Article

A Comparative Analysis of Data Source's Impact on Renewable Energy Scenario Assessment—The Example of Ground-Mounted Photovoltaics in Germany

Elham Fakharizadehshirazi * and Christine Rösch

Karlsruhe Institute of Technology (KIT), 76344 Karlsruhe, Germany; christine.roesch@kit.edu
* Correspondence: elham.fakharizadehshirazi@kit.edu

Abstract: The German government aims to convert its energy system to renewable energy by 2045. This requires a comprehensive understanding of land eligibility for renewables, particularly land-intensive ground-mounted photovoltaics (GM-PV). Federal states must set aside at least 2% of their land for renewable energy. This target value was derived using a top-down energy demand approach. Georeferenced land use data can be used to make bottom-up estimates. This study investigates how the choice of data source influences the bottom-up evaluation of land eligibility for GM-PV installations in Germany. This study evaluates the quality of data sources and their applicability for GM-PV scenario assessment by comparing the official data source Basis-DLM as the reference with the open-access data sources OpenStreetMap (OSM), Corine Land Cover (CLC), and Copernicus Emergency Management Service (CEMS). The intersection over union (IoU) and Matthews correlation coefficient (MCC) methods were used to analyse the differences in land use and eligibility due to the quality of the data sources and to compare their accuracy. The study's results show the crucial role of data source selection in estimating the potential for GM-PV in Germany. The results indicate that open-access data overestimate land eligibility by 4.0% to 4.5% compared to the official Basis-DLM data. Spatial similarities and discrepancies between the OSM, CEMS CLC, and Basis-DLM land uses were identified. The CLC data exhibit higher consistency with Basis-DLM. These findings emphasise the importance of selecting the appropriate data source depending on the purpose and the use of official data sources for accurate and spatially differentiated decision-making and project planning at different scales. Open-access data sources can be applied for initial orientation and large-scale rough assessment as they balance data accuracy and accessibility.

Keywords: ground-mounted photovoltaics; land use; land eligibility; open-access data; renewable energy; data source; Germany

1. Introduction

The German government aims to achieve climate neutrality by 2045, with a key focus on transitioning to renewable energies [1]. In order to achieve its climate targets, the government has set specific targets for the consumption of renewable energies. The goal is to increase the share of renewable energies in gross electricity consumption from 46% in 2022 to 80% by 2030 [2,3]. Ground-mounted photovoltaics (GM-PV) are a crucial element in Germany's renewable energy strategy but face several challenges, including the need for significant land due to the lower energy density of solar energy. This transition will require relatively a significant amount of land. Based on scenario assessments, the German government assumes that 2% of the national area is needed to achieve the political goals of renewable energy. The states must set aside land for renewable energy, with each federal state having its specific area target [4]. The implementation of GM-PV is influenced by various land use considerations and stakeholder acceptance. It is subject to socio-technical, environmental, and economic constraints and is influenced by competition for land use and public acceptance. The potential for GM-PV varies greatly

between the different studies. Depending on the assumptions made, the results of the studies on the GM-PV potential in Germany significantly differ, e.g., between 220 GW and 520 GW [5–9]. It is difficult to compare the studies and understand the differences because the assumptions and data sources are not always the same or disclosed. The choice of data sources plays a decisive role in the accuracy of the assessment of the GM-PV potential. Most scenarios are built on the demand of the energy transformation, but a few analyse spatial land use scenarios with restriction and suitability criteria [10]. The comparability of studies regarding data sources is largely unknown, even though the data source influences the results of scenario assessment [11]. The importance of data sources in renewable energy assessments has been highlighted in previous research. This finding is supported by Risch, Maier [3], showing the suitability of different land use data and their impact on the assessments and indicating that the results of renewable energy potential studies are inconsistent due to different assumptions, methods, and data sources. Risch, Maier [3] analysed the impact of data sources for wind energy scenarios but not for GM-PV scenarios, which they assessed using the same data they found best for analysing wind energy. Ryberg, Robinius [12] also found inconsistencies in assumptions and data, which impairs a uniform land eligibility assessment for GM-PV and makes it difficult to derive robust recommendations for decision-makers. Ryberg, Caglayan [13] highlight the study from McKenna, Pfenninger [11] as best practice in this context.

McKenna, Pfenninger [11] and Masurowski, Drechsler [14] indicate that, in particular, the land use category settlements (residential and commercial areas) differ among data sources and highly influence the renewable energy potential analysis in Germany. They conclude that the eligible area for GM-PV can be overestimated as not all residential and commercial areas are adequately represented in open-access data sources. The error multiplies in substantial orders of magnitude if buffer zones around a settlement are applied in the scenarios. Data resolution varies across different sources, leading to either underestimation or overestimation of areas. This depends on whether a pixel in coarse resolution datasets is assigned to a specific land use category or represents a different land use type.The study of McKenna, Pfenninger [11] on the impact of data sources with varying periods and spatial resolutions on assessing the German onshore wind potential shows that the impact of data sources is significant but has not been evaluated to date. The scenarios on GM-PV potential, e.g., from Luderer, Günther [15], are developed in accordance with German legislation, in particular, the Renewable Energy Act (EEG), considering specific areas, e.g., side strips of motorways and railways, to be eligible and poor soil quality to be superior for GM-PV [10,16]. The results of the studies are not the same due to different land use data sources applied and restrictions and defined suitability criteria. Some federal states, e.g., North Rhine-Westphalia and Baden-Württemberg, provide the official data source Basis-DLM to assess the GM-PV potential at regional and local scale [17,18].

