remote sensing

Review

# Spatial Models of Solar and Terrestrial Radiation Budgets and Machine Learning: A Review †

**Julián Guillermo García Pedreros [1,*], Susana Lagüela López [2] and Manuel Rodríguez Martín [3]**

[1] Doctorate in Geotechnologies Applied to Construction, Energy and Industry, Higher Polytechnic School of Avila, University of Salamanca, C/Hornos Caleros 50, 05003 Ávila, Spain
[2] Department of Cartographic and Land Engineering, Higher Polytechnic School of Avila, University of Salamanca, C/Hornos Caleros 50, 05003 Ávila, Spain; sulaguela@usal.es
[3] Department of Mechanical Engineering, Higher Polytechnic School of Zamora, University of Salamanca, Av. Cardenal Cisneros 34, 49029 Zamora, Spain; ingmanuel@usal.es
* Correspondence: julian.garped@usal.es; Tel.: +34-603228980

† To Daniel Santiago García Skinner "Siempre estarás con y entre nosotros Santi".

**Abstract:** Currently, spatial modeling is of particular relevance as it enables the understanding of the patterns and spatial variability of an event, the monitoring and prediction of the spatial behavior of a variable, the optimization of resources, and the evaluation of the impacts of a phenomenon of interest. Research carried out recently on variables related to solar energy budgets has been of special relevance due to its applications and developments in machine learning (ML) and deep learning (DL). These algorithms are crucial to improve the efficiency, precision, and applicability of remote sensing, allowing greater decision making with more reliable and timely data. Thus, this work proposes a systematic and rigorous methodology for searching research articles about the latest advances and contributions related to the modeling of radiative energy budgets using novel techniques and algorithms in some of the most relevant international scientific databases (Scopus, ScienceDirect, ResearchGate). Search parameters were applied using tracking methods through various filters, specific classifiers, and Boolean operators. The results allowed for an analysis of search trends and citations in the last 5 years related to the topic of interest and the number of most relevant articles for this research, analyzed through specialized metrics and graphs. Additionally, this methodology was classified into four categories of importance for refined and articulated searches in this evaluation: (i) according to the applied interpolation methods, (ii) according to the remote sensors used, (iii) according to the applications in different fields of knowledge. As a relevant fact and with an essentially prospective purpose, a subchapter of this review was dedicated to the latest advances and developments applied to (iv) spatial modeling of terrestrial radiation through ML, this method being a tool that opens multiple alternatives for data processing and analysis in the development and achievement of objectives in the field of geotechnologies. A quantitative comparison was conducted on the predictive performance results between the classification/regression algorithms found in the studies explored in this review. The evaluation confirmed the existence of a persistent shortage of studies in recent years within the geotechnologies field, particularly concerning the comparison of spatial distribution modeling techniques developed and implemented through ML for incident solar and terrestrial radiation. Therefore, this work provides a synthesis and analysis of the most used and novel techniques in the modeling of solar energy budgets, their limitations, and biggest challenges.

**Keywords:** remote sensing; artificial intelligence; machine learning; spatial modeling; interpolation; terrestrial radiation; solar radiation

check for updates

**Citation:** García Pedreros, J.G.; Lagüela López, S.; Rodríguez Martín, M. Spatial Models of Solar and Terrestrial Radiation Budgets and Machine Learning: A Review. *Remote Sens.* **2024**, *16*, 2883. https://doi.org/10.3390/rs16162883

Academic Editor: Stephan Havemann

Received: 25 April 2024
Revised: 15 July 2024
Accepted: 5 August 2024
Published: 7 August 2024



## 1. Introduction

Solar radiation is the energy emitted by the sun, which spreads in all directions of space through electromagnetic waves. The solar energy that reaches the limit of the earth's atmosphere is called extraterrestrial radiation and consists of incident shortwave solar radiation (wavelengths between 0.3 and 2.8 μm), which enters the earth's atmosphere and is energized through flows and complex processes, such as absorption, dispersion, and reflection, among other components. For its part, terrestrial radiation is the energy emitted by the earth's surface in longwave form (wavelengths between 3.5 and 50 μm), which also involves the aforementioned processes in different layers of the atmosphere. The sum of incident shortwave solar radiation and longwave terrestrial radiation is understood as total radiation [1]. Currently, a large number of disciplines focus their interests on understanding these energy budgets due to their importance in physical, biological, and climatic systems; corresponding to these systems, the works presented in references [2–4] address and underscore the significance of the theoretical knowledge of mass and energy budgets, the efficient utilization of bioenergy in agricultural fields, and the role of turbulent heat fluxes in surface energy balances. Therefore, fields such as geomatics and geotechnologies have a superlative role due to their contributions to the increasingly accurate spatial modeling of these complex variables in the earth's system.

A model is defined as a partial representation of reality that reflects some of its properties due to the need to reduce the complexity of the real processes and objects [5]. Spatial models are a set of tools that apply operations (known as "geoprocesses") to their data to create new results and are based on the study of patterns or spatial structures of some phenomenon of interest and its variations over time. Spatial modeling has been consistently implemented in various fields of knowledge with the aim of representing reality through spatial data [6]. Regarding spatial distribution models that have been developed or applied in recent years, studies such as the one carried out by [7] on the spatial variation of air gas concentrations ($H_2S$ and $NH_3$) in Costa Rica; the spatial distribution of weeds under climate change scenarios by [8]; the application of artificial intelligence models to estimate the distribution and risks of heavy metal contamination in agricultural soils conducted by [9]; and much closer works applied to solar radiation, among them the one developed by [10] in which they evaluated two models of spatial interpolation of monthly global solar radiation in Tunisia and another carried out by [11] in which they modeled spatiotemporal interpolations of solar irradiance in the USA.

These approaches are a synthesis of rigorous bibliographic exploration; the evidence, which is interesting and of great contribution to various disciplines, has also allowed for the specification of spatial models developed and applied through ML to incident solar and terrestrial radiation (total radiation) at a regional level. The work of [12] should be highlighted, as the authors generated an ML model for solar irradiance forecasting in Seattle and Medford, USA. Of equal importance is the research of [13], in which a time-continuous land surface temperature (LST) data fusion based on DL with microwave remote sensing and ground truth observations was developed in China. Also important is the work of [14], which compared different ML methods to reconstruct daily evapotranspiration as estimated by thermal-infrared remote sensing in the Heihe River Basin in northwest China. The research by [15] developed a high-resolution spatiotemporal assessment of solar potential from remote sensing data using DL in Maribor city, Slovenia.

However, there is currently no evidence of works conducted in Spain that collect, classify, and compare the spatial modeling of this specific variable. Therefore, this review is relevant and focuses on identifying a broad spectrum of concepts, such as machine learning, maximum entropy, geostatistics, artificial intelligence, geoprocessing, spatial interpolation, capture algorithms, parameterization, and spatial distribution. In addition, we identified a considerable number of variables related to different types of techniques; this study focuses on and delves into these concepts and techniques.

This research focuses on the analysis of the main and most recent studies on spatial modeling of energy budgets in the earth's system and emphasizes the importance of various disciplines on the influence of radiative balance dynamics for informed decision making in critical sectors, especially given the current climate crisis. The contributions derived from this article have a wide range of possibilities, such as understanding spatial modeling techniques (especially those based on ML) regarding incident and terrestrial radiation and their use as analytical decision-making tools in the fields of agriculture [16], hydrology [17], biology [18], meteorology [19], epidemiology [20], climatology [21], energy [22], and economics [23], among many other disciplines.

## 2. Materials and Methods

Emphasis was placed on a systematic and rigorous methodology for searching research articles and literature reviews in the most impactful and relevant scientific and academic databases internationally, such as Scopus, ScienceDirect, and ResearchGate. Search parameters were applied by tracking various filters, specific classifiers, and Boolean operators (Figure 1); the subsequent classification and management of bibliographic resources were handled by the reference management software Mendeley (Version 2.120.0) [24].

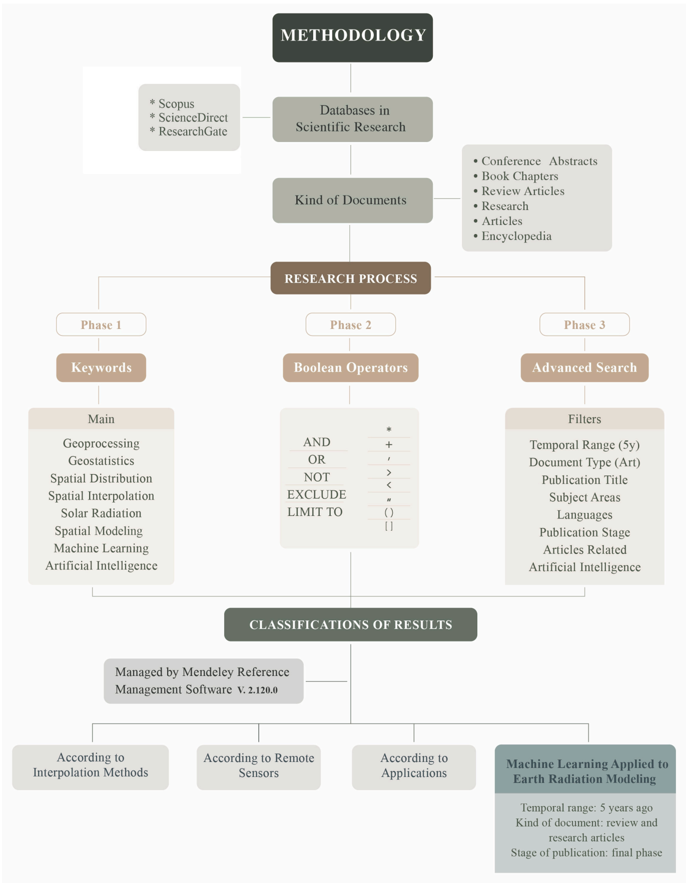


**Figure 1.** Diagram of stages and processes used in the methodology of this review article.

Initially, access to the aforementioned scientific databases was gained, and basic searches were conducted using the keywords "machine learning", "spatial modeling", "radiation", and "interpolation". Once these results were obtained and a high quantity and variety of related documents were found, the search iterations phase began, in which the entry of these keywords was combined with the integration of Boolean operators, that is, words and symbols that allow for expanding or narrowing the search parameters when using a database or search engine, according to each consulted database. Some examples of Boolean operators used were: AND, NOT, LIMIT TO.

