Energy Strategy Reviews 50 (2023) 101205

Contents lists available at ScienceDirect

Energy Strategy Reviews

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

Optimal location selection for a distributed hybrid renewable energy system in rural Western Australia: A data mining approach

Rain Holloway [a], Darren Ho [a], Colin Delotavo [a], William Yau Xie [a], Iman Rahimi [a], Mohammad Reza Nikoo [b], Amir H. Gandomi [a,c,*]

[a] Faculty of Engineering and Information Technology, University of Technology Sydney, Australia
[b] Department of Civil and Architectural Engineering, Sultan Qaboos University, Muscat, Oman
[c] University Research and Innovation Center (EKIK), Óbuda University, 1034, Budapest, Hungary

ARTICLE INFO

Handling Editor: Mark Howells

Keywords:
Distributed hybrid renewable energy system
Western Australia
Data mining
Clustering
Location selection

ABSTRACT

The adverse effects of coal and gas energy production with the subsequent rapid increase in energy consumption emphasize the importance for Australia to adopt more renewable energy sources to counteract these dismissive contributions to climate change. This work presents a data mining approach for optimally selecting the best locations for installing a distributed hybrid renewable energy generation system for rural regions in Western Australia. The K-Means and K-Medoids clustering algorithms were used to divide the constructed dataset into clusters. In total, 69 locations were selected for the overall dataset, proceeding with the filtering process. The returned cluster data were graphically rendered on a Western Australia map for the region. The clustering algorithms were evaluated using the Dunn index, such that K-Means performed to a higher degree than K-Medoids, given our dataset's nature. After passing the generated clusters to HOMER software to generate the potential wind and solar energy output for each centroid, K-Medoids produced a set of locations that generated higher solar and wind energy on average. However, due to the reduced internal validation, K-Medoids might not be as valuable as K-Means, it does not cluster the data points very well, and within-cluster location energy requirements are not considered in our study.

1. Introduction

Global energy consumption has seen a dramatic increase in recent years. Non-renewable sources remain some of the most inefficient energy sources and significantly contribute to climate change [1]. Despite the clear environmental benefits of renewable energy, its adoption is greatly hindered by its reliance on natural factors, such as hours of solar irradiation, wind speed, and wind strength. As such, the energy production by renewable energy is inconsistent [2].

Several sectors have been identified for renewable energy concepts, including sunlight, wind, rain, tides, waves, and geothermal heat [3,4]. Nowadays, many countries in the world use renewable energy, and it is expected that the growing renewable energy markets will continue and improve powerfully in the future [5,6]. Australia, as a developed country, possesses the potential to use renewable energies; furthermore, solar energy has been known as a sound source of renewable energy in Australia [7,8].

When it comes to solar energy, the utilization of solar photovoltaic (PV) systems has gained global recognition and acceptance in recent years for generating hot water. These systems operate efficiently, as evidenced by research conducted by Refs. [3,9]. Additionally, solar thermal heating and cooling systems harness thermal energy from sunrise and find applications in commercial settings [10]. Conversely, solar panels are influenced by weather conditions, making their energy production reliant on sunlight. Consequently, the effectiveness of solar energy collection can be noticeably impacted by rainy or cloudy days, affecting the power output of PV systems.

Also, wind energy, as one of the most useable renewable energies in the world, possesses the most negligible negative impact on the climate compared to other energies, such as fossil fuels. This is because wind energy generates zero-emission energy. Demolli et al. [11] presented a technique for predicting long-term wind power with the aid of machine learning algorithms, utilizing daily wind speed data. The findings demonstrated the successful application of machine learning in efficiently forecasting wind power values, even for locations different from those used to train the model.

* Corresponding author. Faculty of Engineering and Information Technology, University of Technology Sydney, Australia.
E-mail address: Gandomi@uts.edu.au (A.H. Gandomi).

https://doi.org/10.1016/j.esr.2023.101205
Received 24 June 2023; Received in revised form 1 September 2023; Accepted 13 September 2023
Available online 22 September 2023

While Australia does utilize certain fossil fuels, a substantial portion of its energy and electricity production is attributed to clean sources such as solar and wind energy. This trend is driven by the numerous favorable environmental effects associated with these renewable energy forms, setting them apart from fossil fuels [12,13]. On the flip side, there exist numerous drawbacks when it comes to the utilization of wind energy. For instance, crucial factors like site-specific weather data encompassing wind speed and load prerequisites play a pivotal role, constituting a notable challenge in the adoption of wind energy. Moreover, the evaluation of an existing system's efficiency mandates the availability of appropriate weather data [14].

Furthermore, a primary drawback associated with sole reliance on renewable energies is their intermittent nature, as their output hinges on weather conditions. Consequently, to enhance overall energy generation, it is advisable to implement a combination of these energy sources. To determine the optimal configuration involving the quantity of photovoltaic panels and wind turbines, a suitable optimization mechanism is essential. An added advantage of the hybrid system is the reduction in the need for battery storage and diesel backup, contributing to its appeal. Thus, the hybridization of photovoltaic and wind turbine technologies emerges as an optimal solution for mitigating energy requirements. Hybrid renewable energy systems typically involve the integration of multiple renewable energy types to bolster system efficiency and uphold a robust equilibrium in energy supply [15–18]. [19] evaluated the performance of wind-solar hybrid renewable energy systems using the Loss of Power Supply Probability (LPSP) index. The study concludes that the LPSP index can be used to optimize the design of hybrid renewable energy systems.

Finding the best locations to install a distributed hybrid renewable energy system is a challenging task, and data mining-based approaches are known as a powerful tool to tackle this problem [20–22]. There exist numerous works based on data mining approaches in the field, for example, geographical information system (GIS)-based spatial data mining approach [23], a data mining-based optimal demand response program [24], heuristic algorithms [25], and other approaches [26,27].

Regarding the planning of hybrid renewable installations, there exists a wealth of literature on the subject ([28–31]. For example [32], proposed a decision support system, which could assist decision-makers in locating single-sources as well as hybrid renewable energy system installations to meet the requirement of energy production. In another study [33], proposed an optimal design and techno-economic assessment of a hybrid renewable energy system on Sebesi Island, South Lampung Regency, Indonesia. The above-mentioned study proposed the planned hybrid system from a technological, economic, and environmental perspective [34]. explored the viability of a hybrid renewable energy system (HRES) consisting of photovoltaic (PV) and biomass energy for a rural village in India. The HRES was designed, optimized, and analyzed for economic, renewable energy, and environmental aspects [35]. discussed the optimal design of off-grid hybrid renewable energy systems (HRESs) by addressing the levelized cost of energy (LCOE) and CO2 emissions simultaneously using the ε-constraint method and particle swarm optimization (PSO) algorithm. The study investigates various combinations of components such as photovoltaic panels, wind turbines, batteries, hydrogen, and diesel generators to develop cost-emissions Pareto fronts of different HRES configurations [36]. proposed a Geographic Information Systems (GIS) and Multi-Criteria Decision Analysis (MCDA)-based methodology to locate optimal FPV sites, exemplified by selecting San Giovanni Dam in Sicily among seven watersheds, demonstrating benefits for high-temperature regions [37]. addressed an extensive review, which examines trends, models, and challenges in hybrid renewable energy research spanning 2000 to 2022. They synthesized diverse energy systems, compared methodologies, addressed uncertainty, and outlined future prospects.

The novelty of this study lies in its pioneering focus on addressing a critical research gap within the realm of hybrid renewable energy optimization. While various studies have explored potential renewable energy sites across Australia, this study distinguishes itself by exclusively targeting the unique landscape of rural Western Australia. By harnessing cutting-edge clustering algorithms and leveraging intricate geographic attributes, such as solar irradiation and wind speeds, we aim to unlock the untapped potential of these underutilized rural regions. Our approach offers a fresh perspective on renewable energy integration, providing a transformative pathway toward sustainable energy independence for these communities. To achieve our research goals, the study's objectives are as follows.

- Identify areas with high solar exposure, temperature, and frequency of sunny days, making them potentially optimal for solar energy generation.
- Identify areas with high wind speed and frequency that might be suitable for wind energy generation.
- Apply clustering algorithms to group locations based on several factors, including existing energy usage and solar and wind energy generation potential.
- Utilize the clustered data to determine the most suitable locations for installing hybrid renewable energy systems.

This paper presents a data mining approach for optimally selecting the best locations for installing a distributed hybrid renewable energy generation system for rural regions in Western Australia. The remainder of the paper is structured as follows. Section 2 presents the motivation for this study. Section 3 shows the research methodology. Section 4 describes data collation. Sections 5 and 6 illustrate modeling and results. Section 7 depicts the validation and verification of the proposed model, while section 8 provides the discussion. Finally, section 9 presents the conclusion, addressing limitations and potential directions for future research.

2. Motivation

According to the Department of Primary Industries and Energy and the Department of Human Services and Health, rural zones are divided into three different categories: large rural centers (25,000–99,999), small rural centers (10,000–24,999), and other rural areas (<10,000). Thus, all rural zones in Australia are considered to have a population of less than 100,000. Rural towns contain approximately 28% of Australia's total population[1], which correlates to a rural total of 7,210,516 based on Australia's overall population of 25, 750, 200 (September 2021) (Statistics 2020). Even though rural towns in Australia have adequate energy connectivity and transmission, the provided electricity is generated from predominantly non-renewable resources, namely fossil fuels (76%), including coal (54%), gas (20%), and oil (2%).[2] Only 7% of electricity generation from 2019 to 2020 was from renewable sources.[3] As such, the aim of this study is to extract the optimal locations for hybrid renewable energy plants utilizing solar photovoltaic and wind energy sources within Western Australia (WA), ultimately allowing maximal utilization of renewable energy resources. Urbanised locations within WA, such as Perth, Bunbury, or Geraldton, were not considered in this work. Wind (17.5%), solar photovoltaic (18.1%), and solar hot water (4.4%) energy contribute a combined 40% of renewable energy production, equating to only 2.8% of Australia's total energy generation

[1] Australian Government: Department of Industry, Science, Energy and Resources. (2020). States and territories. https://www.energy.gov.au/data/states-and-territories.

