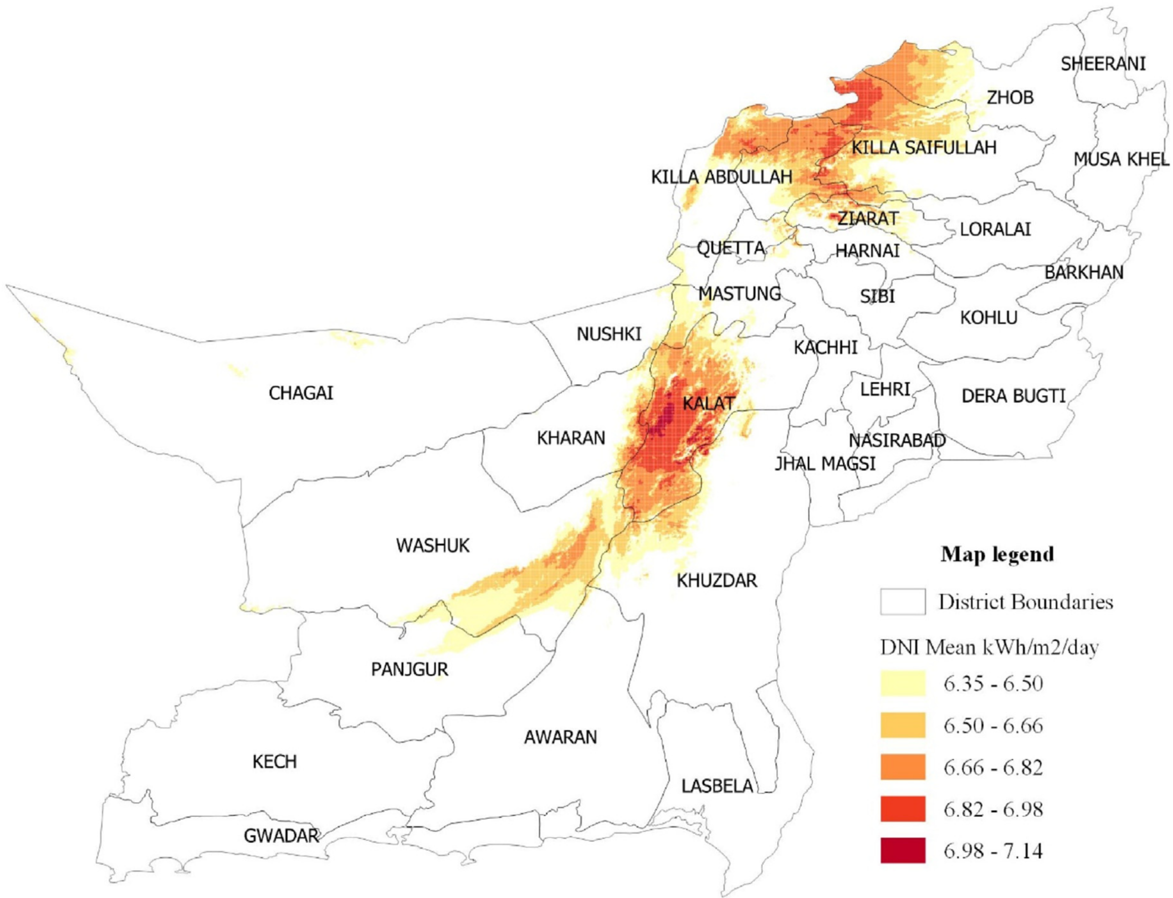


Fig. 3. Density-based clustering mechanisms.

elevation are imported from ESMAP dataset [29]. The three datasets are then loaded into the QGIS (Step 1, Fig. 2) and transformed into a vector grid of 1x1 km^2 (Step 2). The choice of 1x1 km^2 mesh size is optional and can be increased or decreased depending on the computational resources and scope of the problem. Minimum available raster data has the resolution of $250 \times 250\, m^2$ for the current area of interest. The output of the two-step procedure is the mapping of solar irradiance data in the geographical boundaries of Pakistan.