The obtained results were considerably fewer. To further refine them and obtain information with greater relevance and utility, new searches were initiated by combining keywords, more complex Boolean operators for limitation or exclusion, and advanced search parameters (Figure 1), and the results were classified according to the following conditions:

1. The temporal range of related publications was unlimited for the first three classifications (Figure 1) and set to no more than five years old for the fourth classification "Machine Learning applied to spatial models of terrestrial radiation" (since 2018).
2. The document type focused only on scientific articles and reviews in the final phase of the search.
3. The title of the related publication should contain at least two of the exposed keywords (Figure 1).
4. The areas or fields of the publications should be highly related to the fields of geotechnologies, geomatics, planetary sciences, Earth observation, or geographic information systems.
5. The language was limited to English results; however, it was later modified to 80% of publications in English and 20% in other languages to avoid biasing possible relevant information.
6. The publication stage was configured to consider only published articles and discard pre-published articles during the review process. This implies that they would have already been through previous processes of review, corrections, or validation of their results, and consequently they would be considered works of higher quality.

Once these conditions and search parameters were applied to the databases, the selection and classification of articles with the highest number of citations and impact in related areas was conducted, and subsequently, the compilation and storage phase of the bibliographic references in the Mendeley reference manager was initiated.

It is important to explain that the application of this methodology allowed us to find a large number of articles and information that could significantly contribute as a reference in this review and its subsequent analysis. However, after evaluating these documents, we determined that it was convenient to group the results into large categories since many of the works were developed based on the application of spatial interpolation techniques (a highly developed field in ML algorithms due to its enormous functionalities); many others showed their applicability to various fields of knowledge; and another group, for its part, emphasized that remote sensors were used for different study interests and various research scenarios. Furthermore, in recent years, ML has had a high level of growth and development in fields related to earth sciences and remote sensing. Thus, it was appropriate to create a fourth category and subchapter dedicated exclusively to these advances and contributions as well as to their limitations and challenges.

This methodological approach allowed us to establish a state-of-the-art spatial model of solar and terrestrial radiation budgets developed through novel techniques and algorithms (especially those based on ML) in recent years. The details of the current status of this bibliographic compilation are presented later and demonstrate how the commonly used geostatistical methods traditionally applied to modeling the spatial distribution of variables associated with solar energy are becoming less and less important. Therefore, algorithms based on ML and DL seem to take greater prominence in the application and development of more complex and precise models of this type today. The answers to these hypotheses are found in the following sections.

## 3. Results

### *3.1. General Description of Results*

The process of analysis and review of the selected articles was carried out; this evaluation allowed us to first generate some initial results (Section 3.2) related to the search procedures carried out and, second, to conceptualize four main classifications regarding relevant terrestrial and solar radiation models for this literature exploration (Section 3.3): (i) according to interpolation methods; (ii) according to remote sensors; (iii) according to applications; (iv) ML applied to Earth radiation modeling.

For the first three typologies, tables were elaborated to mainly synthesize the objective and results obtained from the developed research with the aim of organizing and specifying such a volume of relevant information more adequately. Subsequently, a general analysis of these summary tables is provided in text form. The fourth typology was exclusively developed in text form, where the most recent and most significant international contributions regarding the application of machine learning algorithms in building spatial distribution models with respect to terrestrial and solar radiation are presented.

Below, the syntactic procedures executed in this literature search are exposed, with special reference to the most outstanding results in some databases, through the analysis of different statistical and graphical tools.

### *3.2. Search Strategy: Procedures and Metrics*

Different search iterations were conducted in the SCOPUS, ResearchGate, and ScienceDirect databases by combining keywords, Boolean operators, and advanced search filters. These combinations were refined until much more precise results were obtained regarding machine learning techniques developed or applied to spatial interpolation geoprocesses for topics related to solar radiation, radiative energy budgets, or climate variables related to radiation.

Below, some examples are shown regarding the syntax applied in the searches conducted in the SCOPUS database specifically as well as some graphical and statistical results that allowed for the proper analysis of the data and information retrieved from it. The results of this exploration in the database were distributed between two phases. The first phase yielded a list of 31 relevant documents based on the following search parameters: TITLE-ABS-KEY ("machine" AND "learning" AND "+" AND "radiation" AND "+" AND "interpolation") AND PUBYEAR > 2017 AND PUBYEAR < 2024 AND (EXCLUDE (SUBJAREA, "MEDI")) AND (LIMIT-TO (PUBSTAGE, "final")).

At this point, the keywords entered in the search engine were "Machine Learning," "Radiation," and "Interpolation," and the temporal range used was from 2018 to 2023 (Figure 2a). Additionally, documents related to the field of medicine were excluded from the search, resulting in much more related and relevant fields of the topic of interest in this exploration (Figure 2b), and it was also limited to documents in the final publication stage.

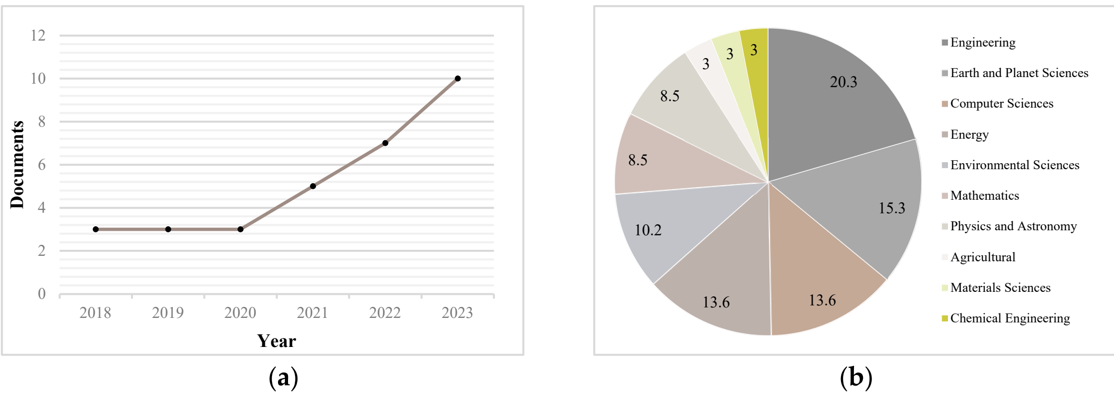


**Figure 2.** (**a**) Documents published in the time range established for the search (2018 to 2023); (**b**) Documents classified by thematic areas of development.

These results provided complementary information where the types of documents classified were highly diverse (Figure 3a), considering articles (71%) and conference papers, reviews, and data papers (29% among these three). The countries with the greatest academic contributions on the mentioned topic were China with 11, the United States with 5, and South Korea with 4 (Figure 3b).

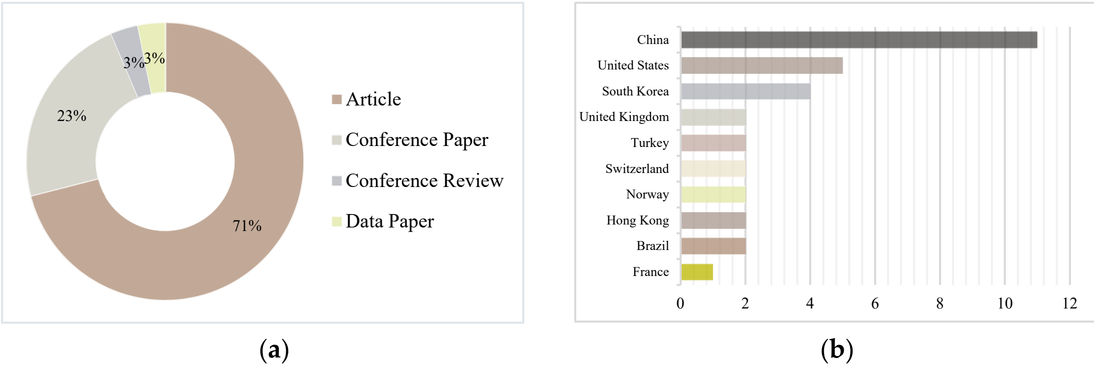


**Figure 3.** (**a**) Classification of search results by document type found. (**b**) Main academic contributions by country.

Subsequently, the second phase of searches related to the topic was carried out, further refining the search by limiting it to other document types besides articles, as conclusions and results in such documents tend to have a higher level of validation, precision, and evaluation. Additionally, the conjunction "solar radiation" was specified in the keywords, as the term "radiation" alone refers to various methods and practices in other fields of knowledge, often unrelated to the focus of this article.

The syntax that defined the results of this second phase of search was TITLE-ABS-KEY ("machine" AND "learning" AND "+" AND "solar" AND "radiation" AND "+" AND "interpolation") AND PUBYEAR > 2017 AND PUBYEAR < 2024 AND (LIMIT-TO (DOCTYPE, "ar")) AND (EXCLUDE (SUBJAREA, "MEDI")). A total of 22 documents associated with the topic of interest were obtained within the same temporal range of 2018–2023 (Figure 4a), and they were classified into a wide variety of knowledge areas, with this occasion being led by the area of planetary and earth sciences (21.1%) (Figure 4b).

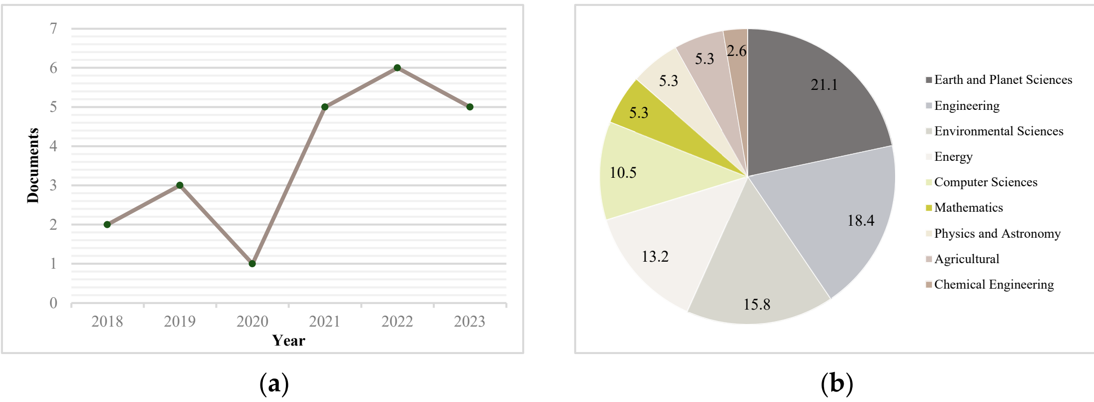


**Figure 4.** (**a**) Documents published in the time range established for the search (2018 to 2023). (**b**) Documents classified by thematic areas of development.