[2] Australian Government: Department of Industry, Science, Energy and Resources. (2020). States and territories. https://www.energy.gov.au/data/states-and-territories.

[3] Australian Government: Department of Industry, Science, Energy and Resources. (2020). States and territories. https://www.energy.gov.au/data/states-and-territories.

(167.6 PJ)[4]. With this in mind, Western Australia's energy consumption from 2019 to 2020 was 1271.9 PJ, which is 21.2% of total energy consumption in Australia.[5] The major disparity between Australia's energy generated from solar and wind plants compared to WA's total energy consumption contributed to our decision to exclude urbanised locations from our study.

While Australia does rely on certain amounts of fossil fuels, the adoption of clean energy sources like solar and wind energy holds substantial importance in the nation's energy and electricity generation. The evaluation of viable sites for implementing renewable energy systems in rural areas of Australia has been a subject of investigation in several studies. In this study, we aim to address a research gap in the hybrid renewable energy optimization field by investigating the potential locations for installing a distributed hybrid renewable energy generation system in rural Western Australia. While several studies have explored the potential locations for renewable energy systems in rural regions across Australia [38–43], to the best of our knowledge, this is the first study specifically focusing on rural Western Australia. Our study recognizes the importance of not only evaluating the technical feasibility and energy output of hybrid renewable energy systems but also understanding their potential to positively impact surrounding populations, economies, and the environment. By assessing areas with high solar exposure and wind speeds for optimal energy generation, we aim to not only contribute to renewable energy integration but also shed light on the significant advantages that these systems offer, such as reduced electricity fees, lower maintenance costs, and decreased greenhouse gas emissions.

By focusing on rural Western Australia, we intend to fill this gap in the literature by providing a comprehensive analysis that encompasses both the technical and socio-economic aspects of renewable energy integration. We believe that this holistic approach is essential for making informed decisions and promoting sustainable energy solutions that benefit not only the energy sector but also the overall well-being of communities.

Each location dramatically affects the performance of renewable energy systems. Western Australia is one location that has received minimal attention compared to other areas. Western Australia also had some of the highest levels of solar irradiation and high wind speeds – this, combined with its low renewable energy usage, made it highly attractive for the topic of our report. Renewable energy systems can significantly benefit rural areas. Numerous articles concentrate solely on analyzing the efficiency and constraints of established energy systems, neglecting to delve into the ways in which these systems contribute to the well-being of the local communities. Specifically, the economic advantages linked to renewable energy systems often go unaddressed, including factors such as the elimination of electricity fees, diminished maintenance expenses, and the absence of greenhouse gas emissions.

Rural areas typically have lower energy demands than suburban areas, meaning it is possible to completely power a rural area using renewable energy. Herein, we intended to extract the optimal locations (locations with high amounts of consistent solar exposure and wind speed) to install hybrid renewable energy plants utilizing solar photovoltaic and wind energy sources. With specific rural areas within Western Australia, we applied HOMER software (free trial version) and various clustering techniques to determine which areas would be able to support hybrid renewable energy systems and provide adequate power to surrounding regions.

Some questions regarding the current study are listed as follows.

- What areas within rural Western Australia experience high amounts of solar irradiation?
- What areas within rural Western Australia experience high wind speeds and wind frequency?
- How do different clustering techniques (K-Means, K-Medoids) affect the organization of clusters?
- What degree of energy output (kWh/yr) should we expect from a hybrid renewable energy system within each cluster?
- What locations within Western Australia best support installing hybrid renewable energy systems?

In conclusion, this study seeks to bridge the research gap in rural Western Australia by providing valuable insights into the potential for renewable energy integration, thus contributing to the sustainable development and energy independence of rural communities in the region.

3. Research methodology

In the methodology section, our primary contribution lies in the novel approach we employed to identify optimal locations for installing hybrid renewable energy systems in rural Western Australia. While we utilized HOMER software and various clustering techniques, including K-Means and K-Medoids algorithms, the main focus of our methodology was to demonstrate how these tools were applied to the specific context of rural Western Australia. To achieve each of our research objectives and answer our research questions, the experimental method involves the analysis of geographical locations and their associated meteorological data through centroidal clustering algorithms. The methodology for this research started with the data collection of rural locations in Western Australia from online sources, including 'Townsites LGATE-248 W A GeoJSON' as well as 'Back4App List of Australian Cities'.

The clustering algorithms categorize locations into groups based on similarities between their field values. The K-Means and K-Medoids clustering algorithms were used to divide the constructed dataset into clusters. The optimal value of K ($K = 7$) was determined by utilizing the elbow curve as well as the silhouette score. Such groups contain a centroid location that can be within the dataset or not, which is identified as the optimal location for installing a solar and wind hybrid renewable energy system. In total, 69 locations were selected for the overall dataset, proceeding with the filtering process. The returned cluster data were graphically rendered on a Western Australia map for the region. Using the Dunn index, the clustering algorithms were evaluated, such that K-Means (0.1458) performed to a higher degree than K-Medoids (0.0715), given the nature of our dataset. The potential energy of this system was calculated by simulating the system through the use of the HOMER Pro software (free trial version) using a control configuration. After passing the generated clusters to HOMER software to generate the potential wind and solar energy output for each centroid, K-Medoids produced a set of locations that generated higher solar and wind energy on average. These data were then used in a cross-comparison between the K-Means and K-Medoids clustering algorithms to determine their effectiveness for datasets within this field of study (Fig. 1). In summary, our contribution to the study goes beyond the mere description of software usage. Instead, it lies in the thoughtful application and adaptation of software tools to address the specific research objectives and challenges in exploring the potential for renewable energy systems in rural Western Australia.

4. Data collation

Our research utilized two main datasets: The Townsites LGATE-248 W A GeoJSON dataset and the Back4App dataset. The former provides an area polygon for each town or city within Western Australia. This polygon is defined by several coordinate points, allowing for a detailed geographic representation of each area. Moreover, it labels each town or

[4] Australian Government: Department of Industry, Science, Energy and Resources. (2020). States and territories. https://www.energy.gov.au/data/states-and-territories.

[5] Western Australia Government. (2020). Townsites (LGATE-248). Retrieved from https://data-downloads.slip.wa.gov.au/LGATE248/GeoJSON.

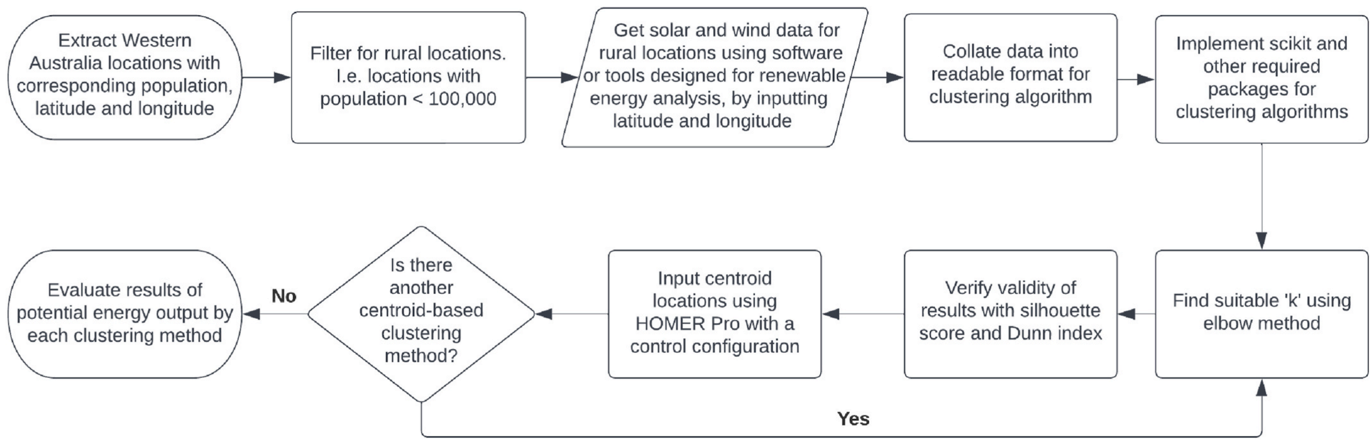


Fig. 1. Flowchart of research methodology.

city's name, providing a list of locales within the Western Australia region. On the other hand, the Back4App dataset, referenced in Energy (2020), offers a broader picture, encompassing all cities across Australia. Unlike the LGATE-248 dataset, it does not provide a polygon representation but assigns a singular coordinate to each location. A unique feature of the Back4App dataset is its inclusion of population data for each city, offering demographic insights not present within the LGATE-248 dataset. For our needs, we pursued a strategy of retrieving the common locations between them. This was necessitated by the lack of specificity in the Back4App dataset concerning a location's situatedness within or outside Western Australia - it merely provided information that a location was somewhere within Australia. However, the LGATE-248 dataset was more precise, including only localities found within WA. By cross-referencing these datasets, we could identify which locations listed in the Back4App dataset were indeed in Western Australia while gaining access to corresponding population data. Using a CSV manipulator package in Python, the combined data from LGATE-248 and 'Back4App List of All Cities in Australia' was compiled and filtered to only include cities containing a population of less than 100,000. Based on the analysis of data sources, there are 69 rural locations within WA. These cities and associated latitudes and longitudes were used to extract monthly averages of the clearness index[6], solar irradiation[7], and wind speed[8] at each location. The data for solar global horizontal irradiance (GHI) feature a clearness index and daily radiation (kWh/m^2/day). The clearness index measures the clearness of the atmosphere, defined as the surface radiation divided by the extra-terrestrial radiation. It is a fraction of the solar radiation transmitted through the atmosphere to strike the surface and ranges between 0 and 1. The clearness index has a high value under clear, sunny conditions and a low value under cloudy conditions[9]. Global horizontal irradiance is the total solar radiation incident on a horizontal surface. HOMER software incorporates sophisticated algorithms and databases that encompass a wide range of geographic, climatic, and technical parameters. It allows us to simulate and assess renewable energy potential within specific regions by utilizing historical weather data, geographical coordinates, and other relevant inputs.