In the next phase of zone evaluation, descriptive statistics for 1x1 km^2 area are computed (Step 3). Multiple scenarios are drawn to put restriction on certain parameters of interest. Solar irradiance is restricted to top 5% (Step 4) in the initial scenario with terrain ruggedness variation allowed up to 25% (Step 5). The mean value of the top 5 percentile came out to be 6.4 $kWh/m^2/day$. Moreover, change in elevation having value less than 25% within a zone of 1x1 km^2 is applied. The previous two sequential steps generated a total of 1466 spatially identical zones in Pakistan which happened to be in Baluchistan province as shown in Fig. 4. These sites were scattered within the geographical boundaries of the province and are located in different administrative units called districts. Prominent zones can be found near Zhob in the north down to Kharan district in the south with dense patches of solar irradiance in Killa Saifullah and Kalat districts.

In the final phase, the density-based clustering is applied. Based on parameters of Algorithm presented in Fig. 3, the minimum solar farm size can be 9 km^2. Complete scanning of the data (1466 solar sites) yielded 23 dense clusters of sizes in the range of 10 km^2 to 289 km^2. More details about the number of clusters in the specific size range, the covered area of each cluster and corresponding location are given in the following section.

6. Results and discussion

Table 1 shows that with a size of 10 km^2, 4 solar farms having solar irradiance above 6.4 kWh/m^2/day can be developed, for instance, at the location with latitude and longitude values of (31.63101, 68.13644), which corresponds to the region near the Afghanistan border to the west of Zhob City in Baluchistan. The other two locations are near Qila Saifullah and Muslim Bagh, while the 4th location of size 10 km^2 lies in Surab Town near Kharan on either side of national highway N25. Similarly, 5 solar farms with sizes in the range of 11–20 km^2 (one 11, one 13, two clusters of 16 and one 15 km^2) have been identified in/near Panjgur, Mangocher, Kalat, between Panjgur and Nag, and halfway between Surab and Besima, respectively. Prominent among the solar farm sizes is 289 km^2 near Nag village.

Fig. 5 depicts the geographic locations of dense solar sites that are color coded to differentiate the sites. The black zones adjacent to the color-coded zones Fig. 5 (left) refer to regions which are outliers of the applied clustering algorithm. By outliers, we mean the regions that could not satisfy the clustering condition of algorithm as explained in Fig. 3. As is evident from Table 1, the largest cluster with a size of 289 km^2 lies at the geographic coordinates of latitude 27.142–27.462 and longitude 64.764–65.124.

The number of potential sites can be increased if one relaxes the criteria. For instance, if the solar irradiance criterion is kept at 90 percentile and above while keeping the same slope, the overall sites

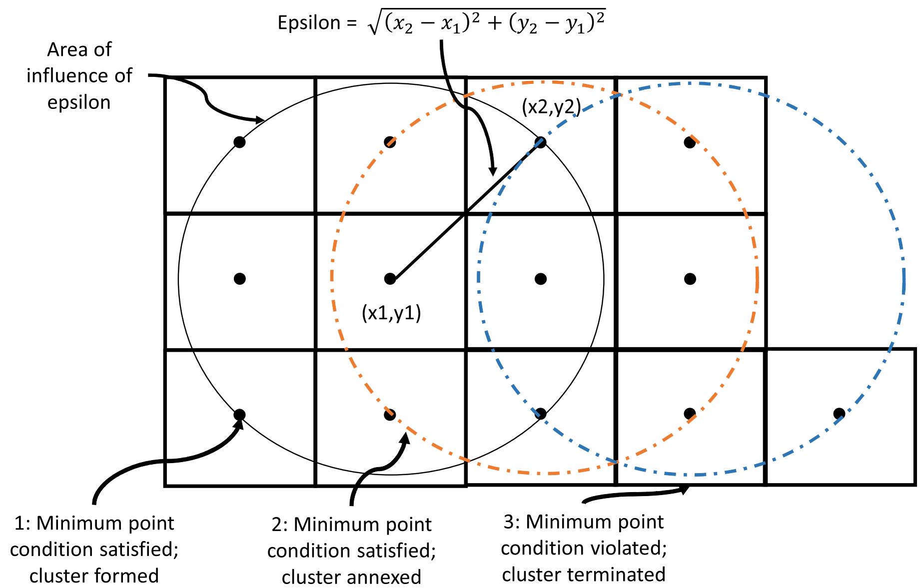


Fig. 4. Potential solar sites in Baluchistan.