In this phase of results, a greater number of statistical data was obtained for analysis. Correspondingly, it was possible to demonstrate which institutions were involved or related to these research efforts, with Zhejiang University in China, The Hong Kong Polytechnic University, and Nord University in Norway leading these developments.

Finally, two results already seen in the first stage of the search were reported: the type of associated document, with only articles considered in this case, including both review and scientific articles (Figure 5a), and the classification of academic contributions by country, with China leading with nine articles, followed by the United States and South Korea with three publications each (Figure 5b).


(**a**) (**b**)

**Figure 5.** (**a**) Classification of search results by document type found. (**b**) Main academic contributions by country.

The search phases outlined were applied and replicated across the different aforementioned databases (resulting in the discovery of 17 additional articles), allowing for the satisfactory refinement and specification of the bibliographic results. This exercise evidenced the growing interest within the scientific community in developing research related to machine learning and spatial distribution models aimed at analyzing complex climate and meteorological variables such as solar radiation. This fact proved beneficial and of complete interest for this literature review document. The subsequent chapter is oriented towards consolidating these results and analyzing the most prominent classifications in this document.

### *3.3. Classification of Results*

This section highlights the primary classifications assigned to the research with the most significant contributions in this review. Firstly, classification was based on the most relevant interpolation methods. Secondly, the results was classified according to the most commonly used remote sensors. In third place, classification was assigned according to the applications with the greatest impact or development.

Finally, a section exclusively devoted to discussing the fourth classification, elaborated, "Machine Learning Applied to Terrestrial Radiation Modeling," was conducted due to its importance in this document as the central axis of discussion and its growing contributions, detailed below.

#### 3.3.1. Classification According to Interpolation Methods

Applying the same search methods previously outlined, advanced searches of recently published works related to interpolation techniques regarding solar radiation, with a greater trend of usage in recent years globally, were conducted. In this classification, primarily spatial autocorrelation methods and traditional geostatistics were considered, without a defined temporal range, in order to compare and analyze the results, utilities, or limitations of each of these investigations from a broader perspective. Below is a synthesis (Table 1) of the most relevant research organized by bibliographic reference, research title, spatial interpolation methods employed, and study conclusions.

**Table 1.** Classification according to interpolation methods.

| Ref. | Study | Method | Conclusions |
|---|---|---|---|
| [25] | "Upscaling of LongWave Downward Radiation (LWDR) from Instantaneous to Any Temporal Scale: Algorithms, Validation, and Comparison" | RFR | The random forest regression (RFR) method enables a precise and highly efficient estimation of the daily averaged LWDR. In comparison to existing methods and products, this method is efficient and exhibits superior applicability and reliable accuracy. |
| [26] | "A Machine Learning Technique for Spatial Interpolation of Solar Radiation Observations" | ML and seven interpolation methods | The average MAE (mean absolute error) of conventional interpolation methods is 21.3 W/m$^2$, more than double that of the RF method, with an average standard deviation of 6.4 W/m$^2$, over four times higher than RF. This underscores the benefits of employing machine learning in environmental research. |
| [27] | "Interpolation Methods Applied to the Spatialization of Monthly Solar Irradiation in a Region of Complex Terrain in the State of Rio de Janeiro in the Southeast of Brazil" | IDW, SpT, OK | The aim of this study was to assess different interpolation methods for the monthly average of daily irradiance. The Inverse distance weighting (IDW), tension spline (SpT), and ordinary kriging (OK) methods were evaluated. The IDW and SpT methods demonstrated satisfactory performance (coefficient of determination: $R^2 > 0.90$ for IDW and >0.80 for SpT). |
| [28] | "El Riesgo de Contaminación por Ozono en Dos Ciudades Españolas (Mad. Sev.) Estudio Hecho con Modelado Espacial y SIG." | IDW | The results obtained indicate that suburban or rural areas are the most contaminated by ozone, constituting risk areas for the population, whereas ozone values are much lower in city centers. |
| [10] | "Performance Comparison of Two Global Solar Radiation Models for Spatial Interpolation Purposes" | IDW, ANN | IDW is easier to implement and slightly more accurate than artificial neural networks (ANNs). Using data from sites with similar climate to the predicted region enhances the accuracy of the IDW model. |
| [29] | "REGNIE Interpolation for Rain and Temperature in the Andean, Caribbean and Pacific Regions of Colombia" | MLR | They employed the rainfall model REGNIE to interpolate rainfall and mean air temperature (Tmed) in Colombia. The model enhanced the accuracy and spatial resolution of the interpolations, particularly for Tmed, which achieved multiple linear regression (MLR) at $R^2 = 0.94$. |
| [30] | "Spatial Interpolation of Climate Variables in Northern Germany—Influence of Temporal Resolution and Network Density" | OK, KED | Geostatistical techniques had the best performance for all climatic variables. |
| [31] | "A Guideline to Select an Estimation Model of Daily Global Solar Radiation between Geostatistical Interpolation and Stochastic Simulation Approaches" | IDW, OK | Geostatistical methods provided better representation than simulation models in regions with a high density of stations. Among the geostatistical methodologies employed, OK exhibited the best performance. |
| [32] | "Mathematical Interpolation Methods for Spatial Estimation of Global Horizontal Irradiation in Castilla-León, Spain" | IDW, Spline, OK, NN | OK is the best interpolation method compared to the others, yielding the lowest errors. |
| [11] | "Spatiotemporal Interpolation and Forecast of Irradiance Data Using Kriging" | OK | Kriging can be applied to both ground-measured data and satellite-derived irradiance data. For interpolation, sparsely distributed ground data are recommended; however, when temporally scaled or forecasted, OK shows promise for both data sources. |

**Table 1.** *Cont.*

| Ref. | Study | Method | Conclusions |
|---|---|---|---|
| [33] | "Assessing the Performance of Several Rainfall Interpolation Methods as Evaluated by a Conceptual Hydrological Model" | IDW, OK, KED | The performance of deterministic methods is comparable to geostatistical methods in daily series. In hourly series, deterministic methods showed significantly better performance. |
| [34] | "Comparing Interpolation Techniques for Monthly Rainfall Mapping Using Multiple Evaluation Criteria and Auxiliary Data Sources: A Case Study of Sri Lanka" | IDW, OK, TPS | Most methods approximated the spatial distribution of precipitation at a high level for May. Thin plate splines (TPS) performed better with high precipitation, and their pattern was smooth. |
| [35] | "Ground-measurement GHI Map for Qatar" | IDW | The combination of satellite and ground-based solar data provided more reliable values with lower uncertainties. |
| [36] | "El Balance de Radiación y Modelos de Radiación Neta para Diferentes Superficies: Estudio Experimental en Mexicali, México" | MLR | Statistical models of net radiation were proposed as functions of incoming solar radiation and net shortwave radiation, with $R^2$ exceeding 0.97. |
| [37] | "Missing Data Imputation of Solar Radiation Data under Different Atmospheric Conditions" | IDW. MICE, MLR | They measured global solar irradiance on a flat surface using a network of stations located in Galicia, Spain. The best results were obtained with the multiple imputation by chained equations (MICE) method. |
| [38] | "Comparison and Evaluation of Spatial Interpolation Schemes for Daily Rainfall in Data Scarce Regions" | IDW, OK | Differences between methods can be significant on a smaller temporal and spatial scale (monthly). The choice of covariates in OK had a large impact on the amount of precipitation and runoff. |
| [39] | "Interpolation of Daily Solar Radiation for New Zealand Using a Satellite Derived Cloud Cover Surface" | TPS | The lowest error was obtained when calculating radiation fields using satellite data. |
| [40] | "Spatial Interpolation and Estimation of Solar Irradiation by Cumulative Variograms" | Cumul. variogram | The authors estimated the solar irradiation value at any point where measurements of global solar irradiation existed. The goal was to determine the change in spatial variability with distance from a given set of irradiation data. |

### 3.3.2. Classification According to Sensors

Following the phase of classifying existing terrestrial radiation models, advanced searches were conducted to establish a categorization according to data obtained from satellite remote sensors associated with climate variables such as solar radiation or covariates closely related to it (Table 2). Below are some of the most commonly used sensors and their prominent applications:

**Table 2.** Classification according to sensors.

| Sensor | Description/Use |
|---|---|
| GeoCarb, OCO-2 OCO-3, CERES RAVAN | Satellite data provide an independent means of investigating global temperature trends, particularly for ocean surface and atmosphere. Sea surface temperatures (SSTs) of oceans, which are directly related to heat transfer between the atmosphere and oceans, serve as important indicators of the state of the climate system [41]. |
| AIRS, CERES, MODIS, AMSR-E, GRACE FO, ICESat-2, SUOMI NPP | Snow and ice cover retreat is a significant indicator of global warming. Seasonal snow and ice melt can cause positive feedback by reducing Earth's surface albedo, thereby contributing to sea level rise [42,43]. |
| ISS-RapidScat GRACE, Jason-3 Jason CS, OSTM | Sea level depends on climatic conditions influenced by climate change (radiative forcing factors) and climate variability [44]. |