HOMER uses solar GHI to calculate the flat-panel photovoltaic output (energy generated from solar panels)[10]. The only feature returned for wind data is the wind speed (m/s). Any type of configuration, not including the latitude and longitude of the location, does not affect the solar and wind data from NASA's Prediction of Worldwide Energy Resource (POWER) database. While our study does not involve direct equipment-based measurements, our utilization of the HOMER software ensures a robust and scientifically rigorous analysis of hybrid renewable energy potential in rural Western Australia. This approach allows us to provide valuable insights into optimal locations for renewable energy integration without the need for physical data collection devices.

The process of collecting and collating the data in a form suitable for using clustering, which in this case was in a spreadsheet, as a CSV file, is explained below.

- Collecting the locations required for clustering: Locations were filtered to only include rural localities — towns/cities with a population of less than 100,000. We extracted locations from Townsites LGATE-245 W A GeoJSON, Back4App list of Australian Cities.
- The HOMER Pro software was implemented with inserted latitude and longitude, and the solar and wind resources were downloaded for the respective locations, which can be found in the Resources Solar GHI and Wind tabs. Within these resources, the three aforementioned variables can be seen with monthly averages in Fig. 2.

These results were manually mapped into a spreadsheet with the structure shown.

- **Rural location:**
  - o **Name:** Location (1), Location (2), …
  - o **Latitude:** Latitude (1), Latitude (2), …
  - o **Longitude:** Longitude (1), Longitude (2), …
- **Clearness Index:**
  - o **Months 1–12:** CI for Jan … Dec (1), CI for Jan … Dec (2), …
- **Daily Radiation:**
  - o **Months 1–12:** DR for Jan … Dec (1), DR for Jan … Dec (2), …
- **Avg. Wind Speed:**
  - o **Months 1–12:** WS for Jan … Dec (1), WS for Jan … Dec (2), …

Each record in the spreadsheet contains information related to one location, including latitude, longitude, and the monthly average of each solar and wind variable.

[6] https://www.homerenergy.com/products/pro/docs/3.11/how_homer_calculates_clearness_index.html.

[7] https://www.homerenergy.com/products/pro/docs/3.11/generating_synthetic_solar_data.html.

[8] https://www.homerenergy.com/products/pro/docs/3.11/generating_synthetic_wind_data.html.

[9] HOMER Energy. Global Horizontal Irradiance (GHI). Retrieved from https://www.homerenergy.com/products/pro/docs/latest/globalhorizontalirradianceghi.html.

[10] Scikit-learn. 2.3 Clustering. Retrieved from https://scikit – learn.org/stable/modules/clustering.htmlk – means.

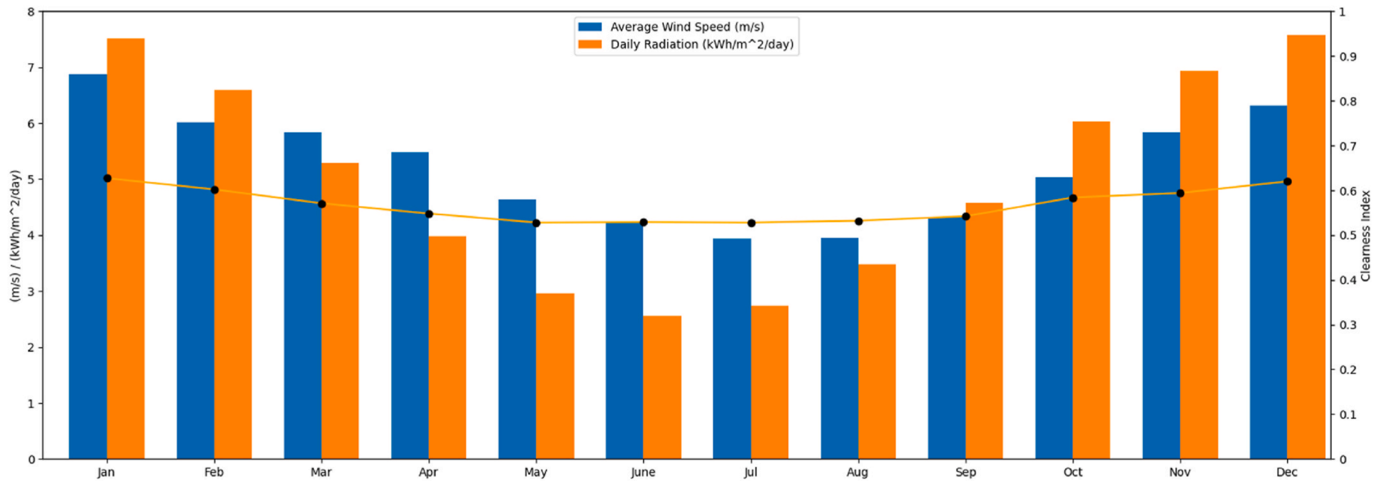


Fig. 2. Clustered column chart indicates monthly average wind speed (m/s) and average solar irradiation (kWh/m2/day). The line graph indicates the clearness index, with corresponding intervals noted on the right y-axis.

5. Modeling

Clustering is an unsupervised machine learning technique of identifying and grouping similar data points in larger datasets; in simpler words, the aim is to segregate groups with similar traits and assign them into clusters. In this research, as we searched for the most optimal location for the plants, we decided to utilize centroid clustering techniques to identify the optimal location within each cluster. Centroid clustering machine learning techniques are iterative clustering algorithms in which the data points are grouped according to how close a data point is to the centroid of the clusters. The K-Means clustering algorithm is a widespread algorithm that falls into this type. In these models, the number of clusters required at the end must be declared beforehand, which makes it vital to have prior knowledge of the dataset. Centroidal clustering methods differ from other clustering methods as they generate or obtain a central data point used to describe a given cluster.

The adoption of both centroid clustering methodologies, encompassing the K-Means algorithm alongside the K-Medoids algorithm, was influenced by their inherent capabilities and alignment with the goals of our research. The K-Means algorithm, which identifies central data points within clusters, aligns seamlessly with our goal of pinpointing optimal locations for hybrid renewable energy plants. Its simplicity, interpretability, and computational efficiency are well-suited for our emphasis on energy generation, enabling us to efficiently analyze and interpret the clustering results. Similarly, the K-Medoids algorithm, a variant of K-Means, was chosen for its ability to identify representative data points that exemplify the characteristics of each cluster, providing a robust and intuitive alternative. This selection is motivated by our study's focus on rural Western Australia and the need for an effective methodology to address our research objectives. By incorporating both centroid clustering techniques and the K-Medoids algorithm, we ensure a comprehensive exploration of potential hybrid renewable energy plant locations, aligning with our overarching goal of sustainable energy integration and development in rural regions. However, other types include distribution-based clustering and hierarchical clustering. Distribution-based clustering assumes all data points belong to a Gaussian distribution and groups samples accordingly; as the distance from the distribution center increases, the probability that a point belongs to the distribution decreases. Hierarchical clustering creates a tree of clusters such that a given cluster can be a child of at most one cluster and be related to several child clusters. This type of clustering is suited for classifying and categorizing points into classes rather than arbitrarily grouping.[11] It is important that our study implemented centroidal clustering methods as they return a set of cluster centroids, which represent potential locations for a network of hybrid renewable energy systems that provide maximum energy generation.

5.1. Choosing K

Due to the nature of our chosen clustering algorithms, the chosen value of K, i.e., the number of clusters, was predetermined when running the algorithms. We adopted the elbow method to select the optimal K value as well as the silhouette coefficient to validate our selected value.

The utilization of both the elbow method and the silhouette coefficient for determining the optimal number of clusters (K) and validating the clustering outcomes is motivated by their complementary contributions. The elbow method offers a visual indication of the point where adding more clusters ceases to significantly enhance clustering performance, aiding in selecting a reasonable K value. Meanwhile, the silhouette coefficient provides a quantitative measure of cluster quality by assessing both cohesion and separation, ensuring that clusters are not only internally coherent but also well-separated from each other. This combined approach enhances the robustness of our clustering analysis, striking a balance between cluster compactness and distinctiveness. While alternative methods like the Davies-Bouldin index or Gap statistic could be considered, the elbow method and silhouette coefficient offer a pragmatic and widely accepted strategy, facilitating both intuitive interpretation and quantitative evaluation of clustering solutions within the context of our research.

5.1.1. Elbow curve method

The elbow method is a commonly-used heuristic that can potentially determine the optimal number of clusters in a dataset. The elbow curve graph plots the inertia for each value of K. Inertia is the summed value of each squared distance from each sample to its cluster center within each cluster. The smaller the value, the more coherent and compact the clusters are. When finalizing results from clustering algorithms, inertia should be minimized. The general method works on picking the "elbow" of the curve as the number of clusters to use. The "elbow" of the curve is defined as the point that initiates a linear decrease in inertia for each proceeding value of K. The rationale is that we use the cut-off point

[11] Google. Clustering Algorithms. Retrieved from https://developers.google.com/machine%20-%20learning/clustering/clustering%20-%20algorithms.

where the diminishing returns are no longer worth the additional cost. In the context of clustering, we should choose a number of clusters where adding another cluster does not improve the data modeling.

5.1.2. The silhouette coefficient

The silhouette coefficient is a popular metric that is used to calculate the goodness of a clustering technique. Specifically, the silhouette value measures how similar an object is to its own cluster compared to other clusters. This value is calculated using the (*a*) mean intra-cluster distance and (*b*) mean nearest-cluster distance for each sample. For a given sample, the silhouette coefficient is calculated using the following mathematical expression:

$$\text{Silhouette coefficient} = (\mathrm{b} - \mathrm{a})/\max(\mathrm{a}, \mathrm{b}) \tag{1}$$

Its value is contained within the range [1, −1], and it can also be used to assist in the calculation of the most optimal value of K. If the silhouette coefficient is calculated and averaged across each sample within the dataset for each value of K, it can be used as a secondary parameter next to the elbow curve method for choosing the optimal K. As the maximized value of the silhouette coefficient correlates with high validity for clustering algorithms, the true maximum and local maximums are preferred when choosing K. If the elbow curve is difficult to interpret, the local maximums of the silhouette coefficient can be observed.