Table 1
Clusters and their sizes and locations.

Size ranges (Km^2)	No. of Clusters	Size, Location (latitude, longitude)
10	4	(31.63101, 68.13644), (31.08501, 68.38544), (30.81301, 67.85444) (28.52601, 66.28344)
11–20	5	Size 11(27.10292, 64.59989), Size 13 (29.36585, 66.59905), Size 16(29.13201, 66.49756; 27.14826, 64.70194), Size 15(28.31734, 66.12311)
21–30	6	Size 21(30.79772, 67.67444), Size 24(27.52826, 65.17319), Size 26(29.35662, 66.51329, Size 29(28.25994, 66.01065; 29.43925,66.34444), Size 30(31.62701, 68.56411)
31–40	2	Size 36(31.75284, 68.49055), Size 37(28.74958, 66.40282)
41–50	3	Size 41(27.66079, 66.00322), Size 42(31.59725, 68.45396; 30.51701, 68.12587)
60–70	2	Size 63 (31.5963, 67.97857), Size 66 (31.31276, 67.74050),
289	1	(27.28672, 64.90299)

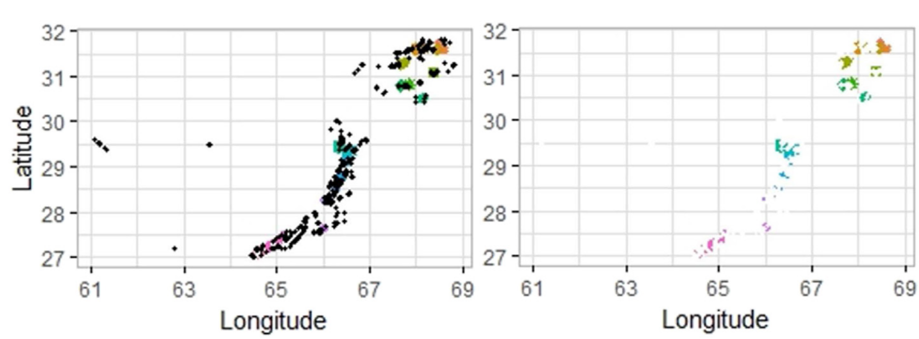


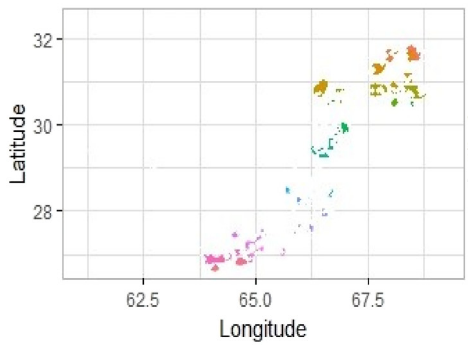
Fig. 5. Cluster locations with omitted sites as block dots (left) and refined locations (right).

area increases to 3300 km^2 which translates into 54 potential clusters.

The number of potential sites can be increased if one relaxes the criteria. For instance, if the solar irradiance criterion is kept at 90 percentile and above while keeping the same slope, the overall sites area increases to 3300 km^2 which translates into 54 potential clusters as shown in Fig. 6(left).

By further relaxing only the irradiance criterion to 80 percentile and above, the area expands to 6064 km^2 where 84 clusters of various sizes can be formed Fig. 6(right).

In short, substantial potential exists for solar power generation in the region particularly in the Baluchistan province of Pakistan. Coincidently most part of the potential sites in the province are not linked with the national grid and so Off-Grid solution using solar energy can be one alternative to provide electricity to communities there. Remoteness of most part of the province from national grid coupled with sparse population but abundance solar power generation potential in the province give energy planner enough reason to focus on solar energy development.