**Table 2.** *Cont.*

| Sensor | Description/Use |
|---|---|
| DSCOVR<br>LIS, RAVAN<br>CERES | Monitoring changes in solar luminosity is important to ascertain whether natural variation in solar radiation has significantly contributed to recent climate change. Additionally, studying complex processes derived in various areas such as plant physiology, photosynthesis, evapotranspiration, absorption of photosynthetically active radiation, or solar light use efficiency is crucial [45]. |
| OCO-3, MLS<br>OMI, TROPOMI<br>PACE, SAGE III<br>SUOMI NPP<br>MODIS, LIDAR | Particles in the atmosphere known as aerosols can generate either cooling or warming effect depending on their type in the system. Recent changes in atmospheric aerosol concentration have been identified through aerosol optical depth (AOD), derived from observations recorded by visible and infrared optical sensors aboard various satellites. Many other sensors monitor gases with significant environmental implications, such as $CO_2$, $NO_2$, $CH_4$, $O_3$, CO, and $SO_2$ [46]. |
| CloudSat<br>CERES, PACE<br>MODIS, VIRS | Climate forcing and cloud feedback adjust energy flux throughout the Earth's system. Cloud dynamics research is fundamental to understanding climatic patterns in periods of climate variability and extreme events [47]. |
| GPM, TOVS<br>RAINCUBE<br>AVHRR, GEOS | Water vapor is a gas with a high heat retention capacity, contributing approximately 50% to the current global greenhouse effect. Models predict that global warming will increase atmospheric specific humidity (resulting in positive feedback) and, in turn, strongly amplify warming [48]. |

#### 3.3.3. Classification According to Applications

The third of the classifications generated in this chapter was established to demonstrate the different applications of spatial models in various fields of knowledge (Table 3). This classification specifically highlighted the importance of studies on spatial distribution of different climatic or meteorological variables according to their prominent applications:

**Table 3.** Classification according to Applications.

| Ref. | Study/Application |
|---|---|
| [7] | The variation in concentrations of hydrogen sulfide ($H_2S$) and ammonia ($NH_3$) in air was studied in a wastewater treatment system in Costa Rica using an air substance dispersion model. Similarity between predicted and observed values for $H_2S$ and $NH_3$ was demonstrated, indicating no risk to health or the environment. |
| [8] | A spatial model of weed distribution under climate change was developed; fifty-nine articles were selected for review, with maximum entropy (MaxEnt) and area under the curve (AUC) being the most popular validation methods and processes. |
| [9] | Spatial autocorrelation was combined with artificial intelligence models to estimate the spatial distribution and risks of heavy metal contamination in agricultural soils. Spatial distribution models were developed using ANN models and adaptive neuro-fuzzy inference systems (ANFIS). |
| [49] | Creation of spatial models using maximum entropy (Maxent) methods. Identification of environmental variables controlling the distribution of greenhouse gas sources in Finland peatlands and prediction of the spatial distribution of these gases. |
| [50] | A method for modeling extreme wind speed distributions using data collected from the Swiss meteorological service. This work applied spatial modeling of distribution parameters of this phenomenon throughout the country. |
| [51] | Analysis of longwave radiation in Earth radiation budgets through topographic models and mesoscale studies of longwave solar radiation. It was determined that most existing satellite-based data capture algorithms are valid only for flat surfaces without considering topographic effects. |
| [52] | A study aimed to parameterize the Bristow–Campbell model to estimate daily global solar radiation (DGSR) on the Tibetan Plateau and propose a method to rasterize it. The spatial pattern of DGSR distribution was estimated and analyzed by combining the solar radiation model (Bristow–Campbell) and a meteorological interpolation model called PRISM. |
| [53] | A model was proposed to estimate the spatial distribution of average monthly global solar radiation for central Chile considering the continental zone from the Coquimbo region to the Araucanía region. |
| [54] | Research based on climate change and spatial distribution of vegetation formations in Colombia. |

### 3.3.4. Machine Learning Applied to Earth Radiation Modeling

In recent years, various areas and subareas derived from artificial intelligence have experienced significant development, being considered high-growth fields contributing to knowledge and scientific discourse. Machine learning is a branch of artificial intelligence [55] that analyzes the use of computational algorithms to convert empirical data into usable models (Figure 6). Additionally, instead of programming each rule within a model, the system utilizes data to train itself and improve its performance, thus facilitating the ability to make predictions or decisions without direct human intervention. The field of machine learning emerged from traditional communities of statistics and artificial intelligence [56].

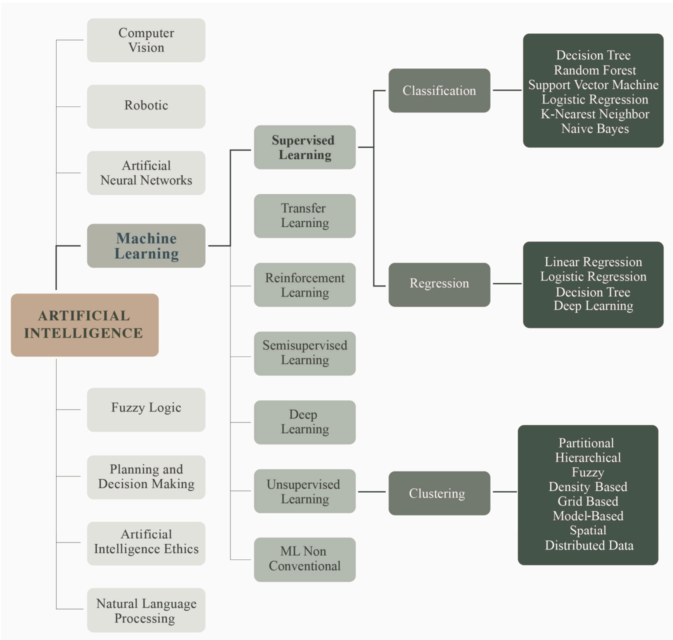


**Figure 6.** Diagram representing the main branches of artificial intelligence and ML.

Machine learning has also made significant contributions to spatial data sciences, geomatics, geoengineering, and geotechnologies [57], as these fields maintain a close relationship with the management and processing of highly complex and voluminous data. Consequently, machine learning algorithms exhibit an appropriate and precise level of adjustment for the tasks inherent in geoprocessing executed on spatial data. The functionalities and tools that machine learning has provided to these fields are currently undergoing rapid expansion. This chapter presents the most relevant studies published in the last five years on the application of machine learning to models of spatial distribution of terrestrial radiation at the international level.

Initially, the research conducted by [58] is highlighted, as they focused on interpreting the feasibility of various ML approaches to assess the spatial distribution of solar radiation on Earth based on the Meteorological, Oceanic, and Communication Satellite's (COMS) geostationary satellite Meteor Imager (MI). Four data-driven models were selected—artificial neural networks (ANNs), random forest (RF), support vector regression (SVR), and deep neural networks (DNNs)—to compare their accuracy and spatial estimation performance. The analysis results showed that RF had the highest accuracy in predicting performance, although the difference between RF and the second-best technique (DNN) was nonsignificant. Temporal variations in root mean square error (RMSE) depended on the number of data samples, while the physical model showed relatively less sensitivity. However, DNN and RF showed less variability in RMSE than the others. The overall analysis showed that deeper layered network approaches (RF and DNN) could better simulate the challenging spatial pattern of thin clouds when using multispectral satellite data.

Another proposal made in 2019 by [59] aimed to estimate the solar radiation potential in an extensive region of China for energy utilization. They proposed developing a novel estimation approach for monthly average daily solar radiation with its complex spatial pattern through machine learning techniques (i.e., a clustering method (k-means) and an advanced case-based reasoning (A-CBR) model. Relevant information was collected in 97 cities in China over nine years (from 2006 to 2015). The average prediction accuracy of the proposed approach was determined to be 93.23%, showing a promising path towards the implementation and development of interpolation methods (e.g., kriging method in geographic information systems).

In 2020, Ref. [60] published research on solar radiation estimation using ordinary kriging and distance weighting, plus an alternative method, based on an unsupervised competitive artificial neural network called a self-organizing map. This neural network generated a map with the most representative nodes and their weights, which were used to obtain the spatial variability of solar radiation. This work provided a new alternative to obtain the spatial variability of a regionalized variable based on a self-organizing map and opened future lines of research, including optimal sensor network design or variability study for different sampling times, with the aim of using mobile sensors to measure or extend to other continuous variables in space, such as temperature and humidity.

In the same year, the work by [61] reviewed the progress of four advanced machine learning methods for handling spatial data, namely, kernel-based learning based on support vector machines, active and semi-supervised learning, ensemble learning, and deep learning. Among the relevant conclusions of this work (Figure 7), the authors determined that the successful application of machine learning methods for handling spatial data suggests a four-level strategy.

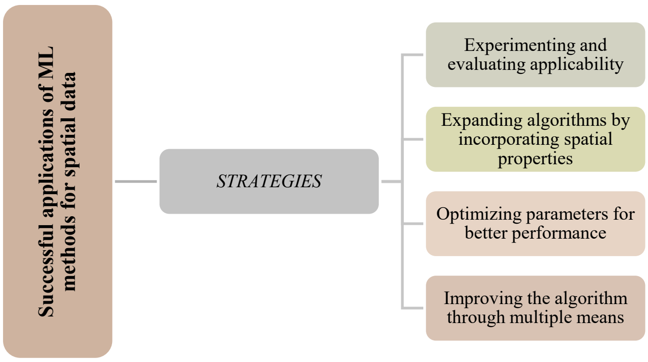


**Figure 7.** Diagram illustrating the main strategies for applying ML to spatial data.

They concluded by affirming that ML methods are highly effective for handling spatial data and have broad potential application in the era of big data.

Equally relevant was the work carried out in 2021 by [26], where they applied statistical methods to interpolate missing values in a dataset of radiative energy fluxes on the Earth's surface. They applied RF algorithms and seven other conventional spatial interpolation models to a global dataset of surface solar radiation (SSR). Tenfold cross-validation showed that RF had a mean absolute error (MAE) of 10.2 W/m$^2$ and a standard deviation of 1.5 W/m$^2$. On the other hand, the average MAE of conventional interpolation methods was 21.3 W/m$^2$, which is more than double that of the RF method, plus an average standard deviation of 6.4 W/m$^2$, which is more than four times higher than the standard deviation of RF. This work highlighted the benefits of using machine learning in climate research.