This study aims to evaluate the impact of different data sources on the assessment of the GM-PV potential in Germany. Given these insights, this research assesses the magnitude of inconsistency when using open-access data, which allows all users to access important information, instead of official data which are difficult to access and must be paid for, and identifies the accuracy of open-access data sources for spatial georeferenced GM-PV scenario assessment based on land use data. The study aims to explain how different data sources affect land eligibility results and to provide valuable knowledge for scientists and decision-makers about the reliability of data sources and their applicability to generate robust and scientifically credible recommendations at different scales and for different purposes. In addition, the study seeks to raise awareness among policymakers and scientists of the opportunities and limitations associated with using open-access data sources in renewable energy decision-making processes. The study examines for the first time the impact of data sources on the assessment of GM-PV

siting in Germany and evaluates four different scenarios using identical criteria but different data sources.

2. Methodology

Open-access data sources such as Corine Land Cover (CLC), OpenStreetMap (OSM), and Copernicus Emergency Management Service (CEMS) are widely used to analyse land eligibility. On the other hand, Basis-DLM stands out as a precise official data source specifically tailored to Germany [19]. Analogous to the study of Risch, Maier [3] on wind energy production comparing different data sources, this study investigates the deviations resulting from different data sources when estimating the GM-PV potential in Germany. To this end, analyses of GM-PV eligibility were carried out for four scenarios, each of which differs in terms of the data sources used: Scenario A (Basis-DLM data), Scenario B (mainly CLC data), Scenario C (mainly OSM data), and Scenario D (all open-access data). In Scenario D, open-access data sources were selected based on comparing CLC, OSM, and CEMS data with the Basis-DLM; the open-access data source that most closely matches the Basis-DLM was selected. The description of which data sources were used for the different land use criteria for the scenarios can be found in Section 3. A more detailed description of the scenarios and the calculated results can be found in the study conducted by Fakharizadehshirazi and Rösch [10]. Scenario A with the Basis-DLM data is the reference to compare the land eligibility results of the other scenarios, B, C, and D, with different shares and sources of open-access data. The methods of intersection over union (IoU) and Matthews correlation coefficient (MCC) were used to compare the scenarios' results and land uses from different sources.

2.1. Data Sources

In this study, the OpenStreetMap (OSM), CORINE Land Cover (CLC), and World Database on Protected Areas (WDPA) data sources were investigated because they are recommended for potential analysis at global and regional scales by McKenna, Pfenninger [11]. The open-access OSM is a crowd-sourced project founded in 2004 [20], and the data are provided by GeoFabrik (Geofabrika, 2023). CLC, a European Commission initiative, provides a land cover inventory with 44 thematic classes updated every 6 years [21]. WDPA, a global directory, is regularly updated and helps to understand how protected areas occur and function [22]. In addition, the European Settlement Map of Copernicus (CEMS) and the Soil Quality Rating (SQR) were included in the study as open-access data sources. CEMS is a spatial raster data source based on the Spot 5&6 satellite data [23]. The SQR from the Federal Institute for Geosciences and Natural Resources is a geo map for assessing soil quality for arable lands [24]. As a reference for the comparison of open-access data sources, the official German data source Basis-DLM [19], characterised by its superior positional accuracy concerning key points and linear features, was used [25]. Derived from 1:25,000 topographic maps, Basis-DLM comprehensively represents various landscape features such as agricultural areas, forests, settlements, and road networks. The accuracy of these features is in the range of ± 3 to ± 15 m, depending on the specific feature. Table 1 gives an overview of the data sources that were used in the comparison analysis. A description of land uses and protected areas data is provided in Supplementary Tables S1 and S2.

Table 1. Overview of data sources applied in the comparison.

Data Sources	Open-Access	Accuracy/Resolution	Data Type	Coverage
Basis-DLM	No	between ± 3 and ± 15 m, depending on the feature	Vector	Germany
CLC	Yes	100 m linear phenomena, 25 ha areal phenomena	Vector	Europe
OSM	Yes	unknown	Vector	Global
CEMS	Yes	10 m	Raster	Global
WDPA	Yes	unknown	Vector	Global
SQR	Yes	250 m	Raster	Germany

2.2. Land Eligibility Analysis for GM-PV

GM-PV land eligibility depends mainly on the Renewable Energy Sources Act [26] and legislation and restrictions to protect the environment and nature. The scenarios used in this study focus on the protection of biodiversity, the preservation of high-quality soils to avoid competition with food production, the preservation of water protection areas, compliance with legal flood plains, and maintaining certain distances from settlements and infrastructure. In determining the areas eligible for GM-PV, restricted areas were excluded, including buffer zones such as roads, railway lines, waterways, forests, protected areas, arable and grassland with high-quality soils, settlement areas, airports, and military areas. For more information regarding the methodological approach, data, and results, see Fakharizadehshirazi and Rösch [10]. Modelling and preprocessing of the data sources used were performed using ArcGIS Desktop 10.8.1, ArcGIS Pro 3.3, and the Python site package arcpy.