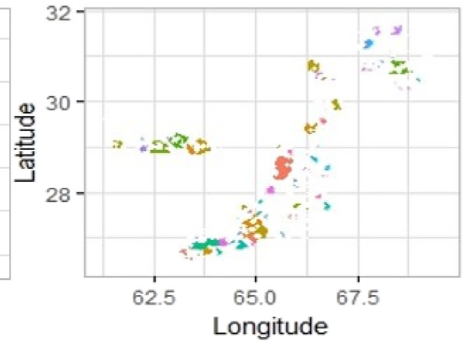


Fig. 6. Solar sites at 90 percentile and above (left), solar sites at 80 percentile and above (right).

7. Conclusion

This paper presented the application of GIS in combination with a density-based clustering technique to identify high irradiance solar sites suitable for solar farm development. The GIS was used to map the solar potential in the specified area of interest employing ESMAP data as spatial database for solar energy information. Multiple scenarios were developed while relaxing the solar irradiance criterion. The determine the precise farm size and their energy potential, clusters are formed using density-based clustering technique. The proposed methodology has been applied to the geographic boundaries of Pakistan. To identify and refine the clusters for solar farm development, two screening criteria were used: solar irradiance and terrain ruggedness. The former criterion filtered out high potential solar sites that have solar irradiance above 6.4 kWh/m2/day. However, to reduce the development cost of the solar farm, a terrain ruggedness criterion was applied (change in elevation $\leq$25%). In the next phase, density-based clustering technique was used to identify farm locations of minimum 9 km^2. This exercise yielded 23 clusters (farm sites) of varying sizes, which are located in the Baluchistan Province in the west of the country along/near the Afghanistan border. It is worthwhile to note that more clusters of smaller sizes (less than 9 km^2) can be added to the existing ones if vector grid having a mesh size of less than $1 \times 1\ km^2$ is used.

The choice of methodology not relying on criteria to skew the sites selection towards developed infrastructure is by design. The idea was to reduce the risk of eliminating potential viable solar sites in terms of high solar irradiance and least terrain ruggedness.

8. Policy implications

The findings of the study have many policy implications for Pakistan. It can contribute in the power market study aimed to identify the energy supply-demand for Pakistan till 2030. The study is in progress and is being conducted with the help of National Energy Commission (NEA) of China. In addition to the supply-demand, the focus of the study is also to increase the share of renewable energy in the energy mix of the country. Current energy mix is more in favor of fossil fuel. Solar energy presents a promising alternative for the sustainable energy mix for the country because of the availability of abundant solar resources especially along the western corridor of China Pakistan Economic Corridor (CPEC) covering Baluchistan Province of Pakistan.

Moreover, the methodology employed in the study does not rely on pre-defined criteria to identify suitable location as is the case in most of the studies which skew the results and the location decision outcomes are more towards developed infrastructure, eliminating the optimum locations from the universe of discourse. In under-developed regions like Baluchistan, the development strategy requires good infrastructure planning which in turns requires the identification of suitable locations for energy generation in addition to provision of water to promote social and economic development. The recently built 500 km Kachhi canal to irrigate 712,750 acres of land in the south-east districts of Bolan, Jhal Magsi, Dera Bugti of Baluchistan is supposed to bring green revolution in the province whose yield can spur food industry including food processing and distribution industry in the region. The industry and other economic development can be powered through solar farms that can be built in the neighboring districts of Kalat and Kharan.

Furthermore, Pakistan's national grid is mostly concentrated towards the eastern provinces whereas, most of the Baluchistan region is not well connected with the national electricity distribution network. Also, the region is sparsely populated. With a total land area of 347,190 km^2, its population is merely 12,34 million. In other words, connecting the sparsely populated communities and towns all over Baluchistan with the national electricity grid is not cost effective. The alternative to it, in the short run, is the off-grid or stand-alone solution in the form of solar farms for the communities and towns which not only makes economic sense but also appeals to current Pakistan's drive to go for green and clean initiatives. Limited interventions of incorporating small scale solar technology at village level have not succeeded due the lack of technical capabilities to maintain the solar equipment. In this scenario, the feasible solution is to have large centralized solar farms to achieve the economies of scale in addition to concentrating the technical and managerial capabilities on few larger sites.