5.2. Selection of optimal K-value for centroidal clustering algorithms

Herein, K-Means and K-Medoids are both centroid-based clustering algorithms, where determined centroids are the selected locations for the hybrid renewable system. K-Means is a centroid-based algorithm that calculates the distance to assign a point to a cluster. In K-Means, each specific cluster is associated with a centroid. The key objective of the K-Means algorithm is to reduce the sum of distances between the points and their respective cluster centroid. In Fig. 3, it is difficult to distinguish the point at which the inertia begins to decrease linearly. When analysing the curve, the point at which it appears to start decreasing linearly is at K = 5. However, the corresponding value in the silhouette coefficient graph (Fig. 4) displays the true minimum value for all values of K. We wish to maximize the silhouette coefficient, so this value of K is not suitable.

A limitation of the K-Means algorithm is that the centroids are the mean of points present in that cluster rather than the actual point. Hence, we decided to explore another method known as the K-Medoids algorithm (Figs. 5 and 6), which selects actual data points as cluster centers (medoids) and, thereby, allows for more straightforward interpretability of the cluster centers.

Applying the same methodology for K-Medoids, we attained the corresponding inertia values by looping through different values of K. The elbow curve (Fig. 5) demonstrates an inflection point when K = 7—this value is further validated by the silhouette coefficient (Fig. 7), indicated by the maximum value at K = 7. Based on the selection process for K-Medoids, we may analogously choose to adopt this value for K-Means. The suitability of K = 7 for K-Means is further supported by its silhouette score (Fig. 4), which depicts a local maximum at K = 7.

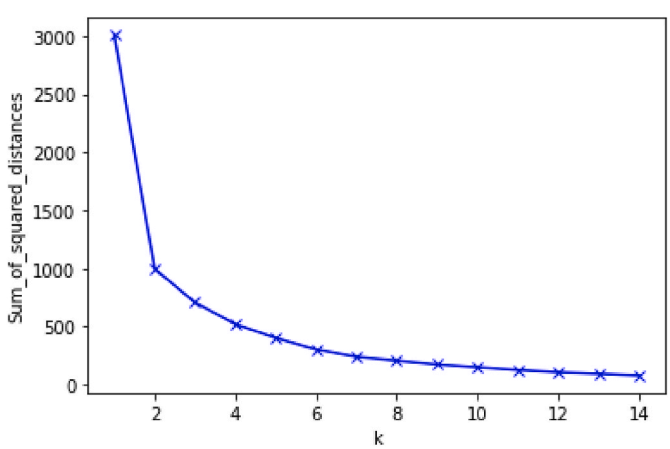


Fig. 3. Elbow curve for K-Means.

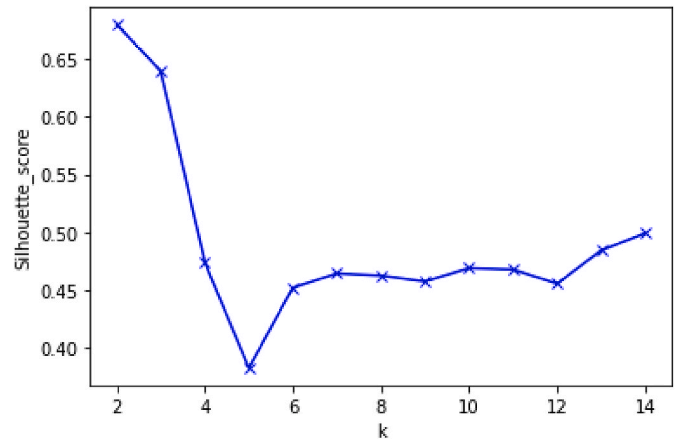


Fig. 4. Silhouette score for K-Means.

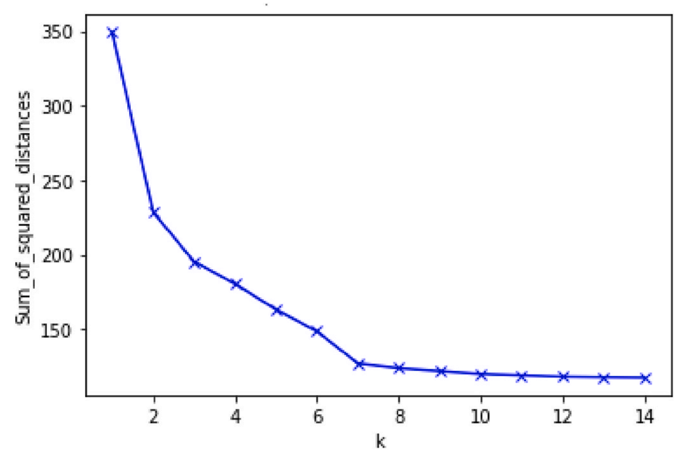


Fig. 5. Elbow curve for K-Medoids.

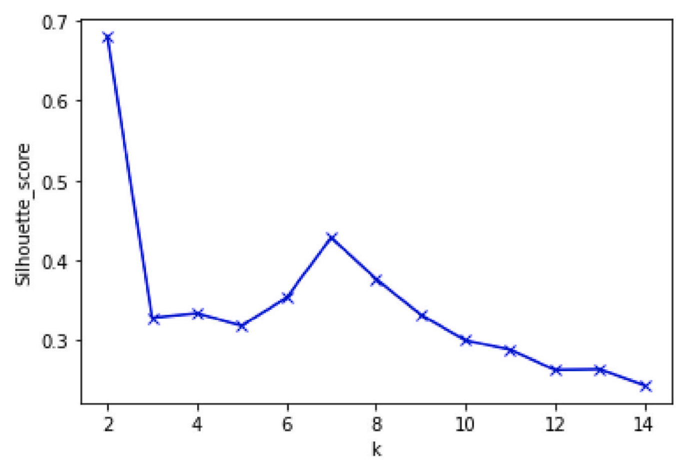


Fig. 6. Silhouette score for K-Medoids.

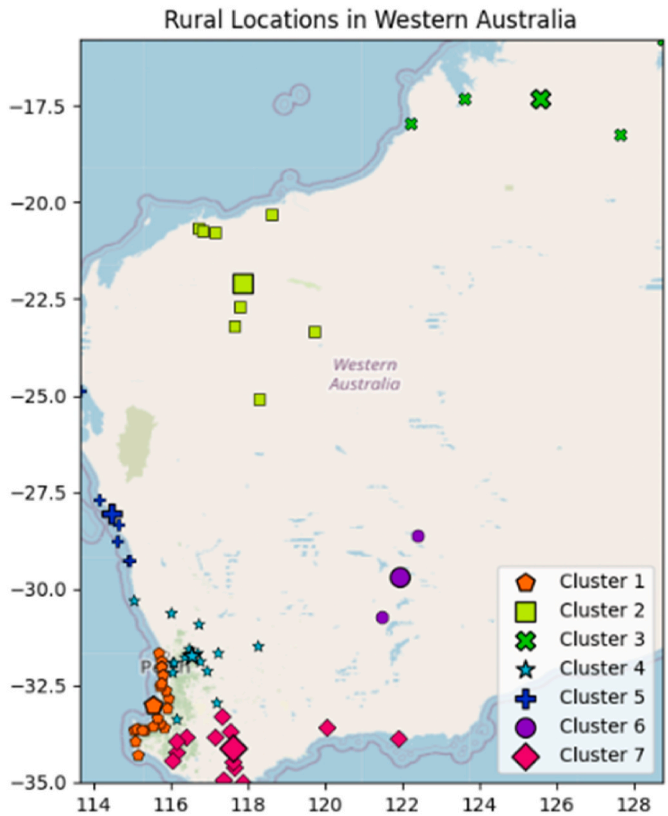


Fig. 7. Generated cluster map for K-Means (rural locations in Western Australia).

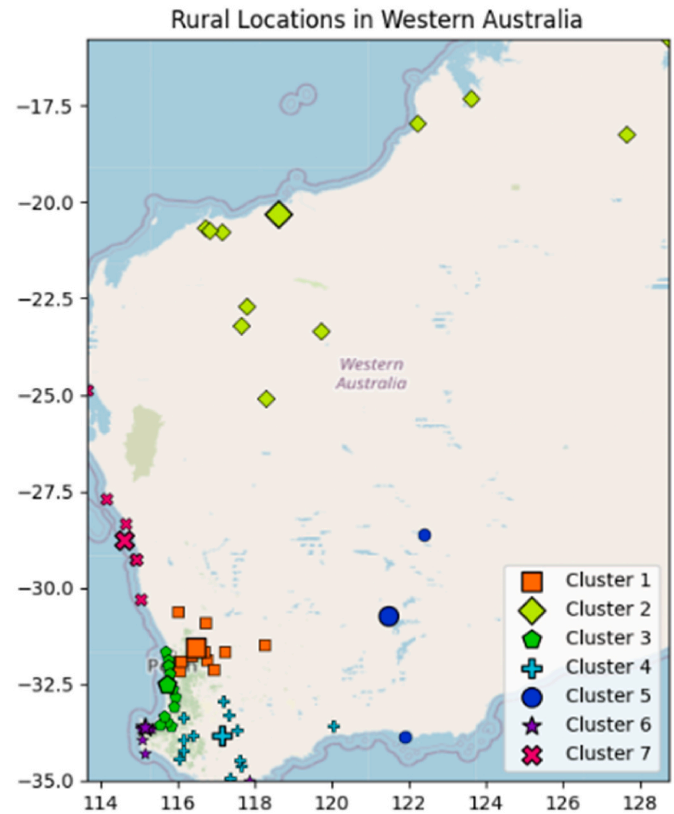


Fig. 8. Generated cluster map for K-Medoids (rural locations in Western Australia).

6. Results

The following subsections present the results of the clustering algorithms along with the simulated solar and wind energy production per annum.

6.1. Geographical cluster maps

Figs. 7 and 8 illustrate the generated clusters for K-Means and K-Medoids, respectively, which can be distinguished via their colors. Centroids within a cluster are clearly seen due to their increased size and thick black outline. Fig. 9 provides a graphical representation of the merged clustering algorithms. Figs. 7 and 8 exhibit several salient differences despite an overall similarity. One such distinction, mentioned earlier, lies in the selection of centroids; while K-Means determines a centroid based on the mean of points within a given cluster, K-Medoids uses an actual data point as the centroid. This difference is particularly noticeable in the purple/circle cluster in Fig. 8 and the dark blue/circle cluster in Fig. 8. Another key feature is the singular lime/diamond cluster, found in northern WA, depicted in Fig. 8, while the K-Means algorithm generated two separate clusters, lime/square and green/X, in the same northern region, as shown in Fig. 7. A similar trend can be observed in the southwest of WA, where K-Means produced a singular cluster, orange/pentagon (Fig. 7), as opposed to the K-Medoids algorithm that separated the same general region into two distinct clusters, green/pentagon and purple/star (Fig. 8).