Once again, ref [26] in 2021 performed spatial interpolation based on the RF machine learning method to interpolate monthly anomalies in downward solar radiation (RSD) from the surface using a range of climatic variables (various temperature indices, cloud cover, etc.), temporal point indicators (years and months of observations of this radiation), and geographical characteristics of locations (latitude, longitude, etc.). The result of spatial interpolation was a monthly dataset on a 0.5° × 0.5° grid of RSD anomalies with global coverage during the period of 1961–2019, which was subsequently used in a comprehensive trend analysis for each continent separately and for the entire globe. The continent-level analysis revealed the major contributors to global dimming, understood as the process of decreasing incoming radiative energy to the earth's system due to greenhouse gases (GHGs) in the top of the atmosphere (TOA) preventing shortwave radiation entry as well as global brightness, understood as the amount of energy entering, being emitted, and reflected (albedo) in the earth's system. Particularly, global dimming before the 1980s is primarily dominated by negative trends in Asia and North America, while Europe and Oceania were the two largest contributors to brightening from 1982 to 2019.

Ref. [62] in 2022 developed a study to predict land surface temperature (LST) using new hybrid models based on spatial interpolation techniques and deep learning models; they employed multilayer perceptron (MLP), long short-term memory (LSTM), and an integrated machine learning model, specifically convolutional LSTM (ConvLSTM). Data were collected from the MODIS sensor with a spatial resolution of 1 km × 1 km. Considering the inability of MODIS sensors to provide LST data under cloudy conditions, IDW, nearest neighbor (NN), and cubic spline interpolation methods were used to overcome the issue of missing pixels. The proposed methods were tested in the northern part of Adana province, Turkey, and model performance was quantitatively evaluated using RMSE and MAE. The selected datasets ranged from 1 January 2017, to 1 November 2020, for daily LST and from 1 January 2015 to 1 November 2020 for an 8-day composite LST. While 60% of the datasets were used as training sets, the remaining 40% were used as validation (20%) and test (20%) sets. RMSE maps were generated to assess the pixel-wise performance of the proposed method. The best RMSE and MAE averages for the daily test set were obtained from the combination of ConvLSTM and NN, yielding 3.62 °C and 2.85 °C, respectively, while those acquired from the combination of MLP and NN for the 8-day composite LST test set were 3.57 °C and 2.69 °C. The results revealed that the proposed hybrid models could be used for spatiotemporal one-step-ahead prediction of LST data.

Also in 2022, Ref. [63] published a study related to the estimation and simplified modeling of solar irradiance at the Earth's surface in Australia and China based on the application of four machine learning models: gradient boosting machine (GBM), RF, support vector regression (SVR), and multilayer perceptron (MLP). Solar irradiance data under clear skies were calculated based on time and location; meanwhile, cloud optical thickness (COT) and aerosol optical thickness (AOT), significantly influencing solar irradiance, were retrieved from Himawari-8 meteorological satellite images, and surface solar irradiance data were obtained from observation stations for training and evaluation. As a comparative study, historical data from six years prior were collected, and solar distribution was estimated with a spatial resolution of 5 km in Australia and China. Based on the coefficient of determination ($R^2$), normalized root mean square error (nRMSE), normalized mean bias error (nMBE), and time consumption (t), the results showed that GBM achieved the highest accuracy with $R^2 > 0.7$, followed by RF, SVR, and MLP.

In May 2023, Ref. [64] published a study in which aerial measurements of latent and sensible heat fluxes collected in the Netherlands in August 2008 were taken, and five machine learning methods were evaluated to construct a gridded map of regional heat fluxes, including ANN, BRT, RFR, DNN, and SVR. The models were trained and tested using a dataset compiled from sensible and latent heat observations, digital elevation model, land surface temperature retrieved by the MODIS sensor, enhanced vegetation index, and albedo data for each flux footprint. The support vector regression model outperformed the other models, with $R^2 = 0.91$, RMSE = 9.6 W/m$^2$ for sensible heat and $R^2 = 0.89$, RMSE = 26.32 W/m$^2$ for latent heat.

Two more studies conducted in 2023 are highlighted for their contributions of machine learning to terrestrial radiation modeling; these were developed by [65], and the authors focused on predicting surface solar radiation and diffuse solar radiation, respectively, using machine learning techniques in radiative transfer models. This initial study developed hybrid models with a high computational speed and high accuracy to estimate global solar radiation (GSR) and quantify the uncertainty in GSR simulations caused by uncertainties in atmospheric and surface parameter measurements. The radiative transfer model library (LibRadtran) was combined with six machine learning models: extreme gradient boosting (XGBoost), random forest (RF), multivariate adaptive regression splines (MARS), multilayer perceptron (MLP), deep neural network (DNN), and light gradient boosting machine (LightGBM).

The estimated GSR was first compared with GSR inversion values provided by the Aerosol Robotic Network (AERONET) and then validated using ground measurements at three locations in China between 2005 and 2018. The results showed that the RTM-RF was superior in terms of efficiency and computational performance, with an MAE and $R^2$ of 15.57 W/m$^2$ and 0.98, respectively. Under clear sky conditions, aerosol optical depth (AOD) contributed the most to the accuracy of GSR estimates, with an average contribution of 57.95%. Measurement uncertainty due to the asymmetry factor, AOD, single scattering albedo, and land surface albedo (LSA) could explain differences in GSR between RTM estimates and GSR observations in Lulin (20.33 vs. 20.91 W/m$^2$), Wuhan ($-$1.40 vs. 14.58 W/m$^2$), and Xianghe (7.28 vs. 14.32 W/m$^2$). In conclusion, this work supported the use of physical models combined with machine learning models to estimate GSR and provided valuable scientific information for solar radiation estimates in large areas.

The second study [66] emphasized the development and comparison of five high-speed, high-precision hybrid models for simulating diffuse radiation based on aerosol optical properties and radiation parameters provided by the Aerosol Robotic Network (AERONET); Baseline Surface Radiation Network (BSRN); Wuhan University, China, Ecosystem Research Network (CERN); GLASS surface albedo data; and a combination of the radiative transfer model (RTM) with ML models, including RF, XGBoost, MLP, DNN, and convolutional neural networks (CNNs). Additionally, the uncertainty in simulated diffuse radiation due to measurement uncertainties in aerosol optical properties and land surface albedo was quantified, and the relative contributions of multiple variables to diffuse radiation were analyzed. The results showed that RTM-RF was the most successful, with determination coefficients ($R^2$) of 0.95, 0.94, and 0.98 and minimum root mean square errors (RMSEs) of 9.56, 10.05, and 13.27 W/m$^2$ in Lulin, Wuhan, and Xianghe, respectively. Overall, the proposed RTM-RF method showed superior performance, making it recommended for estimating diffuse radiation in China.

Regarding the most recent and highly contributive works collected in this review, Ref. [12] presented a new method for predicting solar irradiance directly correlated with photovoltaic energy production; in this work, a fast statistical learning technique based on a truncated-regularized kernel ridge regression model was employed. The proposed method exceled in predicting solar irradiance, especially during highly intermittent weather periods and by incorporating multiple historical climatic parameters to generate accurate predictions of future values. The model's performance was evaluated using datasets from cloudy and sunny days in Seattle and Medford, USA, which were compared using three forecasting models: persistence, modified 24-h persistence, and least squares. Based on three widely accepted statistical performance metrics (mean squared error, mean absolute error, and coefficient of determination), the hybrid model demonstrated superior predictive accuracy under different weather conditions and forecast horizons.

Ref. [13] published a study aimed at retrieving LST data generated from high-density station observation data and remote microwave sensors for deep learning-based monitoring (U-Net family). Given the high spatial and temporal variability of LST and its sensitivity to various factors according to radiation transmission equations, this study incorporated climate, anthropogenic, geographic, and vegetation datasets to facilitate a multi-source data fusion approach for estimating this variable. The accuracy of the LST simulation showed favorable results in spatial and temporal dimensions. Additionally, temporal changes in simulation accuracy fluctuated in a W-shaped pattern due to the limited simulation capacity of deep learning in extreme scenarios. Anthropogenic factors had the greatest influence on LST changes in China, followed by climate, geography, and vegetation. This study highlighted the application of deep learning in remote sensing monitoring in the context of "big data" and provided a scientific basis for climate change response to human activities, ecological protection of the environment, and sustainable social and economic development.

Another research of particular interest in this dissertation was conducted by [14]. Although the solar radiation variable was not directly analyzed, it was indirectly considered for daily evapotranspiration (ET) estimates with a novel approach in the Heihe River Basin on the border between Russia and China. Remote sensing-based models generally struggle to generate continuous spatiotemporal ET due to cloud cover and model failures. In this study, four popular machine learning methods (DF, DNN, RF, XGB) were used to reconstruct the ET product. The reconstructed ET by these four methods was evaluated and compared; the results showed that all methods performed well in the Heihe River Basin, but the RF method was particularly robust. Not only did it perform well compared to ground measurement ($R^2 = 0.73$), but it also reconstructed ET across the entire basin. Validation based on ground measurements showed that the DNN and XGB models performed well ($R^2 > 0.70$). However, gaps still remained in the desert after reconstruction, especially in the case of the XGB model. The DF model filled these gaps across the entire basin, but the model had lower consistency compared to ground measurement ($R^2 = 0.66$) and produced many low values. The results of this study suggested that machine learning methods have considerable potential in ET reconstruction at a regional scale.

Lastly, Ref. [15], with the aim of finding suitable locations for future photovoltaic systems placement, presented a method based entirely on convolutional neural networks for high-resolution spatiotemporal evaluation of solar potential using remote sensing data. The method was trained and validated in a 32 km$^2$ area of Maribor, Slovenia, and tested at six different locations in Slovenia and Germany. The proposed method was tested with a simulation algorithm, which used either the Liu Jordan isotropic diffuse model or the Pérez anisotropic model with shading based on the sky view factor. On average, a normalized RMSE (nRMSE) of 6.06% and a mean absolute percentage error (MAPE) of 4.29% were achieved at test locations compared to the Liu–Jordan model-based simulation. By training fully convolutional neural network models against ground truth observations generated with a more advanced Pérez diffuse model, the average nRMSE was 8.37% and a MAPE of 10.90% across all test locations (Table 4). Additionally, it was demonstrated that the proposed method can assess solar potential at a high resolution from an input at a lower resolution more accurately than the simulation algorithm, while being up to 1500 times faster.