Additionally, the expected demand for the Gwadar city by 2030 per Gwadar Development Authority documents (GDA) is 560 MW. The current reliance on Iran for power supply to the city faces challenges. The GDA is considering options of power supply to support the growth of the city in terms of economic development particularly the Gwadar Free Zone. One option is to connect Gwadar to national grid via Khuzdar passing through Kharan district which hosts various optimal locations for solar development including 289 km^2 solar farm in the Nag community along the national high way. Potential exists to power the city through solar in the short run while the surplus power will be evacuated to the national grid once it is built.

Another issue which our research can be applied is the use of solar power to run the water desalination operations at Gwadar, currently Gwadar is relying on the Ankara, Swad, and Shadi Kour dams to meet its demand in the relative short term. Future water needs have to be catered for and planned in advance so that a sustainable fresh water supply can be maintained. Our research methodology can be applied to identify the most suitable solar potential sites within the region and coupling them to the desalination operation so that future sustainable water supply to the city can be maintained.

In short, the current study identifies the solar potential of Pakistan using a structured and systematic methodology and provides guidelines to formulate policies with respect to energy infrastructure of the country. The study also highlights the prospects for the social and economic development of Baluchistan province. Other development projects like Kachhi Canal coupled with China-Pakistan Economic Corridor projects along the western route and Gwadar city in the presence of immense solar potential therein present ideal conditions to formulate comprehensive sustainable energy policy for the uplift of the Baluchistan province.

Future studies will rank the sites on the basis of desirability by applying progressively multiple screening criteria. The criteria will include but not limited to proximity to population, proximity to roads, transmission lines, water availability. The objective is to prioritize regions and make short, medium and long-term plans for solar sites development. Priority, in development, should be given to regions where infrastructure is relatively better to reduce sites development cost. At this point, our study could be use conventional MCDM to rank and classify sites based on infrastructure criteria. The authors also aim to work on the potential of rainwater harvesting along the western corridor in Baluchistan. Water availability coupled with abundance solar energy in the region will provide the necessary ingredients for economic activities and social uplift under China-Pakistan Economic Corridor.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

CRediT authorship contribution statement

Fahd Amjad: Conceptualization, Methodology, Project administration, Supervision, Writing - original draft, Writing - review & editing. **Liaqat Ali Shah:** Conceptualization, Methodology, Project administration, Supervision, Writing - original draft, Writing - review & editing.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.renene.2020.03.083.