This study investigated the influence of data quality and source reliability on eligibility assessments by maintaining consistent restriction and exclusion criteria taken from Fakharizadehshirazi and Rösch [10] and only varying data sources. Scenarios A, B, and C are based on the data sources Basis-DLM, CLC, and OSM, respectively. However, not for all constraints, information was equally available in each data source. For example, the CLC data did not include information on roads and railways. The Basis-DLM data on roads and railroads were included in all three scenarios to ensure consistent comparability. In addition, the CLC data lacked information on protected areas, and OSM only displayed natural reserve data. Although the protected areas were defined and provided in Basis-DLM, WPDA was used to provide information on protected areas in all three scenarios to ensure comparability of results. This study investigated the applicability and accuracy of open-access data to assess land eligibility for GM-PV and applied the GIS eligibility model to Scenario D, using only open-access data. In Scenario D, all restriction layers were sourced from open-access data, which significantly overlapped with official data. To determine the most accurate open-access land use in Scenario D, land uses from open-access sources were compared to Basis- DLM. The analysis focused on land uses of two key freely available data sources, CLC and OSM; however, in the case of settlements, CEMS data were included in the analysis. Compared to the scenario using Basis-DLM data (Scenario A), the results of Scenario D highlight the divergence between the use of open-access and official data.

2.3. Comparison Method

Basis-DLM with precise positional accuracy [3] is an ideal choice for assessing the quality of CLC, OSM, and CEMS-derived land use data. Therefore, Basis-DLM was used as a reference data source to assess the alignment between land use from different data sources and compare the land eligibility analysis outcome results across four scenarios. Intersection over union (IoU) and Matthews correlation coefficient (MCC) were used to make the comparison. The IoU (Equation (1), [27]) is a metric widely used in object recognition and image classification applications. It measures the overlap between two data sources compared with values between 0 and 1. An IoU of 1 describes the specified areas of two data sources as equal. Matthews [28] initially developed the MCC or phi coefficient to

compare chemical structures. It is used in machine learning to assess the accuracy of binary (two-class) classifications [29]. Chicco and Jurman [29] argue that the MCC should be adopted as the preferred measure for evaluating binary classification tasks throughout the scientific community.

$$IoU = \frac{\text{Area of Overlap}}{\text{Area of Union}} \tag{1}$$

3. Results

Table 2 shows the IoU scores. The forest category shows consistently high IoU scores for both the CLC and OSM data sources, indicating a strong match with the Basis-DLM reference data. Specifically, forest achieves IoU scores of 0.85 with CLC and 0.88 with OSM data sources, confirming the reliability of both data sources for representing forest land use. Arable land shows a higher IoU score when comparing CLC and OSM data sources, suggesting that CLC may offer superior accuracy in delineating this land use type, with scores of 0.75 and 0.66, respectively. Conversely, grassland has lower IoU scores in both the CLC and OSM data sources, indicating potential difficulties in accurately identifying and categorizing grassland areas, with scores of 0.53 and 0.43, respectively. Settlements have relatively consistent IoU scores of 0.54 with both CLC and OSM data sources, indicating moderate agreement with the Basis-DLM reference data. However, there is a lower IoU score of 0.42 with the CEMS dataset, indicating less agreement between the CEMS data and the Basis-DLM.

Table 2. IoU comparison of land use classes with Basis-DLM data.

Land Uses	CLC	OSM	CEMS
Forest	0.85	0.88	-
Arable land	0.75	0.66	-
Grassland	0.53	0.43	-
Settlements	0.54	0.54	0.42
Airports	0.67	0.69	-

Based on MCC analysis in Table 3, the highest agreement was found for the classification of forests (CLC: 0.89, OSM: 0.91), followed by arable land (CLC: 0.78, OSM: 0.70), indicating strong compatibility between these land use classes and the Basis-DLM data. In contrast, grassland showed the lowest agreement (CLC: 0.62, OSM: 0.55), indicating a lower consistency of classification between the different data sources. Furthermore, settlements showed a moderate agreement with CLC (MCC = 0.68) and OSM data (MCC = 0.68), but a lower agreement with CEMS data (MCC = 0.55).

Table 3. Matthews (phi) Coefficients comparing land use classes with Basis-DLM data.

Land Uses	CLC	OSM	CEMS
Forest	0.89	0.91	-
Arable land	0.78	0.70	-
Grassland	0.62	0.55	-
Settlements	0.68	0.68	0.55
Airports	0.73	0.71	-

In terms of IoU and CMM results, the CLC data sources have a higher overall overlap with the Basis-DLM.