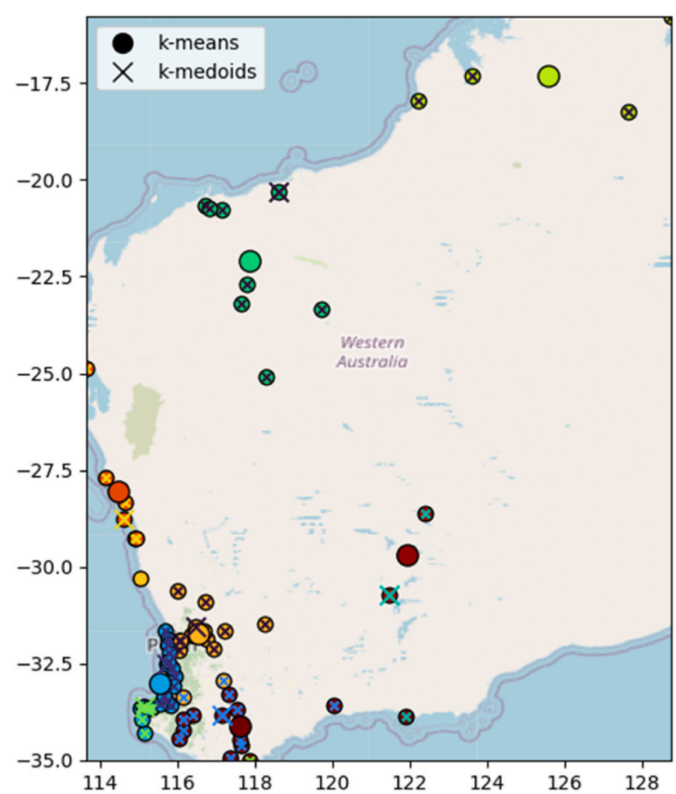


Fig. 9. Generated cluster map with merged clustering methods.

6.2. Energy results from HOMER

All results were retrieved using the HOMER Pro Microgrid Analysis Tool (free trial version). The settings for the configuration are listed below.

- Wind Turbine: Generic 3 kW
- Solar Panel: Generic Flat Plate PV
- Battery: Generic 1-kWh Lithium-Ion Battery

The selected parameters, namely the Generic 3 kW Wind Turbine, Generic Flat Plate PV Solar Panel, and Generic 1-kWh Lithium-Ion Battery, were chosen based on a balance between representing typical technologies and conducting a feasibility assessment within the scope of our study.

While a design process for the maximization of renewable sources within the identified clusters could indeed be pursued, it is important to consider our research's practical constraints and objectives. Our primary focus was to evaluate the potential energy output and feasibility of hybrid renewable energy systems within specific rural areas of Western Australia. By utilizing representative values for the wind turbine, solar panel, and battery, we aimed to offer a preliminary analysis of the energy potential, rather than an exhaustive optimization of each individual component.

Given the wide range of possible configurations and the complexities of optimizing renewable sources at a microgrid scale, we made a deliberate choice to use generic parameters to provide a foundational assessment of the renewable energy potential. This approach allowed us to focus on the overarching goal of identifying optimal locations for hybrid renewable energy systems, while acknowledging that more detailed design processes and optimization studies could be pursued in future research to fine-tune the system parameters.

6.2.1. *K-means clustering results*

The bar graph in Fig. 10 represents the culmination of the K-Means clustering results, in which the simulated energy production is in kWh per annum. Each unit located on the y-axis is an exponent of 7 (i.e., eˆ7). It is pertinent to point out that the green/X and lime/square centroids have a clear difference in solar panel energy production, which will be clarified in the analysis of the results. Additionally, an average of all the centroids is located in the last side-by-side bars. Table 1 presents the

Table 1
K-Means: Cluster centroids' colour/shape and location.

Cluster Centroid Colour/Shape	Coordinates
Orange/Pentagon	−33.03464714, 115.56349762
Lime/Square	−22.10418125, 117.8515975
Green/X	−17.3153575, 125.5691025
Light Blue/Star	−31.74793733, 116.53323933
Dark Blue/Plus	−28.04230667, 114.48807
Purple/Circle	−29.68602, 121.93785
Magenta/Diamond	−34.14196538, 117.63305846

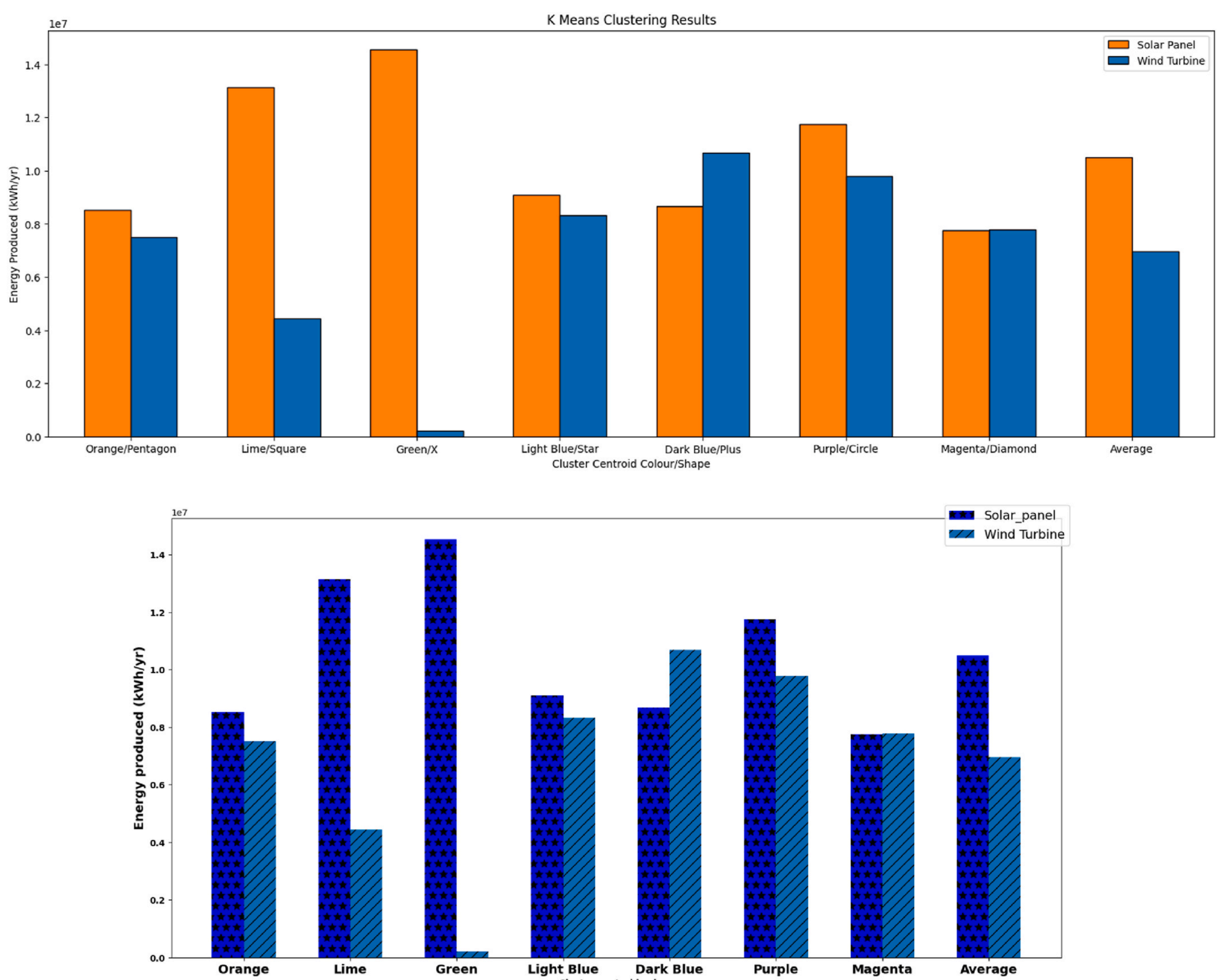


Fig. 10. K-Means clustering results of produced energy by wind and solar.

cluster centroids' colour/shape and location of K-Means.

6.2.2. K-medoids clustering results

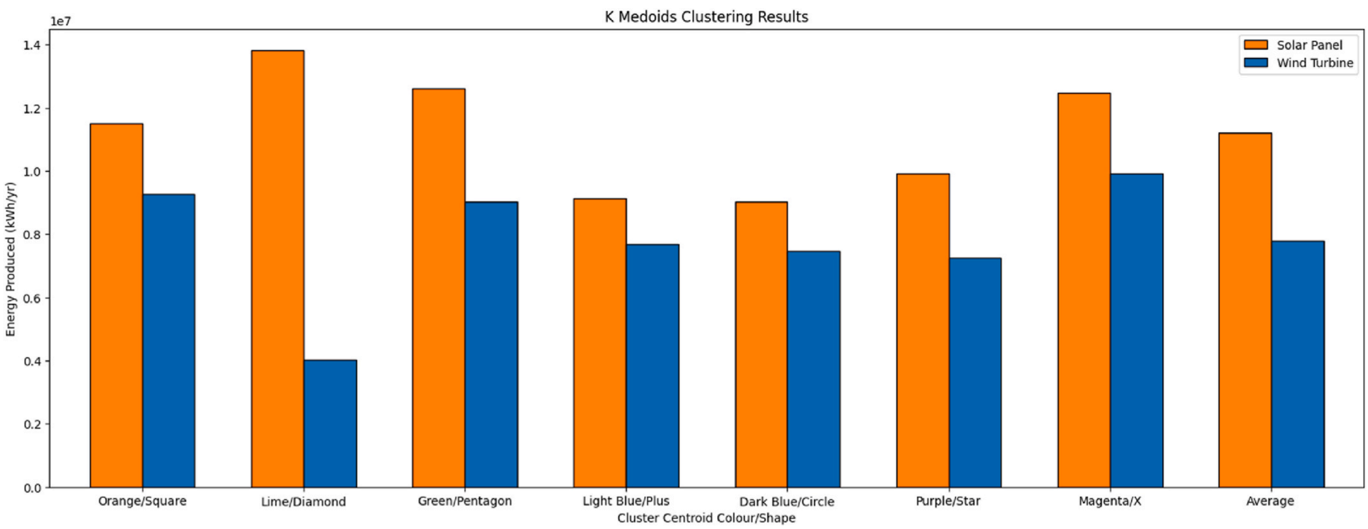
Similarly, Fig. 11 presents the bar graph for the K- Medoids clustering results. Unlike the bar graph of K-Means, it is evident that there is only one significant disparity in solar panel and wind turbine energy production for the lime/diamond centroid. Table 2 shows the cluster centroids' colour/shape and location of K-Medoids.