References

[1] O. Edenhofer, R. Pichs-Madruga R, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, Adler, Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, 2014.
[2] D. Sperling, J.S. Cannon, Driving Climate Change: Cutting Carbon from Transportation, Elsevier, 2010.
[3] S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K.B. Averyt, M. Tignor, H.L. Miller, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, 2007, Cambridge University Press, Cambridge, 2007.
[4] I. Dincer, Renewable energy and sustainable development: a crucial review, Renew. Sustain. Energy Rev. 4 (2) (2000) 157–175.
[5] J.R. Janke, Multicriteria GIS modeling of wind and solar farms in Colorado, Renew. Energy 35 (10) (2010) 2228–2234.
[6] E. Kabir, P. Kumar, S. Kumar, A.A. Adelodun, K.-H. Kim, Solar energy: potential and future prospects, Renew. Sustain. Energy Rev. 82 (Feb. 2018) 894–900, https://doi.org/10.1016/j.rser.2017.09.094.
[7] J.A. Carrión, A. Espín Estrella, F. Aznar Dols, A.R. Ridao, The electricity production capacity of photovoltaic power plants and the selection of solar energy sites in Andalusia (Spain), Renew. Energy 33 (4) (Apr. 2008) 545–552, https://doi.org/10.1016/j.renene.2007.05.041.
[8] S.B. Darling, F. You, T. Veselka, A. Velosa, Assumptions and the levelized cost of energy for photovoltaics, Energy Environ. Sci. 4 (9) (Aug. 2011) 3133–3139, https://doi.org/10.1039/C0EE00698J.
[9] J. L. Sawin et al., "Renewables 2017 global status report".
[10] B. Parida, S. Iniyan, R. Goic, A review of solar photovoltaic technologies, Renew. Sustain. Energy Rev. 15 (3) (Apr. 2011) 1625–1636, https://doi.org/10.1016/j.rser.2010.11.032.
[11] Y. Charabi, A. Gastli, PV site suitability analysis using GIS-based spatial fuzzy multi-criteria evaluation, Renew. Energy 36 (9) (Sep. 2011) 2554–2561, https://doi.org/10.1016/j.renene.2010.10.037.
[12] H.Z. Al Garni, A. Awasthi, Solar PV power plant site selection using a GIS-AHP based approach with application in Saudi Arabia, Appl. Energy 206 (Nov. 2017) 1225–1240, https://doi.org/10.1016/j.apenergy.2017.10.024.
[13] M. Cannemi, M. García-Melón, P. Aragonés-Beltrán, T. Gómez-Navarro, Modeling decision making as a support tool for policy making on renewable energy development, Energy Pol. 67 (Apr. 2014) 127–137, https://doi.org/10.1016/j.enpol.2013.12.011.
[14] G. Ghasemi, Y. Noorollahi, H. Alavi, M. Marzband, M. Shahbazi, Theoretical and technical potential evaluation of solar power generation in Iran, Renew. Energy 138 (Aug. 2019) 1250–1261, https://doi.org/10.1016/j.renene.2019.02.068.
[15] I. Potić, R. Golić, T. Joksimović, Analysis of insolation potential of Knjaževac Municipality (Serbia) using multi-criteria approach, Renew. Sustain. Energy Rev. 56 (Apr. 2016) 235–245, https://doi.org/10.1016/j.rser.2015.11.056.
[16] J.P. Huang, K.L. Poh, B.W. Ang, Decision analysis in energy and environmental modeling, Energy 20 (9) (Sep. 1995) 843–855, https://doi.org/10.1016/0360-5442(95)00036-G.
[17] E. Ilbahar, S. Cebi, C. Kahraman, A state-of-the-art review on multi-attribute renewable energy decision making, Energy Strategy Rev. 25 (Aug. 2019) 18–33, https://doi.org/10.1016/j.esr.2019.04.014.
[18] S. Ozdemir, G. Sahin, Multi-criteria decision-making in the location selection for a solar PV power plant using AHP, Measurement 129 (Dec. 2018) 218–226, https://doi.org/10.1016/j.measurement.2018.07.020.
[19] J.M. Sánchez-Lozano, M.S. García-Cascales, M.T. Lamata, Comparative TOPSIS-ELECTRE TRI methods for optimal sites for photovoltaic solar farms. Case study in Spain, J. Clean. Prod. 127 (Jul. 2016) 387–398, https://doi.org/10.1016/j.jclepro.2016.04.005.
[20] D. Jun, F. Tian-tian, Y. Yi-sheng, M. Yu, Macro-site selection of wind/solar hybrid power station based on ELECTRE-II, Renew. Sustain. Energy Rev. 35 (Jul. 2014) 194–204, https://doi.org/10.1016/j.rser.2014.04.005.