Table 4 provides a detailed overview of the exclusion criteria applicable to the different scenarios and the corresponding data sources used. In Scenario A, the eligibility analysis is based on official data (Basis-DLM), while Scenarios B and C primarily use CLC and OSM, respectively. CEMS was excluded from the eligibility analysis due to its insufficient compliance with Basis-DLM. The results presented in Tables 2 and 3 played a key role in

the data selection decision process for the implementation of Scenario D, known as the open-access data scenario, which aims to analyse the accuracy of using open-access data in the absence of commercial data sources. Based on these results, the CLC data sources were found to have a higher degree of coherence with the official data (Basis-DLM) compared to other available sources, leading to their selection for Scenario D. It is important to note that the lack of road data within this data source limits the use of the CLC data. Complementary road data from OSM was included to overcome this limitation while remaining consistent with the open data access principles of Scenario D. In addition, information on protected areas has been sourced from the WPD in conjunction with Scenario D.

Table 4. Restriction criteria for analysing GM-PV land in different scenarios.

Restriction Layers	Scenario A	Scenario B	Scenario C	Scenario D
High-quality arable lands: SQR $\geq$ 40	Basis-DLM	CLC	OSM	CLC
Forests	Basis-DLM	CLC	OSM	CLC
Protected area	WDPA	WDPA	WDPA	WDPA
Landscape area	WDPA	WDPA	WDPA	WDPA
Settlement areas	Basis-DLM	CLC	OSM	CLC
Waterways and waterbody	Basis-DLM	CLC	OSM	OSM
Transport routes: Roads	Basis-DLM	Basis-DLM	Basis-DLM	OSM
Transport routes: Railway lines	Basis-DLM	Basis-DLM	Basis-DLM	OSM
Airports	Basis-DLM	CLC	OSM	CLC
Military zones	Basis-DLM	CLC	OSM	CLC

By applying the eligibility model across the four scenarios, areas eligible for GM-PV were identified. Table 5 shows the results of land eligibility for the different scenarios in Germany. Scenario A shows a significant contrast in the proportion of non-eligible land compared to the other scenarios. Specifically, 91.1% of Germany's total area belongs to this category, while 8.9% remains unrestrictedly accessible, significantly differing from the other scenarios. This difference in Scenario A distinguishes it from Scenarios B, C, and D, where the percentage of non-eligible areas with restrictions ranges from 86.6% to 87.1%.

Table 5. Eligibility analysis results as a percentage of total Germany (%).

Scenarios	Not Eligible	Eligible
Scenario A	91.1	8.9
Scenario B	86.6	13.4
Scenario C	86.8	13.2
Scenario D	87.1	12.9

To illustrate the spatial distribution of the four scenarios, Figure 1a–d was created, showing the eligible areas' locations. The spatial distribution of areas eligible for GM-PV shows a consistent pattern in all four scenarios, with no significant deviations. A significant share of the eligible areas is located in the north-eastern and southern regions of Germany. In scenarios with freely accessible data (Figure 1b–d), the yellow colour of the eligible areas is more pronounced compared to the official data (Figure 1a).
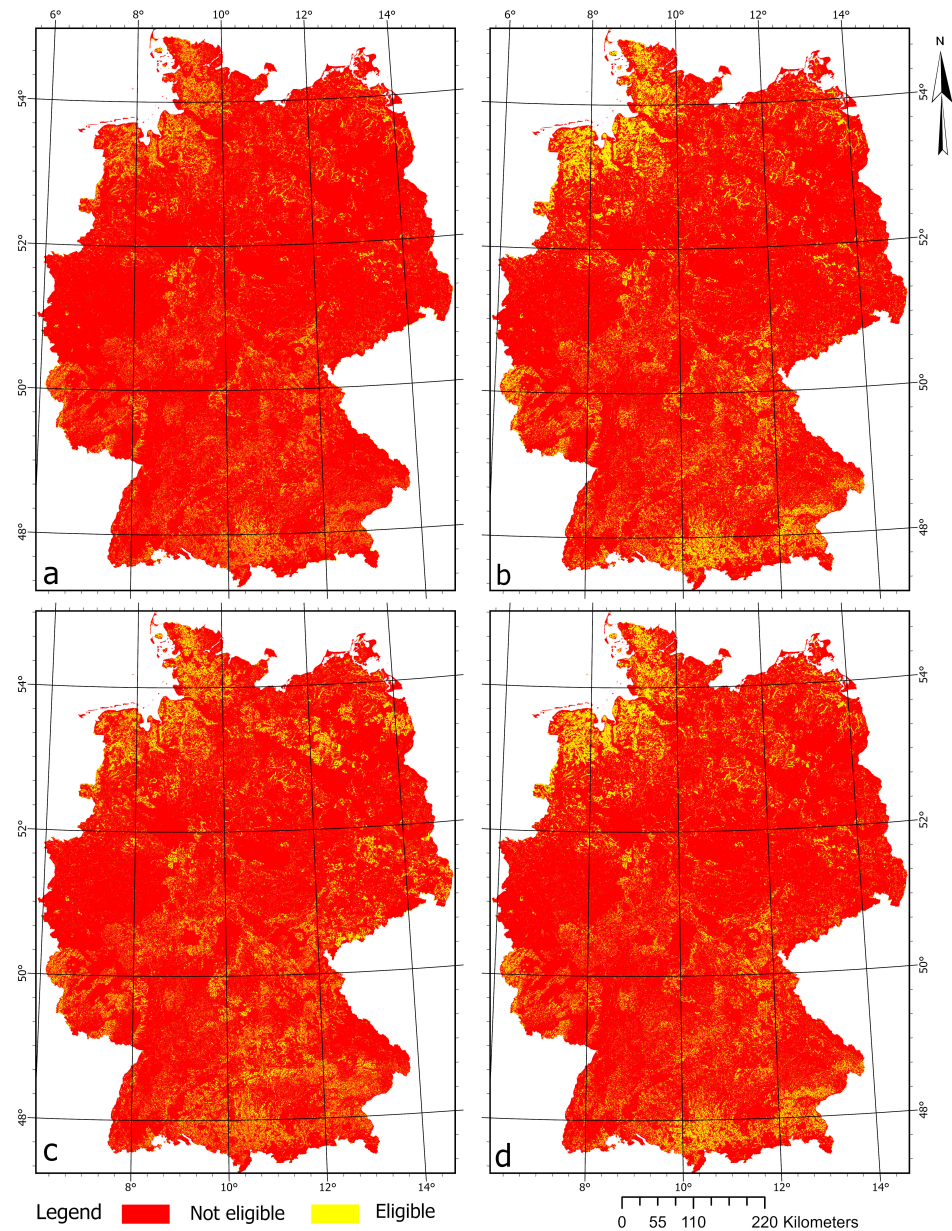
Figure 1. Geographical coverage for GM-PV eligibility in Germany: (**a**): Scenario A (Basis–DLM), (**b**): Scenario B (CLC), (**c**): Scenario C (OSM) and (**d**): Scenario D (open-access data).