6.3. Results analysis

Across the large majority of clusters, solar panel energy production was higher than their wind turbine counterpart. From Figs. 7–9, the exception to this can be seen in the dark blue/plus cluster (−28.04230667, 114.48807) and magenta/diamond cluster (−34.14196538, 117.63305846) from the K-Means cluster centers. The K-Medoids clusters contain no such instances of wind turbine energy production exceeding solar panel energy production. From K-Means, the green/X cluster center (−17.3153575, 125.5691025) has the single highest solar panel energy production across both sets of clusters, producing 14, 550, 168 kW h of energy per year. Inversely, this center's wind turbine energy production is the lowest out of any center. This cluster sits in the center of the Wunaamin Conservation Park, a largely uninhabited area in the northern region of Western Australia. Comparing the light blue/star K-Means center (−31.74793733, 116.53323933) with the orange/pentagon K-Medoids center (−31.54977, 116.46743), there is a dramatic difference in energy production despite a minor difference in location. Solar energy production increased by 2,409,144 kWh, while wind energy production increased

Table 2
K-Medoids: Cluster centroids' colour/shape and location.

Cluster Centroid Colour/Shape	Coordinates
Orange/Square	−31.54977, 116.46743
Lime/Diamond	−20.31215, 118.61059
Green/Pentagon	−32.5269, 115.7217
Light Blue/Plus	−33.8305, 117.15946
Dark Blue/Circle	−30.74614, 121.4742
Purple/Star	−33.6356, 115.14899
Magenta/X	−28.77897, 114.61459



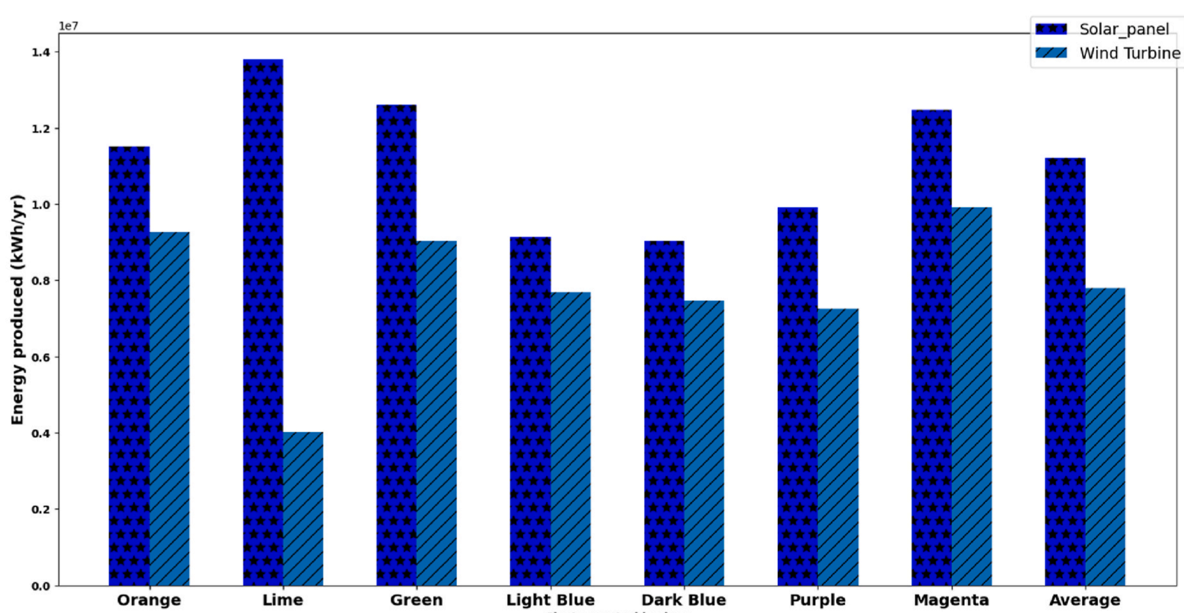


Fig. 11. K-Medoids clustering results of produced energy by wind and solar.

by 1,769,897 kWh. Compared to the dark blue/plus K-Means cluster (−28.04230667, 114.48807), the magenta/X K-Medoids cluster (−28.77897, 114.61459) produced dramatically higher solar energy of 3,802,432 kWh. However, wind energy production is higher in the dark blue/plus K-Means center, with an increase of 776,848 kWh. Overall, the highest combined energy production from solar and wind energy was obtained by the magenta/X K-Medoids cluster center (located at −28.77897, 114.61459), which points to the coastal city of Geraldton. From the K-Means clustering results, the highest combined energy production was achieved by the purple/circle center (−29.68602, 121.93785), which sits in a largely uninhabited region approximately 630 km northeast of Perth.

7. Validation and verification

The results obtained using the HOMER software and clustering algorithms were further validated and verified heuristically by examining and comparing cluster attributes of each clustering method. In other words, the results were compared to experimental or synthetic data to assess how well the clustering algorithms performed. As previously mentioned, the silhouette coefficient, or silhouette score, is a form of validation for selecting an optimal K value. In this section, we describe the Dunn index, another validation technique that evaluates the performance of clustering algorithms for artificially made datasets.

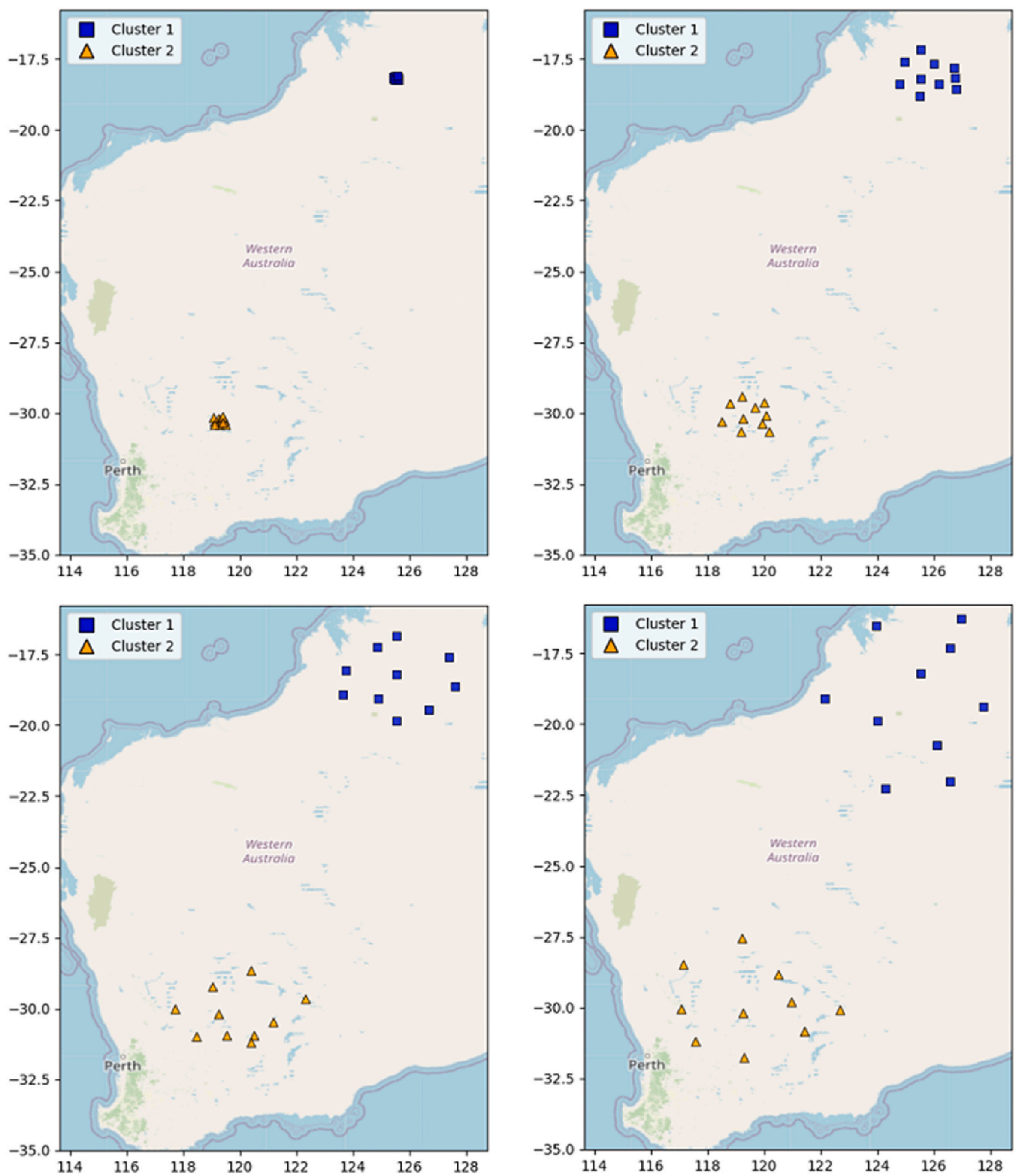


Fig. 12. Artificial clusters generated using dummy datasets. Each cluster contains 10 data points. Dunn index values are 1.5652, 0.4794, 0.2155, and 0.1516, respectively.