[21] J.M. Sánchez-Lozano, C. Henggeler Antunes, M.S. García-Cascales, L.C. Dias, GIS-based photovoltaic solar farms site selection using ELECTRE-TRI: evaluating the case for Torre Pacheco, Murcia, Southeast of Spain, Renew. Energy 66 (Jun. 2014) 478–494, https://doi.org/10.1016/j.renene.2013.12.038.
[22] A. Asakereh, M. Soleymani, M.J. Sheikhdavoodi, A GIS-based Fuzzy-AHP method for the evaluation of solar farms locations: case study in Khuzestan province, Iran, Sol. Energy 155 (2017) 342–353, https://doi.org/10.1016/j.solener.2017.05.075.
[23] J.C. Mourmouris, C. Potolias, A multi-criteria methodology for energy planning and developing renewable energy sources at a regional level: a case study Thassos, Greece, Energy Pol. 52 (Jan. 2013) 522–530, https://doi.org/10.1016/j.enpol.2012.09.074.
[24] N. Sarmiento, S. Belmonte, P. Dellicompagni, J. Franco, K. Escalante, J. Sarmiento, A solar irradiation GIS as decision support tool for the Province of Salta, Argentina,, Renew. Energy 132 (Mar. 2019) 68–80, https://doi.org/10.1016/j.renene.2018.07.081.
[25] J.R.S. Doorga, S.D.D.V. Rughooputh, R. Boojhawon, Multi-criteria GIS-based modelling technique for identifying potential solar farm sites: a case study in Mauritius, Renew. Energy 133 (Apr. 2019) 1201–1219, https://doi.org/10.1016/j.renene.2018.08.105.
[26] J. Arán Carrión, A. Espín Estrella, F. Aznar Dols, M. Zamorano Toro, M. Rodríguez, A. Ramos Ridao, Environmental decision-support systems for evaluating the carrying capacity of land areas: optimal site selection for grid-connected photovoltaic power plants, Renew. Sustain. Energy Rev. 12 (9) (Dec. 2008) 2358–2380, https://doi.org/10.1016/j.rser.2007.06.011.
[27] A. Fausett, M.E. Celebi, An accelerated nearest neighbor search method for the K-means clustering algorithm, in: FLAIRS Conference, 2013.
[28] Global administrative areas | boundaries without limits [Online]. Available: http://gadm.org/ [Accessed: 04-Oct-2017].
[29] Global solar atlas - downloads [Online]. Available: http://globalsolaratlas.info/downloads/pakistan [Accessed: 15-Aug-2017].
[30] QGIS Development Team, QGIS Geographic Information System. Open Source Geospatial Foundation.
[31] J. MacQueen, Some methods for classification and analysis of multivariate observations, in: Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, 1967, pp. 281–297.
[32] L. Kaufman, P.J. Rousseeuw, Partitioning around medoids (program pam), in: Find. Groups Data, 1990, pp. 68–125. Introd. Clust. Anal.
[33] A. Kassambara, Practical Guide to Cluster Analysis in R: Unsupervised Machine Learning, vol. 1, STHDA, 2017.
[34] Pakistan Energy Yearbook, Hydrocarbon development inistitute of Pakistan [Online]. Available: https://www.hdip.com.pk/contents.php?cid=31, 2018. (Accessed 7 January 2020).
[35] Prospects of renewables penetration in the energy mix of Pakistan, Renew. Sustain. Energy Rev. 29 (Jan. 2014) 693–700, https://doi.org/10.1016/j.rser.2013.08.083.
[36] S. Stökler, C. Schillings, B. Kraas, Solar resource assessment study for Pakistan,

Renew. Sustain. Energy Rev. 58 (2016) 1184–1188.
[37] U. Zafar, T. Ur Rashid, A.A. Khosa, M.S. Khalil, M. Rashid, An overview of implemented renewable energy policy of Pakistan, Renew. Sustain. Energy Rev. 82 (Feb. 2018) 654–665, https://doi.org/10.1016/j.rser.2017.09.034.
[38] Energy | China-Pakistan economic corridor (CPEC) official website [Online]. Available: http://cpec.gov.pk/energy. (Accessed 5 January 2020).
[39] 14-Energy 22-05-17 FINAL composed - 14-Energy.pdf [Online]. Available: http://www.finance.gov.pk/survey/chapters_17/14-Energy.pdf. (Accessed 15 August 2017).
[40] Pakistan Engineering Council, PEC alternate energy resource portal [Online]. Available: http://energy.pec.org.pk/about.php. (Accessed 10 June 2018).
[41] AEDB, Alternative energy development board [Online]. Available: http://www.aedb.org/. (Accessed 17 June 2018).