The four scenarios' results were compared in pairs using IoU and MCC calculations. The results of these comparisons were then combined to create an IoU and MCC matrix, shown in Table 6.

Table 6. Comparing land eligibility scenario results (A, B, C, and D) using IoU and MCC methods.

Scenarios	Methods	Scenario A	Scenario B	Scenario C	Scenario D
Scenario A	IoU	1.00	0.60	0.53	0.53
	MCC		0.74	0.68	0.68
Scenario B	IoU		1.00	0.52	0.81
	MCC			0.64	0.89
Scenario C	IoU			1.00	0.47
	MCC				0.59
Scenario D	IoU				
	MCC				

The calculation of the IoU value between the scenarios (Scenario A, B, C, and D) shows that Scenario B (CLC), with the highest IoU [MCC] value of 0.60 [0.74], is the most similar to Scenario A (Basis-DLM), which indicates a considerable degree of spatial correspondence between the two scenarios and suggests a substantial degree of harmonisation in their representations of land eligibility. Scenario C (OSM) follows closely with an IoU [MCC] of 0.53 [0.68], showing a moderate degree of spatial correspondence with Scenario A. In contrast, Scenario D (open-access) has a similar IoU [MCC] value of 0.53 [0.68] to both Scenario A and Scenario C, indicating a moderate but comparable spatial correspondence. Among the pairs, Scenario B and Scenario D have the highest IoU [MCC] (0.81 [0.89]), indicating that these two scenarios are most similar in terms of space. This comparison provides insights into the varying degrees of spatial concordance between different scenarios when capturing land eligibility patterns.

4. Discussion

In this study, three data sources, Basis-DLM, OSM, and CLC, in addition to CEMS settlement data, the SQR soil quality rating map, and WDPA-protected areas, were used to investigate the influence of the choice of data source for GM-PV scenario assessment. McKenna, Pfenninger [11] state that OSM, CLC, and WDPA are mainly used at global and regional scales to analyse the eligibility of land for wind energy. Risch, Maier [3] discussed the influence of the land use category settlements on potential analyses in Germany. He emphasised that settlements are often used synonymously with residential land uses in land eligibility analyses. McKenna, Pfenninger [11], and Masurowski, Drechsler [14] (cited in Risch, Maier [3]) support the finding that settlements have a significant impact on the land eligibility analysis. For this reason, in addition to the Basis-DLM, OSM, and CLC data, the CESM settlement data were analysed. CESM is based on satellite data and has the lowest level of accuracy in comparison to the Basis-DLM data, which is why it is not used in this study for the analysis of land eligibility in Scenario D. The variance in land eligibility analysis between the use of official data and open-access data ranges from 4.0 to 4.5% of Germany's total area. This corresponds to 1.4 to 1.6 million hectares (Mha), a significant area compared to Germany's arable land, which totals approximately 11.66 Mha.