**Table 4.** Summary of predictive performances in the application of algorithms with ML.

| Author | Method or Algorithm | Predictive Performance (Better Performances) |
|---|---|---|
| [26] | RF | **RF:** MAE: 10.2 W/m$^2$<br>MAE Conventional Methods: 21.3 W/m$^2$ |
| [63] | GBM | **$R^2$**: 0.7 |
| [64] | ANN, BRT, RFR, DNN and SVR | **SVR:** $R^2 = 0.91$, RMSE = 9.6 W/m$^2$ in sensitive heat,<br>$R^2 = 0.89$, RMSE = 26.32 W/m$^2$ in latent heat. |
| [65] | RTM-RF | **RTM-RF** MAE = 15.57 W/m$^2$<br>$R^2 = 0.98$ |
| [64] | XGBoost, MLP, DNN, CNN RTM-RF | **RTM-RF** $R^2 = 0.95$ Lulin;<br>0.94 Wuhan;<br>0.98 Xianghe<br>RMSE = 9.56W/m$^2$ Lulin;<br>10.05 W/m$^2$ Wuhan;<br>13.27 W/m$^2$ Xianghe |

**Table 4.** *Cont.*

| Author | Method or Algorithm | Predictive Performance (Better Performances) |
|---|---|---|
| [14] | DF, DNN, RF, XGB | **RF:** $R^2$ = 0.73<br>**DNN** and **XGB** $R^2$ => 0.70<br>**DF**: $R^2$ = 0.66 |
| [15] | CNN | **CNN**: nRMSE average = 8.37%<br>MAPE = 10.90% |
| [59] | k-means and A-CBR model | **Average prediction** = 93.23% |
| [58] | ANN, RF, SVR, DNN | **RF & DNN** Superior prediction and better RMSE than **ANN SVR** |
| [62] | MLP, LSTM, ConvLSTM | **ConvLSTM & NN**: RMSE = 3.62 °C/MAE = 2.85 °C<br>**MLP & NN**: RMSE = 3.57 °C/MAE = 2.69 °C |

Machine learning has made significant contributions in the past five years in the field of spatial modeling, specifically in the study of variables such as solar and terrestrial radiation; a considerable number of complex algorithms have been applied in various disciplines with mostly scientific and technological interests. The following chapter discusses these advances, details some of the limitations and challenges that this area of artificial intelligence still faces, and highlights developments with significant contributions and trending topics today.

## 4. Discussion

In summary, the studies mentioned in this article demonstrate a search for suitable methodologies for solar parameter modeling (mainly interpolations), showing that other interpolation methods applied to solar and terrestrial radiation yielded better results than the traditionally used IDW method. One of the geostatistical methods with better results, as demonstrated in this literature review, was the kriging interpolator; for the analysis of variables such as solar radiation and atmospheric gas modeling [67], its widespread use has been almost undisputed due to favorable results that have optimized the data's capacity for spatial analyses. This technique is based on statistical assumptions and is employed in advanced surface predictions.

The method works by using a specific number of points or all points within a coverage radius to determine the resulting value for each location and is more suitable when the distance is known [68]. In kriging techniques, some probability is associated with the predictions; that is, values are not perfectly predictable from a statistical model. For example, in a sample of net radiation values measured in a geographic space, obviously, even with a considerable number of samples, the exact value of net radiation at an unmeasured location cannot be predicted. Therefore, kriging not only attempts to predict it but also evaluates the error of that prediction. As seen, the applications of this technique are diverse in atmospheric sciences, soil geology, and hydrology.

The classifications conducted enabled the establishment of chronological and succinct differences that deserve to be highlighted. In the first typology, "According to Interpolation Methods," it was of significant interest to demonstrate how, prior to the rise of algorithms from artificial intelligence and machine learning, techniques such as inverse distance weighting (IDW) and ordinary kriging (OK) predominated in a substantial proportion (52.94%) of the studies collected across various fields of knowledge globally (Figure 8).

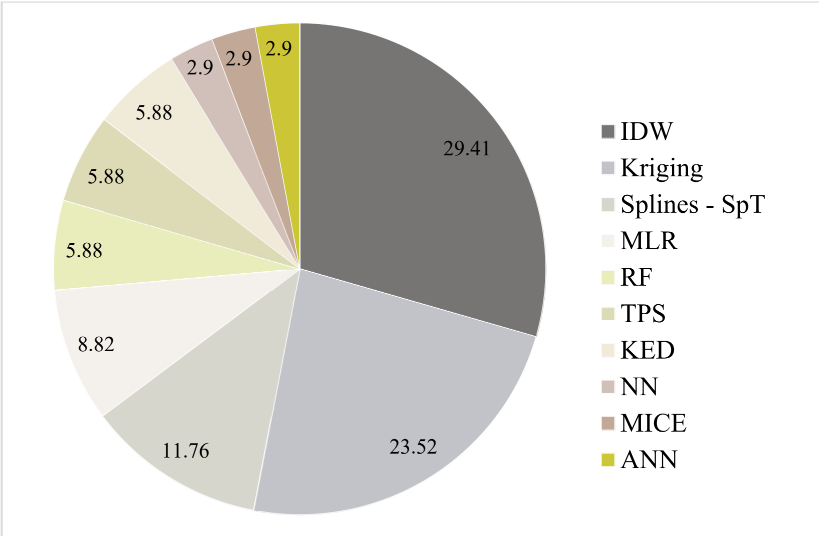


**Figure 8.** Spatial interpolation methods traditionally used in search results (percentages).

These techniques were based on principles of geostatistical inference and spatial autocorrelation and showed satisfactory or promising results in their application to climatic or meteorological variables such as global solar radiation; daily solar radiation; ozone concentrations; monthly, daily, and hourly precipitation; average temperature; absolute humidity; and global horizontal irradiance.

On the other hand, the second classification, 'According to Sensors', allowed for the distinction of the presence of some satellite projects or missions that acquired significant data of high interest to the scientific community on topics related to climate change, land and ocean observation/monitoring, meteorological data acquisition, detection of changes, and studies related to the presence and distribution of pollutant gases. The following sensors were noteworthy for the reliability of their data and the solid consolidation of their projects: OCO, CERES, MODIS, GRACE, LANDSAT, TROPOMI, PACE, SENTINEL, IceSAT, SUOMI, LIDAR.

The third classification, 'According to Applications', revealed a wide variety of disciplines interested in spatial modeling across diverse variables globally, ranging from the identification of greenhouse gas sources to the distribution of plant species in certain regions. This further underscores the urgent need to continue making contributions to the development of the field of spatial distribution models due to their extensive utility, as evidenced in this review.

Furthermore, based on the bibliometric analysis and literature review, it has been established that machine learning algorithms are increasingly taking center stage in the development and application of techniques related to the spatial modeling of radiative energy worldwide, as their results and applications have demonstrated higher precision parameters and broader diversification spectra. Some of the reviewed research has made interesting contributions in:

1. Solar and terrestrial radiation prediction: Models successfully incorporated factors such as geographical location, time of day, date, and weather conditions to make more accurate predictions. An example of this is [12], which presented a new method for predicting solar irradiance (SI) directly correlated with photovoltaic energy production. The proposed method exceled in predicting SI, especially during highly intermittent weather periods, by incorporating multiple historical climatic parameters to generate accurate future value predictions.
2. Spatial interpolation: Solar and terrestrial radiation values were estimated at intermediate locations between known measurement stations, thereby improving the spatial resolution of the data. Two studies were particularly relevant: the first by [25] applied

RFR, enabling precise and highly efficient estimation of daily averaged longwave downward radiation (LWDR). Compared to existing methods and products, this method is efficient, exhibits superior applicability, and provides reliable accuracy. The second study by [26] compared RF algorithms with seven traditional interpolation methods to determine the benefits of using machine learning in solar radiation observations.

3. Data correction: Machine learning algorithms were used to correct errors or inconsistencies in collected solar and terrestrial radiation data, enhancing the quality of datasets and making predictions more reliable. Notably, [37] contributed to the imputation and correction of missing solar radiation data under different atmospheric conditions through the application of the multiple imputation by chained equations (MICE) method.
4. Solar system design optimization: This involves assisting in the planning of the location and orientation of solar panels to maximize their efficiency considering the spatial and temporal variability of solar radiation. An exemplary contribution is the research by [59], which proposed a novel estimation approach for monthly average daily solar radiation with a complex spatial pattern using machine learning techniques, focusing on estimating the solar radiation potential in an extensive region of China for energy utilization. Additionally, the research by [15] is noteworthy, presenting a method based on CNN for high-resolution spatiotemporal evaluation of solar potential using remote sensing data. The proposed method was tested with a simulation algorithm to find suitable locations for future photovoltaic systems placement.
5. Adaptation to changing weather conditions: Models were dynamically adapted to changing weather conditions to provide real-time estimates of meteorological variables. The contribution by [14] is particularly valuable, focusing on estimating daily evapotranspiration (ET) with a novel approach based on four popular machine learning methods (DF, DNN, RF, XGB) to reconstruct the ET product. Remote sensing-based models generally struggle to generate continuous spatiotemporal ET due to cloud cover and model failure. The results showed that all methods performed well, with the RF method being particularly robust.
6. Uncertainty analysis: Machine learning algorithms evaluated the reliability of solar radiation budget estimates, facilitating informed decision making. The work by [65] is exemplary, focusing on predicting surface solar radiation and diffuse solar radiation by combining six ML techniques (XGBoost, RF, MARS, MLP, DNN, and LightGBM) in radiative transfer models. This initial study developed hybrid models with high computational speed and high accuracy to estimate global solar radiation (GSR) and quantify the uncertainty caused by ambiguities in atmospheric and surface parameter measurements.
7. Deep learning models: Techniques such as convolutional and recurrent neural networks were applied to capture complex and nonlinear patterns in solar and terrestrial radiation data, leading to more accurate predictions. Therefore, the work of [15], which focused on the high-resolution spatiotemporal assessment of solar potential from remote sensing data using deep learning, is important to highlight.
8. Integration of multiple data sources: The fusion of data from various sources, such as satellite images, weather stations, and historical data, improved the accuracy of solar radiation estimation. In this regard, the research by [26] in 2021 stands out for integrating various types of data, including a range of climatic variables (various temperature indices, cloud cover, etc.), temporal point indicators (years and months of radiation observations), and geographical characteristics of locations (latitude, longitude, etc.). This research focused on interpolating monthly anomalies in downward solar radiation (RSD) from the surface based on the RF machine learning method.