7.1. Cluster attributes

Compactness, connectedness, and separation of the cluster partitions are three cluster attributes reflected in our internal validation measures. Herein, compactness and separation are pertinent components of the Dunn index. Specifically, compactness measures how close data points are within the same cluster (inertia/within-cluster variation). This value is also utilized in the silhouette coefficient to assist in choosing the optimal K value. Separation measures how well-separated a cluster is from other clusters.

7.2. Dunn index

The Dunn index is a metric that can be used to evaluate clustering algorithms, whereby a higher value is considered better. The Dunn index was chosen for its ability to quantify cluster compactness and inter-cluster separation, applying to two practical implications of our study. More compact clusters promote efficient energy distribution, as energy can be shared efficiently within each cluster to meet local demands. Simultaneously, the strategic placement of centroids within the respective cluster serves to minimize transmission distances, thereby minimizing energy loss. The index is calculated as the lowest intercluster distance (i.e., the smallest distance between any two cluster centroids) divided by the highest intracluster distance (i.e., the largest distance between any two points in any cluster). We then utilized the finalized model above to generate our clusters, cluster centers, and cluster maps. To further validate our results, we passed our generated labels into a customized algorithm to calculate the Dunn index of 0.1458 for K-Means and 0.0715 for K-Medoids. We can assess the validity of these algorithms based on how closely our measurements adhere to our expectations and understanding of clustering algorithms. Accordingly, there is a high correlation between higher index values and overall well-defined clusters. However, the Dunn index has its limitations. While it can represent the overall degree of clustering in our algorithm, it does not provide a complete picture of the distribution of the clusters. In other words, the Dunn index may not necessarily reflect the shape of the clusters. This can be likened to how an average may produce misleading results if the data is extremely skewed. Nevertheless, the high correlation does give a good indication of the accuracy of clustering in our dataset. To verify the correlation between more well-defined and less-defined clusters, we compared the Dunn index with corresponding artificial datasets of varying degrees of spread data points. Fig. 12 presents examples of such artificial clusters, ranging from extremely tight to sparsely distributed clusters. Further research may be required to comprehend the subtleties between one Dunn index and another, the relevance of the difference between two Dunn index values, and the nature of the clusters. Additionally, we assessed the extent to which our measurement results correspond to other valid clustering measurements. To verify this, we analyzed the artificial clusters to compare the results of the Dunn index based on the spread of clustering and data distribution. As mentioned previously, by comparing these clusters with our results, we observed a high correlation between higher Dunn index values and more well-defined clusters. Therefore, our Dunn index values accurately represent the nature of our clusters and their relation to other clusters. In cases where there are more complex cluster shapes, we would also compare the Dunn index of artificial datasets that emulate the clusters so that we can more accurately assess the validity of the Dunn index on unusual cluster shapes. Therefore, based on our assessment of the last two points, we can verify that our results are accurate and valid for the purposes of this study (Fig. 12).

8. Discussion

With the dramatic increase in global energy consumption, there is an increased urgency to research the higher efficacy of different renewable energy sources. Even though renewable energy offers clear environmental benefits, there are some key flaws mainly stemming from its reliance on natural factors, such as hours of solar irradiation, wind speed, and wind strength. Herein, our proposed method identifies the optimal locations (particularly locations with high amounts of consistent solar exposure and wind speed) to install hybrid renewable energy plants, utilizing solar photovoltaic and wind energy sources. We collected data from various online sources, determined the wind and solar energy produced from our chosen rural locations using HOMER, and finalized our dataset for the K-Means and K-Medoids clustering algorithms. Our results show that K-Means achieved a higher Dunn index than K-Medoids, where a higher value means a better cluster. We also utilized the elbow curve method and silhouette coefficient to choose the correct K value. In addition, by representing the clusters and their centroids on cluster maps, we can see that the clusters generated by K-Means are logically more sensible in terms of energy production compared to K-Medoids. Hence, it can be suggested that K-Means will be the better clustering algorithm to utilize in our specific use case. Our clustering results were then inserted into the HOMER software to generate the potential energy output from these proposed locations. On average, solar panel and wind turbine energy production from K-Medoids was higher than the cluster centers found from K-Means clustering. The green K-Means cluster center (−17.3153575, 125.5691025) has the single highest solar panel energy production out of all the clusters from both algorithms, while the highest combined energy production from solar and wind energy can be seen at the magenta K-Medoids cluster center (−28.77897, 114.61459). While the K-Medoids algorithm has a lower Dunn index and does not cluster the dataset as well as K-Means, it generated more solar and wind energy on average compared to K-Means. Although K-Medoids might generate more energy on average, a specifically given cluster may only provide energy to a few locations, while another cluster may provide energy to more locations. This can be observed in the purple K-Means cluster, which sits in a largely uninhabited region 630 km northeast of Perth that generates the highest combined energy production, whereas the dark blue K-Means cluster located in the center of a coastal city produces limited energy. With a poorer clustering score from dark blue K-Medoids, the energy generated might not be able to be provided to enough locations and, thus reducing its overall value.

The proposed methodology focused on identifying optimal locations for hybrid renewable energy plants in rural Western Australia, considering solar photovoltaic and wind energy sources. Herein, we also employed the K-Means and K-Medoids clustering algorithms for regional grouping and explored the potential energy output of the proposed locations. Regarding the energy production comparison between K-Means and K-Medoids, our findings showed that K-Medoids, on average, generated higher solar and wind energy. The variation in energy production can be attributed to the geographical and climatic differences between regions, emphasizing the need for location-specific analyses. An interesting observation is the discrepancy between energy production and cluster locations. For instance, the purple K-Means cluster, situated in a sparsely populated area northeast of Perth, exhibited the highest combined energy production. In contrast, the dark blue K-Means cluster, located in a coastal city center, had limited energy production. Such discrepancies highlight the importance of not solely relying on clustering scores but considering the actual energy supply capacity to nearby regions.

To summarize, our study provides valuable insights into renewable energy optimization in rural Western Australia. The comparison of clustering algorithms and the assessment of energy production in specific locations contribute to the ongoing research in this field. However, it is crucial to recognize the limitations of our study, such as the reliance on specific datasets and the regional focus on Western Australia. Future research could benefit from broader datasets and comparisons across various geographical regions to gain a more comprehensive understanding of renewable energy site selection.

9. Conclusion

The clustering and energy generation analysis results are inconclusive due to the disparity between the better clustering performance of K-Means and the potential energy production of K-Medoids. We hypothesized that the clustering algorithm with the greatest Dunn index and silhouette coefficient values would return the maximum potential energy production. Despite our inconclusive results, we obtained insight into the effectiveness of both clustering algorithms in terms of their validity scores, establishing that K-Means has a greater clustering performance for geographical and meteorological data. The obtained cluster centroids and associated cluster map can be used as an assistive resource for the national planning of energy generation systems to aid in the ongoing global aim of increasing the use of renewable energy while diminishing the use of coal, gas, and other non-renewables. This data has not been generated within the field of hybrid renewable energy in Western Australia, thus promoting and providing an opportunity for deeper analysis and improved research.

One limitation of this study is our lack of focus on the overall cost of building such energy plants. Since we mainly investigated which locations had high amounts of solar energy and wind exposure, we did not manage to consider the overall cost or annual operation costs of such a plant. Therefore, these topics could be a potential direction for future research. Also, another limitation of the methodology presented in this paper is HOMER software. This software has a detailed economic calculation that takes account of all economic factors, but the detailed calculation is not revealed. Also, it is just a black box with limited flexibility in changing the input data without the capability to check and change the economic calculation method. It is worth mentioning that this study focuses on a specific region of Australia. The insights gained from this research provide a valuable foundation for future studies within the country.

There are several facets of our methodology that we did not include in our study to maintain an achievable study scope. These factors can be improved in future research. Specifically, we failed to recognize the importance of parameter optimization for the clustering methods. If given the opportunity, the K-Means and K-Medoids algorithms may have performed to a higher degree and returned a greater average silhouette coefficient and Dunn index if we had chosen the optimal parameters. Furthermore, we did not consider optimizing the installment and operational costs of the hybrid renewable energy systems; this is due to the fact that a co-component of this topic is the inclusion of energy demand and cost of transport for each location. Researching each location's energy requirements would provide satisfactory results to optimize installment and operational costs. In an optimized system, each cluster's hybrid renewable system would generate approximately the summed required energy of each location in the given cluster.

Future studies incorporating the facets mentioned above would highly contribute to this field as it would be suitable for real-world application in Western Australia. Also, the work could be expanded to evaluate the feasibility of using hybrid renewable energies in different areas of Australia, considering different case studies. Moreover, it would be valuable to investigate how the approach works for larger test cases with thousands of variables and what's the performance of the method in a large-scale optimization problem. Besides, the uncertainty is another challenging topic for this case as future work, and we will try to answer these questions: "how the approach can handle the uncertainties" and "What's the model's sensitivity to the change of computational parameters?"

Credit author statement

Darren Ho, Rain Holloway, Colin Delotavo, and William Yau Xie: Methodology, Formal analysis, Software, Visualization, Writing- Original draft preparation, Data curation, Investigation. Iman Rahimi: Ideas, Supervision, Revising, Editing, Mohammad Reza Nikoo: Revising, Editing, Visualization. Amir H. Gandomi: Supervision, Investigation, Editing.

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.

Data availability

Data will be made available on request.