Numerous studies have examined the siting of renewable energy sources at the international, national, and regional levels. As part of a systematic review, Spyridonidou and Vagiona [30] carefully analysed and evaluated existing methods for selecting sites suitable for both solar photovoltaic and concentrated solar power. Of the 10,121 studies reviewed, 152 met the eligibility criteria and underwent comprehensive evaluation, including site suitability analysis. Despite the high amount of international research on land suitability for renewable energy, the selection of data sources for such analyses has received relatively little attention.

Risch, Maier [3] examined the influence of the data on the potential for wind energy, which can be considered the most relevant study to the current research. The present study analyses the impact of data sources on GM-PV in more detail and with more data as a continuation of Risch, Maier [3]'s study. Risch, Maier [3], investigating the impact of the choice of data type in wind potential analysis, concluded that there was a significant overlap between OSM and Basis-DLM data in most land use categories. The study of Risch, Maier [3] notes that significant biases exist in the land use data sources commonly applied by the energy systems community. They conclude that the use of CLC leads to a significant overestimation of the potentially usable area for renewable energy technologies compared to Basis-DLM and OSM, respectively, and is, therefore, not recommended for renewable energy potential analyses. They state that high-quality data sources are needed for reliable input for policymakers or energy system models. Risch, Maier [3] found that in the case of Germany, Basis-DLM and OSM provide similar information for several categories, especially line-like features such as power lines or railways. For others, Basis-DLM and OSM show significant differences, e.g., for industry/commercial areas, lakes, or rivers.

This study evaluated the IoU and the MCC for Scenarios B, C, and D compared to Scenario A (Basis-DLM). The IoU values were between 0.53 and 0.6, while the MCC ranged from 0.7 to 0.68. These metrics were used to evaluate the results of the land eligibility analysis, highlighting differences between the use of official and open-access data sources. The IoU [MCC] resulting from the comparison between Scenarios B and C was approximately 0.5 [0.6], indicating a moderate difference in the impact of using CLC and OSM to analyse land eligibility in Germany. These findings could be applied across Europe. Nevertheless, Risch, Maier [3] results suggest that CLC underestimates smaller settlements, which leads to an overestimation of the potential area, with the total potential area being about 80% higher.

In the study of Risch, Maier [3] comparing the main land uses (arable, forest, and grassland) from CLC and OSM with Basis-DLM, the largest overlap occurred with forests, while the least overlap was observed with grassland, which is consistent with the current study's results. Although OSM boasts infrastructure data, including roads and railways, unlike CLC, it is essential to note that OSM relies on contributions from volunteer users, raising concerns about data completeness [31]. Previous OSM assessments have shown that road reliability in many Western European countries is over 95%. Still, the completeness of building data is considerably lower, for example, in Saxony, Germany, where it was only 23% in 2013 (Hecht, Kunze [32] cited in McKenna, Ryberg [31]). In the current research, the IoU for comparing settlement data between OSM and Basis-DLM is estimated to be 0.55. Consequently, using OSM data depends on including supplementary data to assess renewable energy potential accurately [31]. On the other hand, CLC data, which cover the whole of Europe, overlap more with the Basis-DLM data. The high consistency between the IoU and MCC comparison results suggests a high degree of agreement between these assessment methods. This alignment underscores the reliability and robustness of the results and strengthens confidence in the assessments' accuracy.

5. Conclusions

This study provides valuable insights into the influence of data source selection on the assessment of the potential for ground-mounted PV in Germany. The influence of data selection on potential analyses was demonstrated by systematically comparing commonly used open-access data sources with the official Basis-DLM data. The results of this study highlight the critical importance of carefully selecting data sources for accurate land-based renewable energy planning. The variance in GM-PV eligible areas calculated from official data versus open-access data ranges from 4% to 4.5% of Germany's total land area. In particular, the use of open-access data tends to overestimate the eligible area. This discrepancy mainly affects arable land, which has a direct impact on food competition and stakeholder acceptance. To address these issues, policy makers and planners should prioritise the use of official Basis-DLM data for accurate assessments of GM-PV potential. When official data are not available, the combination of open-access datasets such as CLC and OSM can provide a practical alternative for preliminary assessments, offering a balance between accuracy and accessibility. Researchers are encouraged to further investigate the impact of different data sources on land suitability, as refining methodologies and improving data accuracy are critical for reliable renewable energy planning. In addition, stakeholders and community advocates should understand how data discrepancies affect land use and the acceptance of renewable energy projects. This understanding can help manage potential conflicts and foster broader support for renewable energy initiatives. Among the openly available data, satellite-based CEMS data show limited consistency with Basis-DLM. Conversely, CLC data, which provide comprehensive coverage of European land use, show greater consistency with Basis-DLM, although they do not include road and railroad data. On the other hand, OSM includes road and railroad data, which increases its usefulness for transport infrastructure efforts. In the absence of official data, a viable approach for land suitability analysis is to combine CLC and OSM data, as explored in Scenario D, resulting in moderate consistency compared to Basis-DLM data. Thus, OSM,

CLC, and WDPA provide options for rough large-scale planning that balance accuracy and accessibility. However, for studies where accuracy is a priority, the use of official Basis-DLM data is recommended. Insights into the quality and applicability of these data sources provide valuable guidance to decision-makers on potential uncertainties in scenario analysis and help scientists choose data sources wisely, tailored to their specific needs. Given the limited research on the impact of data sources in determining land suitability for renewable energy analysis results, continued research is essential to fill knowledge gaps and improve the accuracy of land suitability analyses, ultimately leading to more informed decisions and effective implementation of renewable energy projects.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/en17153766/s1, Table S1: Description of land use data from different data sources.; Table S2: Description of protection areas data.