However, despite all these advances and contributions of machine learning in terrestrial and solar radiation models with high impacts on scientific research, it is also important to highlight that it is a growing field that still faces significant challenges and improvements. To highlight some that have been identified in this dissertation:

- Temporal and spatial variability: Solar radiation varies over time and space. Models must be able to capture this variability to provide accurate estimates.
- Missing or noisy data: A lack of data for certain locations or times can affect prediction accuracy. Additionally, the presence of outliers may require preprocessing techniques such as statistical assessments of data quality.
- Model calibration: It is important to calibrate and validate models using independent datasets to ensure their robustness and generalization.
- Model interpretation: The inherent complexity of some machine learning models, such as neural networks, can make interpreting results challenging. Developing interpretable models is crucial for gaining user and urban planner confidence.
- Scalability: Model scalability should be considered for large-scale applications, such as city or regional planning.

On the other hand, considering the uses and applications of different machine learning algorithms in various disciplines, it is interesting to highlight some sectors that would experience significant effects thanks to the practical contributions of these trending techniques as identified in this literature review. In energy resource management, for example, ML is improving efficiency in managing solar energy-based resources such as solar parks and distributed generation systems [69]. Also, in urban planning, ML is helping with sustainable urban organization by considering solar radiation in building and urban area design [70]. In electrical grid operation, ML could facilitate solar energy integration into power grids through accurate solar radiation prediction. Finally, in scientific research, which is the focus of this discussion, ML contributes to scientific development and innovation regarding solar radiation variability and its impact on climate and the environment.

Additionally, the trends established after analyzing the data and information collected in this review point to topics such as multispectral data integration, where recent research has explored integrating this type of satellite-derived data to improve model accuracy [71]. These data include more detailed information about atmospheric composition and land surface conditions. Transfer learning approaches are another trending example, as they are being developed to improve model generalization to different geographical regions. These approaches seek to leverage knowledge learned in one region to improve accuracy in another.

In situ model validation, on the other hand, is another growing field; here, by comparing predictions with data from weather stations and local measurements, the aim is to ensure model robustness and reliability [72]. On the other hand, integrated platform development also stands out as a trending field, as recent research has sought to develop platforms that combine spatial solar radiation distribution models with urban planning and renewable energy tools. These platforms facilitate decision making for effective implementation of solar infrastructures.

The use of high-resolution remote sensing data is another area with significant development expectations, as through state-of-the-art satellite images, it allows for better capture of spatial variability and topography [73], thus improving model accuracy. Moreover, applications in solar energy for vehicles are also growing interestingly. Recent research exploring the application of spatial solar radiation distribution models in the design and planning of charging infrastructure for solar-powered electric vehicles [74] is currently of paramount concern in relation to the energy transition and the global elimination of fossil fuels. Lastly, climate impact analysis is a burgeoning field in scientific discourse, as evidenced by the applications reviewed in this document. Some studies have focused on understanding the impact of climate change on the spatial distribution of associated phenomena and how models can adapt to account for these variations, enabling the development of informed mitigation strategies [75].

The utilization of machine learning in models of terrestrial and solar radiation spatial distribution offers substantial benefits in terms of availability and functionalities. However, addressing the associated challenges is crucial to ensure the accuracy and practical applicability of these models in advancing the field of remote sensing. Collaboration between data scientists, energy experts, and geomatics is essential to advance in this area, maximizing understanding of solar energy budgets and its potential while simultaneously contributing to significant advances in scientific research to build viable solutions to the complex climate variability and change scenarios currently projected.

## 5. Conclusions

The most relevant conclusions and the greatest contributions of this document are described below. The first is that this review, which belongs to the field of geotechnologies, has contributed to the reduction of deficiencies in the production of comparative studies conducted in Spain concerning the development and applications of ML in spatial modeling of solar and terrestrial radiation. Consequently, it demonstrates the primary research and contributions on this topic and presents at least a promising panorama of how novel techniques and algorithms (e.g., ANN, DNN, or MLP) and the combination of sections in their processes have generated interesting results according to particular application interests.

Regarding the proposed methodology, it allowed us to satisfactorily find relevant results that contributed to this discussion of geostatistical techniques and ML algorithms compared in the modeling of solar and terrestrial radiation budgets in recent years. The combination of keywords, Boolean operators, and advanced searches in scientific databases resulted in an efficient method, not only for finding validated and quality information but also for limiting and excluding works that were not convenient to include in this analysis for various reasons, such as having concepts in common but focused on other fields of knowledge or being close due to transversal topics of discussion but nevertheless distant in the central objectives of radiation modeling at a general level.

Another additional conclusion from the review is that in the last five years, developments in ML algorithms and their combination of techniques (which are ultimately sophisticated and refined geostatistical processes) have gained significant prominence in the fields of geomatics, remote sensing, and applied geotechnologies. These advancements have, for example, transformed traditional spatial interpolation models into tools with high levels of precision and adjustment in semivariogram reference models. Therefore, they are now presented as powerful groups of geostatistical tools in spatial autocorrelation analyses.

Furthermore, this review highlights the positive quantitative performance of the results and predictive yields by the ML methods and algorithms employed in the various explored studies. Primarily, the RF algorithms exhibited satisfactory values for the analyzed determination coefficients ($R^2 \approx >0.90$). Regarding the performance of different error tests, root mean square error, mean absolute error, mean absolute percentage error, RF, DNN, and in some specific cases, NN methods achieved better performances when compared with the results of other methods used in the studies. It is important to clarify that other ML methods such as support vector regression, extreme gradient boosting, MLP, and decision forest also obtained interesting results in terms of their predictive performances (though not all positive cases) depending on the specific interests measured or compared by the researchers.

Finally, this conclusion aims to address the hypothesis raised in this work. Indeed, this review has made it clear that geostatistical methods commonly and traditionally applied to the spatial modeling of solar radiation and terrestrial radiation budgets have recently been surpassed in applicability and functionalities by techniques based on ML and DP algorithms due to their lower error yields and higher levels of precision in some statistical geoprocesses, such as spatial interpolation at the digital levels of optical scenes recovered by remote sensors. However, this does not mean that geostatistical methods such as IDW, kriging, or NN are not currently used (as they continue to perform satisfactorily for some specific variables, such as temperature or precipitation at a general level). Rather, many recent works seem to have refined these geostatistical techniques and even considered partially or totally combining these techniques for working with complex and specific variables. This suggests very promising scenarios in which the science of spatial data will have more prominence in these discussions every day but also more challenges in a context largely influenced by the current climate variability and crisis.

**Funding:** The authors would like to thank the Ministerio de Ciencia e Innovación—Gobierno de España, for the funding provided in the Project RAILTWIN, number: PID2021-127089OB-I00. Project Reference: TDT-TSI-065100-2022-002.