References

[1] G.M. Shafiullah, Hybrid renewable energy integration (HREI) system for subtropical climate in central queensland, Australia, Renew. Energy 96 (2016) 1034–1053.
[2] Mukund K. Deshmukh, S Deshmukh Sandip, Modeling of hybrid renewable energy systems, Renew. Sustain. Energy Rev. 12 (1) (2008) 235–249.
[3] Omar Ellabban, Haitham Abu-Rub, Frede Blaabjerg, Renewable energy resources: current status, future prospects and their enabling technology, Renew. Sustain. Energy Rev. 39 (2014) 748–764.
[4] Javad Nematian, Iman Rahimi, Feasibility study of using renewable energies in Iranian seas: a comparative study, in: Renewable Energy vol. 189, of Statistics, Australian Bureau. 2020. "National, State and Territory Population." The Bureau Canberra, 2022, pp. 383–391.
[5] Sven Teske, Alex Fattal, Christine Lins, Hullin Martin, Laura E. Williamson, Renewables Global Futures Report: Great Debates towards 100% Renewable Energy, 2017.
[6] Iman Rahimi, Javad Nematian, A bibliometric analysis on optimization solution methods applied to supply chain of solar energy, Arch. Comput. Methods Eng. 1–19 (2022).
[7] Mark A. Hemer, Richard Manasseh, Kathleen L. McInnes, Irene Penesis, Tracey Pitman, Perspectives on a way forward for ocean renewable energy in Australia, Renew. Energy 127 (2018) 733–745.
[8] Iman Rahimi, Mohammad Reza Nikoo, Amir H. Gandomi, Techno-economic analysis for using hybrid wind and solar energies in Australia, Energy Strategy Rev. 47 (2023), 101092.
[9] John Byrne, Lado Kurdgelashvili, Manu Mathai, Ashok Kumar, J. Yu, Xilin Zhang, Jun Tian, Rickerson Wilson, G. Timilsina, World Solar Energy Review: Technology, Markets and Policies, Center for Energy and Environmental Policies Report, 2010.
[10] Alireza Bahadori, Chikezie Nwaoha, A review on solar energy utilisation in Australia, Renew. Sustain. Energy Rev. 18 (2013) 1–5.
[11] Halil Demolli, Ahmet Sakir Dokuz, Alper Ecemis, Murat Gokcek, Wind power forecasting based on daily wind speed data using machine learning algorithms, Energy Convers. Manag. 198 (2019), 111823.
[12] G M Joselin Herbert, Selvaraj Iniyan, E. Sreevalsan, S. Rajapandian, A review of wind energy technologies, Renew. Sustain. Energy Rev. 11 (6) (2007) 1117–1145.
[13] A. Ajanovic, R. Haas, N. Nakicenovic, Economic analysis of production and use of hydrogen from solar energy, wind, hydropower and biomass, in: ISES Solar World Congress 2007 vol. 4, ISES 2007, 2007, pp. 2496–2500.
[14] Hermann-Josef Wagner, Introduction to wind energy systems, in: EPJ Web of Conferences, 2018, p. 5, 189.
[15] Jinze Li, Pei Liu, Li Zheng, Optimal design and techno-economic analysis of a solar-wind-biomass off-grid hybrid power system for remote rural electrification: a case study of west China, Energy 208 (2020), 118387.
[16] Mohammad Hossein Jahangir, Seyed Ali Mousavi, Rad Mohammad Amin Vaziri, A techno-economic comparison of a photovoltaic/thermal organic rankine cycle with several renewable hybrid systems for a residential area in rayen, Iran, Energy Convers. Manag. 195 (2019) 244–261.
[17] Jahangir, Mohammad Hossein, Ramin Cheraghi, Economic and environmental assessment of solar-wind-biomass hybrid renewable energy system supplying rural settlement load, Sustain. Energy Technol. Assessments 42 (2020), 100895.
[18] Clementine L. Chambon, Tanuj Karia, Philip Sandwell, Jason P. Hallett, Techno-economic assessment of biomass gasification-based mini-grids for productive energy applications: the case of rural India, Renew. Energy 154 (2020) 432–444.
[19] A. Ecemiş, H. Demolli, M. Gökçek, A.Ş. Dokuz, Evaluation of wind-solar hybrid renewable energy systems using Loss of power supply probability, in: International Conference on Engineering Technologies, 2019, pp. 377–381. ICENTE'19.
[20] Samayan Narayanamoorthy, Thirumalai Nallasivan Parthasarathy, Subramaniam Pragathi, Ponnan Shanmugam, Baleanu Dumitru, Ahmadian Ali, Daekook Kang, The novel augmented fermatean MCDM perspectives for identifying the optimal renewable energy power plant location, Sustain. Energy Technol. Assessments 53 (2022), 102488.
[21] Ouafae Jbaihi, Ouchani Fatima-zahra, Ahmed Alami Merrouni, Cherkaoui Mohamed, Abdellatif Ghennioui, Mohammed Maaroufi, An AHP-GIS based site suitability analysis for integrating large-scale hybrid CSP+ PV plants in Morocco: an approach to address the intermittency of solar energy, J. Clean. Prod. 369 (2022), 133250.
[22] Behzad Masoomi, Iman Ghasemian Sahebi, Masood Fathi, Figen Yldrm, Shahryar Ghorbani, Strategic supplier selection for renewable energy supply chain

under green capabilities (fuzzy BWM-WASPAS-COPRAS approach), Energy Strategy Rev. 40 (2022), 100815.
[23] Deepak Paramashivan Kaundinya, P. Balachandra, N.H. Ravindranath, Veilumuthu Ashok, A GIS (geographical information system)-based spatial data mining approach for optimal location and capacity planning of distributed biomass power generation facilities: a case study of tumkur district, India, Energy 52 (2013) 77–88.
[24] Masoud Babaei, Ahmadreza Abazari, Mohammad Mahdi Soleymani, Mohsen Ghafouri, S.M. Muyeen, Mohammad T.H. Beheshti, A data-mining based optimal demand response program for smart home with energy storages and electric vehicles, J. Energy Storage 36 (2021), 102407.
[25] Halil Demolli, Ahmet Sakir Dokuz, Alper Ecemis, Murat Gokcek, Location-based optimal sizing of hybrid renewable energy systems using deterministic and heuristic algorithms, Int. J. Energy Res. 45 (11) (2021) 16155–16175.
[26] E Laxmi Lydia, B Prasanna Kumar, D. Ramya, Generation of dynamic energy management using data mining techniques basing on big data analytics isssues in smart grids, Int. J. Eng. Technol. 7 (26) (2018) 85–89, 2.
[27] Sidique Gawusu, Rhoda Afriyie Mensah, Oisik Das, Exploring distributed energy generation for sustainable development: a data mining approach, J. Energy Storage 48 (2022), 104018.
[28] S. Bahramara, M Parsa Moghaddam, M.R. Haghifam, Optimal planning of hybrid renewable energy systems using HOMER: a review, Renew. Sustain. Energy Rev. 62 (2016) 609–620.
[29] Ahmed S. Menesy, Hamdy M Sultan, Ibrahim O. Habiballah, Hasan Masrur, Kaisar R. Khan, Muhammad Khalid, Optimal configuration of a hybrid photovoltaic/wind turbine/biomass/hydro-pumped storage-based energy system using a heap-based optimization algorithm, Energies 16 (9) (2023) 3648.
[30] Alain C. Biboum, Ahmet Yilanci, Ruben Mouangue, Comparative analysis of hybrid renewable energy systems based on concentrating solar and biomass technologies for faro-poli, Cameroon, Environ. Prog. Sustain. Energy 42 (1) (2023), e13933.
[31] Mojan Maleki Tehrani, Mohammadreza Akhtari, Alibakhsh Kasaeian, Rad Mohammad Amin Vaziri, Ashkan Toopshekan, Maryam Sadeghi Motlagh, Techno-economic investigation of a hybrid biomass renewable energy system to achieve the goals of SDG-17 in deprived areas of Iran, Energy Convers. Manag. 291 (2023), 117319.
[32] J.K. Kazak, M. Mrówczyńska, M. Skiba, A. Bazan-Krzywoszańska, M. Świader, K. Tokarczyk-Dorociak, S. Szewrański, A decision support system for the planning of hybrid renewable energy technologies, in: IOP Conference Series: Earth and Environmental Science, 2021, 12012, 701.
[33] Sabhan Kanata, Syamsyarief Baqaruzi, Muhtar Ali, Purwono Prasetyawan, Winata Toto, Optimal planning of hybrid renewable energy system using homer in Sebesi Island, Indonesia, Int. J. Renew. Energy Resour. 11 (4) (2021) 1507–1516.
[34] Raman Kumar, Harpreet Kaur Channi, A PV-biomass off-grid hybrid renewable energy system (HRES) for rural electrification: design, optimization and techno-economic-environmental analysis, J. Clean. Prod. 349 (2022), 131347.
[35] Paolo Marocco, Domenico Ferrero, Andrea Lanzini, Massimo Santarelli, The role of hydrogen in the optimal design of off-grid hybrid renewable energy systems, J. Energy Storage 46 (2022), 103893.
[36] Grazia, Di Salvatore, Giuseppe Marco Tina, Optimal site selection for floating photovoltaic systems based on geographic information systems (GIS) and multi-criteria decision analysis (MCDA): a case study, Int. J. Sustain. Energy (2023) 1–23.
[37] Pengyu Wei, Olusola Bamisile, Adun Humphrey, Dongsheng Cai, Sandra Obiora, Jian Li, Qi Huang, Bibliographical progress in hybrid renewable energy systems' integration, modelling, optimization, and artificial intelligence applications: a critical review and future research perspective, Energy Sources, Part A Recovery, Util. Environ. Eff. 45 (1) (2023) 2058–2088.
[38] Julian Clifton, Bryan J. Boruff, Assessing the potential for concentrated solar power development in rural Australia, Energy Pol. 38 (9) (2010) 5272–5280.
[39] Jarra Hicks, Nicky Ison, Community-owned renewable energy (CRE): opportunities for rural Australia, Rural Soc. 20 (3) (2011) 244–255.
[40] Roberta Fornarelli, Farhad Shahnia, Anda Martin, Parisa A. Bahri, Goen Ho, Selecting an economically suitable and sustainable solution for a renewable energy-powered water desalination system: a rural Australian case study, Desalination 435 (2018) 128–139.
[41] Mark P. McHenry, Remote area power supply system technologies in western Australia: new developments in 30 Years of slow progress, Renew. Energy 34 (5) (2009) 1348–1353.
[42] Emmanuel K. Yiridoe, others, Social acceptance of wind energy development and planning in rural communities of Australia: a consumer analysis, Energy Pol. 74 (2014) 262–270.
[43] Liam Byrnes, Colin Brown, Liam Wagner, John Foster, Reviewing the viability of renewable energy in community electrification: the case of remote western Australian communities, Renew. Sustain. Energy Rev. 59 (2016) 470–481.