Author Contributions: Conceptualization, E.F. and C.R.; methodology, E.F. and C.R.; software, E.F.; investigation, E.F. and C.R.; data curation, E.F.; writing—original draft preparation, E.F. and C.R.; writing—review and editing, E.F. and C.R.; visualization, E.F.; supervision, E.F. and C.R. All authors have read and agreed to the published version of the manuscript.

Data Availability Statement: The original contributions presented in the study are included in the article, further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Bundesregierung. Gesetz für den Ausbau Erneuerbarer Energien (Erneuerbare-Energien-Gesetz-EEG 2021). 2021. Available online: https://dserver.bundestag.de/btd/19/302/1930230.pdf (accessed on 19 July 2024).
2. Kullmann, F. Energieperspektiven 2030. 2023. Available online: https://www.fz-juelich.de/de/iek/iek-3/aktuelles/meldungen/energieperspektiven-2030/handout_energieperspektiven-2030.pdf (accessed on 19 July 2024).
3. Risch, S.; Maier, R.; Du, J.; Pflugradt, N.; Stenzel, P.; Kotzur, L.; Stolten, D. Potentials of renewable energy sources in germany and the influence of land use datasets. *Energies* **2022**, *15*, 5536. [CrossRef]
4. May, J. Raumplanung und erneuerbare Energien. Renews Kompakt. *Agent. Erneuerbare Energ.* **2023**, *58*, 1–15.
5. Wirth, H. Aktuelle Fakten zur Photovoltaik in Deutschland. 2024. Available online: https://www.ise.fraunhofer.de/content/dam/ise/de/documents/publications/studies/aktuelle-fakten-zur-photovoltaik-in-deutschland.pdf (accessed on 19 July 2024).
6. *Energieszenarien für ein Energiekonzept der Bundesregierung: Projekt Nr. 12/10 des Bundesministeriums für Wirtschaft und Technologie*; Studie 2012; Prognos AG: Basel, Switzerland; Energiewirtschaftliches Institut an der Universität zu Köln (EWI): Cologne, Germany; Gesellschaft für Wirtschaftliche Strukturforschung (GWS): Osnabrück, Germany, 2010.
7. Peters, D.; Schicketanz, S.; Hanusch, D.; Rohr, A.; Kothe, M.; Kinast, P. Räumlich differenzierte Flächenpotentiale für erneuerbare Energien in Deutschland. Bundesministerium für Verkehr und digitale Infrastruktur (BMVI). Berlin (08). Online verfügbar unter. Available online: https://www.bbr.bund.de/BBSR/DE/Veroeffentlichungen/ministerien/BMVI/BMVIOnline/2015/DL_BMVI_Online_08_15.pdf (accessed on 19 July 2024).
8. Bründlinger, T.; König, J.E.; Frank, O.; Gründig, D.; Jugel, C.; Kraft, P.; Krieger, O.; Mischinger, S.; Prein, P.; Seidl, H. *Dena-Leitstudie Integrierte Energiewende: Impulse für die Gestaltung des Energiesystems bis 2050*; Deutsche Energie-Agentur GmbH (dena), ewi Energy Research & Scenarios gGmbH: Berlin/Cologne, Germany, 2018.
9. BDI. Bundesverband der Deutschen Industrie (BDI) durchgeführt von The Boston Consulting Group (BCG) und Prognos. 2021. Available online: https://www.vci.de/vci/downloads-vci/media-weitere-downloads/2018-01-18-bdi-studie-klimapfade-fuer-deutschland.pdf (accessed on 19 July 2024).
10. Fakharizadehshirazi, E.; Rösch, C. A novel socio-techno-environmental GIS approach to assess the contribution of ground-mounted photovoltaics to achieve climate neutrality in Germany. *Renew. Energy* **2024**, *227*, 120117. [CrossRef]
11. McKenna, R.; Pfenninger, S.; Heinrichs, H.; Schmidt, J.; Staffell, I.; Bauer, C.; Gruber, K.; Hahmann, A.N.; Jansen, M.; Klingler, M.; et al. High-resolution large-scale onshore wind energy assessments: A review of potential definitions, methodologies and future research needs. *Renew. Energy* **2022**, *182*, 659–684. [CrossRef]
12. Ryberg, D.S.; Robinius, M.; Stolten, D. Evaluating land eligibility constraints of renewable energy sources in Europe. *Energies* **2018**, *11*, 1246. [CrossRef]

13.   Ryberg, D.S.; Caglayan, D.G.; Schmitt, S.; Linßen, J.; Stolten, D.; Robinius, M. The future of European onshore wind energy potential: Detailed distribution and simulation of advanced turbine designs. *Energy* **2019**, *182*, 1222–1238. [CrossRef]
14.   Masurowski, F.; Drechsler, M.; Frank, K. A spatially explicit assessment of the wind energy potential in response to an increased distance between wind turbines and settlements in Germany. *Energy Policy* **2016**, *97*, 343–350. [CrossRef]
15.   Luderer, G.; Günther, C.; Sörgel, D.; Kost, C.; Benke, F.; Auer, C.; Koller, F.; Herbst, A.; Reder, K.; Böttger, D. *Deutschland auf dem Weg zur Klimaneutralität 2045-Szenarien und Pfade im Modellvergleich (Zusammenfassung)*; Potsdam Institute for Climate Impact Research: Potsdam, Germany, 2021. [CrossRef]
16.   Amme, J.; Kötter, E.; Janiak, F.; Lancien, B. *Der Photovoltaik-und Windflächenrechner—Methoden und Daten*; Agora Energiewende: Berlin, Germany, 2021.
17.   LUBW. Energieatlas Baden-Württemberg. 2024. Available online: https://www.energieatlas-bw.de/sonne/freiflachen (accessed on 19 July 2024).
18.   NRW. Energy4Climate. Photovoltaik auf Freiflächen. Leitfaden. 2023. Available online: https://www.energy4climate.nrw/energiewirtschaft/photovoltaik/freiflaechen-pv (accessed on 19 July 2024).
19.   Basis-DLM. Digitales Basis-Landschaftsmodell (Ebenen) (Basis-DLM). 2022. Available online: https://gdz.bkg.bund.de/index.php/default/digitale-geodaten/digitale-landschaftsmodelle.html (accessed on 19 July 2024).
20.   Fonte, C.C.; Minghini, M.; Patriarca, J.; Antoniou, V.; See, L.; Skopeliti, A. Generating up-to-date and detailed land use and land cover maps using OpenStreetMap and GlobeLand30. *ISPRS Int. J. Geo-Inf.* **2017**, *6*, 125. [CrossRef]
21.   CLC. CORINE Land Cover. Available online: https://land.copernicus.eu/en/products/corine-land-cover (accessed on 19 July 2024).
22.   WDPA. The World Database on Protected Areas. 2023. Available online: https://www.protectedplanet.net/en/thematic-areas/wdpa?tab=WDPA (accessed on 19 July 2024).
23.   Pesaresi, M.; Ehrlich, D.; Ferri, S.; Florczyk, A.; Freire, S.; Halkia, M.; Julea, A.; Kemper, T.; Soille, P.; Syrris, V. *Operating Procedure for the Production of the Global Human Settlement Layer from Landsat Data of the Epochs 1975, 1990, 2000, and 2014*; Publications Office of the European Union: Luxembourg, 2016; pp. 1–62.
24.   BGR. Ackerbauliches Ertragspotential der Böden in Deutschland. Available online: https://www.bgr.bund.de/DE/Themen/Boden/Ressourcenbewertung/Ertragspotential/Ertragspotential_node.html (accessed on 19 July 2024).
25.   BKG. *Digitales Basis-Landschaftsmodell Basis-DLM*; Bundesamt für Kartographie und Geodäsie: Frankfurt am Main, Germany, 2022.
26.   EEG. *Gesetz für den Ausbau Erneuerbarer Energien (ErneuerbareEnergien-Gesetz-EEG 2023): Erneuerbare-Energien-Gesetz vom 21. Juli 2014 das Zuletzt durch Artikel 6 des Gesetzes vom 4. Januar 2023 Geändert Worden ist*; EEG: Sottrum, Germany, 2023.
27.   Matthews, B.W. Comparison of the predicted and observed secondary structure of T4 phage lysozyme. *Biochim. Biophy. Acta (BBA)-Protein Struct.* **1975**, *405*, 442–451. [CrossRef]
28.   Chicco, D.; Jurman, G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genom.* **2020**, *21*, 6. [CrossRef] [PubMed]
29.   Pyimagesearch. Intersection over Union (IoU) for Object Detection. Available online: https://pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/ (accessed on 23 July 2024).
30.   Spyridonidou, S.; Vagiona, D.G. A systematic review of site-selection procedures of PV and CSP technologies. *Energy Rep.* **2023**, *9*, 2947–2979. [CrossRef]
31.   McKenna, R.; Ryberg, D.S.; Staffell, I.; Hahmann, A.N.; Schmidt, J.; Heinrichs, H.; Höltinger, S.; Lilliestam, J.; Pfenninger, S.; Robinius, M.; et al. On the socio-technical potential for onshore wind in Europe: A response to Enevoldsen et al. (2019), Energy Policy, 132, 1092-1100. *Energy Policy* **2020**, *145*, 111693. [CrossRef]
32.   Hecht, R.; Kunze, C.; Hahmann, S. Measuring Completeness of Building Footprints in OpenStreetMap over Space and Time. *ISPRS Int. J. Geo-Inf.* **2013**, *2*, 1066–1091. [CrossRef]