**Data Availability Statement:** No new data were created or analyzed in this study.

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

## References

1. Kirkham, M.B. Solar radiation, black bodies, heat budget, and radiation balance. In *Principles of Soil and Plant Water Relations*; Academic Press: Cambridge, MA, USA, 2023; pp. 497–517. [CrossRef]
2. Fraga, E. *Concepts of Chemical Engineering for Chemists*, 2nd ed.; Simons, S., Ed.; The Royal Society of Chemistry: Cambridge, UK, 2016; pp. 1–21.
3. Simon, R.O.; Hülsbergen, K.-J. Energy Balance and Energy Use Efficiency of Annual Bioenergy Crops in Field Experiments in Southern Germany. *Agronomy* **2021**, *11*, 1835. [CrossRef]
4. Mauder, M.; Foken, T.; Cuxart, J. Surface-Energy-Balance Closure over Land: A Review. *Bound. Layer Meteorol.* **2020**, *177*, 395–426. [CrossRef]
5. Mateo, R.G.; Felicísimo, M.; Muñoz, J. Modelos de distribución de especies: Una revisión sintética. *Rev. Chil. Hist. Nat.* **2011**, *84*, 217–240. [CrossRef]
6. Valdés, J.; Londoño, L. *Análisis y Modelamiento Espacial*; Editorial Bonaventuriana: Medellín, Colombia, 2016; Volume 1.
7. Delgado-Villalobos, A.; Sibaja-Brenes, J.P.; Zúñiga-Castro, W.; Vega-Guzmán, I.; Alfaro-Solís, R.; Mora-Barrantes, J.C.; Borbón-Alpízar, H. Spatial distribution of hydrogen sulfite and ammonia emissions from a wastewater treatment plant in Costa Rica, using the AERMOD air pollution dispersion model. *Uniciencia* **2023**, *37*, 1–16. [CrossRef]
8. Tirado, J.; Gonzalez-Andujar, J. Spatial weed distribution models under climate change: A short review. *PeerJ* **2023**, *11*, e15220. [CrossRef]
9. Günal, E.; Budak, M.; Kılıç, M.; Cemek, B.; Sırrı, M. Combining spatial autocorrelation with artificial intelligence models to estimate spatial distribution and risks of heavy metal pollution in agricultural soils. *Environ. Monit. Assess.* **2023**, *195*, 317. [CrossRef]
10. Loghmari, I.; Timoumi, Y.; Messadi, A. Performance comparison of two global solar radiation models for spatial interpolation purposes. *Renew. Sustain. Energy Rev.* **2018**, *82*, 837–844. [CrossRef]
11. Jamaly, M.; Kleissl, J. Spatiotemporal Interpolation and Forecast of Irradiance Data Using Kriging. *Sol. Energy* **2017**, *158*, 407–423. [CrossRef]
12. Almarzooqi, A.M.; Maalouf, M.; El-Fouly, T.H.M.; Katzourakis, V.E.; El Moursi, M.S.; Chrysikopoulos, C.V. A hybrid machine-learning model for solar irradiance forecasting. *Clean Energy* **2024**, *8*, 100–110. [CrossRef]
13. Han, J.; Fang, S.; Mi, Q.; Wang, X.; Yu, Y.; Zhuo, W.; Peng, X. A time-continuous land surface temperature (LST) data fusion approach based on deep learning with microwave remote sensing and high-density ground truth observations. *Sci. Total Environ.* **2024**, *914*, 169992. [CrossRef]
14. Zhao, G.; Song, L.; Zhao, L.; Tao, S. Comparison of Different Machine Learning Methods to Reconstruct Daily Evapotranspiration Estimated by Thermal-Infrared Remote Sensing. *Remote. Sens.* **2024**, *16*, 509. [CrossRef]
15. Žalik, M.; Mongus, D.; Lukač, N. High-resolution spatiotemporal assessment of solar potential from remote sensing data using deep learning. *Renew. Energy* **2024**, *222*, 119868. [CrossRef]
16. Pelacani, S.; Maerker, M.; Tommasini, S.; Moretti, S. Combining biodiversity and geodiversity on landscape scale: A novel approach using rare earth elements and spatial distribution models in an agricultural Mediterranean landscape. *Ecol. Indic.* **2024**, *158*, 111583. [CrossRef]
17. Wagner, P.D.; Waske, B. Importance of spatially distributed hydrologic variables for land use change modeling. *Environ. Model. Softw.* **2016**, *83*, 245–254. [CrossRef]
18. Pennino, M.G.; Zurano, J.P.; Hidalgo, M.; Esteban, A.; Veloy, C.; Bellido, J.M.; Coll, M. Spatial patterns of β-diversity under cumulative pressures in the Western Mediterranean Sea. *Mar. Environ. Res.* **2024**, *195*, 106347. [CrossRef]
19. Valdivielso, S.; Hassanzadeh, A.; Vázquez-Suñé, E.; Custodio, E.; Criollo, R. Spatial distribution of meteorological factors controlling stable isotopes in precipitation in Northern Chile. *J. Hydrol.* **2022**, *605*, 127380. [CrossRef]
20. Wood, S.M.; Alston, L.; Beks, H.; Mc Namara, K.; Coffee, N.T.; Clark, R.A.; Shee, A.W.; Versace, V.L. Quality appraisal of spatial epidemiology and health geography research: A scoping review of systematic reviews. *Health Place* **2023**, *83*, 103108. [CrossRef]
21. Chen, C.; Dubovik, O.; Schuster, G.L.; Fuertes, D.; Meijer, Y.; Landgraf, J.; Karol, Y.; Li, Z. Characterization of temporal and spatial variability of aerosols from ground-based climatology: Towards evaluation of satellite mission requirements. *J. Quant. Spectrosc. Radiat. Transf.* **2021**, *268*, 107627. [CrossRef]
22. Gatto, A.; Mattera, R.; Panarello, D. For whom the bell tolls. A spatial analysis of the renewable energy transition determinants in Europe in light of the Russia-Ukraine war. *J. Environ. Manag.* **2024**, *352*, 119833. [CrossRef]
23. Olner, D.; Mitchell, G.; Heppenstall, A.; Pryce, G. The spatial economics of energy justice: Modelling the trade impacts of increased transport costs in a low carbon transition and the implications for UK regional inequality. *Energy Policy* **2020**, *140*, 111378. [CrossRef]

24. Reichelt, J.; Henning, V.; Foeckler, P. Mendeley Reference Manager, Versión 2.120.0 [Computer Software]. 2023. Available online: https://www.mendeley.com (accessed on 6 August 2024).
25. Du, Y.; Wang, T.; Zhou, Y.; Li, D.; Wang, S.; Xian, Y. Upscaling of longwave downward radiation from instantaneous to any temporal scale: Algorithms, validation, and comparison. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *117*, 103196. [CrossRef]
26. Leirvik, T.; Yuan, M. A Machine Learning Technique for Spatial Interpolation of Solar Radiation Observations. *Earth Space Sci.* **2021**, *8*, e2020EA001527. [CrossRef]
27. Pessanha, M.S.; dos Santos, L.M.; Lyra, G.B.; Lima, A.O.; Lyra, G.B.; de Souza, J.L. Interpolation methods applied to the spatialization of monthly solar irradiation in a region of complex terrain in the state of Rio de Janeiro in the southeast of Brazil. *Model. Earth Syst. Environ.* **2021**, *7*, 1565–1578. [CrossRef]
28. Cañada Torrecilla, M.R. El riesgo de contaminación por ozono en dos ciudades españolas (Madrid y Sevilla). Un estudio realizado con técnicas de modelado espacial y SIG. *Geographicalia* **2021**, *73*, 195–212. [CrossRef]
29. Velásquez, D.F.A.; Carrillo, G.A.A.; Barbosa, E.O.R.; Latorre, D.A.G.; Maldonado, F.E.M. Interpolacion Regnie Para Lluvia y Temperatura En Las Regiones Andina, Caribe y Pacífica de Colombia. *Colomb. For.* **2018**, *21*, 102–118. [CrossRef]
30. Berndt, C.; Haberlandt, U. Spatial interpolation of climate variables in Northern Germany—Influence of temporal resolution and network density. *J. Hydrol. Reg. Stud.* **2018**, *15*, 184–202. [CrossRef]
31. Jeong, D.; St-Hilaire, A.; Gratton, Y.; Bélanger, C.; Saad, C. A guideline to select an estimation model of daily global solar radiation between geostatistical interpolation and stochastic simulation approaches. *Renew. Energy* **2017**, *103*, 70–80. [CrossRef]
32. Rodríguez-Amigo, M.; Díez-Mediavilla, M.; González-Peña, D.; Pérez-Burgos, A.; Alonso-Tristán, C. Mathematical interpolation methods for spatial estimation of global horizontal irradiation in Castilla-León, Spain: A case study. *Sol. Energy* **2017**, *151*, 14–21. [CrossRef]
33. Mendez, M.; Calvo-Valverde, L. Assessing the Performance of Several Rainfall Interpolation Methods as Evaluated by a Conceptual Hydrological Model. *Procedia Eng.* **2016**, *154*, 1050–1057. [CrossRef]
34. Plouffe, C.C.; Robertson, C.; Chandrapala, L. Comparing interpolation techniques for monthly rainfall mapping using multiple evaluation criteria and auxiliary data sources: A case study of Sri Lanka. *Environ. Model. Softw.* **2015**, *67*, 57–71. [CrossRef]
35. Bachour, D.; Perez-Astudillo, D. Ground-measurement GHI Map for Qatar. *Energy Procedia* **2014**, *49*, 2297–2302. [CrossRef]
36. Garcia Cueto, R.; Santillán-Soto, N.; Haro, Z.; Bojorquez, G.; Ojeda-Benitez, S.; Quintero-Nuñez, M. *El Balance de Radiación y Modelos de Radiación Neta Para Diferentes Superficies: Estudio Experimental en Mexicali, México*; Spanish Association of Climate: Almería, Spain, 2014.
37. Turrado, C.C.; López, M.D.C.M.; Lasheras, F.S.; Gómez, B.A.R.; Rollé, J.L.C.; Juez, F.J.d.C. Missing Data Imputation of Solar Radiation Data under Different Atmospheric Conditions. *Sensors* **2014**, *14*, 20382–20399. [CrossRef] [PubMed]
38. Wagner, P.D.; Fiener, P.; Wilken, F.; Kumar, S.; Schneider, K. Comparison and evaluation of spatial interpolation schemes for daily rainfall in data scarce regions. J. Hydrol. 2012, 464–465, 388–400. [CrossRef]
39. Tait, A.; Liley, B. Interpolation of daily solar radiation for New Zealand using a satellite derived cloud cover surface. *Weather. Clim.* **2009**, *29*, 70–88. [CrossRef]
40. Şen, Z.; Şahīn, A.D. Spatial interpolation and estimation of solar irradiation by cumulative semivariograms. *Sol. Energy* **2001**, *71*, 11–21. [CrossRef]
41. Ustin, S.L.; Middleton, E.M. Current and near-term advances in Earth observation for ecological applications. *Ecol. Process.* **2021**, *10*, 1. [CrossRef]
42. Goudar, S.P.; Shah, S.S.; Shirali, G.S. ASTER measurement of supraglacial lakes in the Mount Everest region of the Himalaya. *Ann. Glaciol.* **2002**, *34*, 399–408. [CrossRef]
43. Stow, D.A.; Hope, A.; McGuire, D.; Verbyla, D.; Gamon, J.; Huemmrich, F.; Houston, S.; Racine, C.; Sturm, M.; Tape, K.; et al. Remote sensing of vegetation and land-cover change in Arctic Tundra Ecosystems. *Remote Sens. Environ.* **2004**, *89*, 281–308. [CrossRef]
44. Chang, L.; Sun, W. Consistency analysis of GRACE and GRACE-FO data in the study of global mean sea level change. *Geod. Geodyn.* **2022**, *13*, 321–326. [CrossRef]
45. Ulsig, L.; Nichol, C.J.; Huemmrich, K.F.; Landis, D.R.; Middleton, E.M.; Lyapustin, A.I.; Mammarella, I.; Levula, J.; Porcar-Castell, A. Detecting Inter-Annual Variations in the Phenology of Evergreen Conifers Using Long-Term MODIS Vegetation Index Time Series. *Remote Sens.* **2017**, *9*, 49. [CrossRef]
46. Bauwens, M.; Compernolle, S.; Stavrakou, T.; Müller, J.-F.; Van Gent, J.; Eskes, H.; Levelt, P.F.; Van Der, A.R.; Veefkind, J.P.; Vlietinck, J.; et al. Impact of Coronavirus Outbreak on $NO_2$ Pollution Assessed Using TROPOMI and OMI Observations. *Geophys. Res. Lett.* **2020**, *47*, e2020GL087978. [CrossRef]
47. Smith, G.; Priestley, K.; Loeb, N.; Wielicki, B.; Charlock, T.; Minnis, P.; Doelling, D.; Rutan, D. Clouds and Earth Radiant Energy System (CERES), a review: Past, present and future. *Adv. Space Res.* **2011**, *48*, 254–263. [CrossRef]
48. Galleguillos, M.; Jacob, F.; Prévot, L.; French, A.; Lagacherie, P. Comparison of two temperature differencing methods to estimate daily evapotranspiration over a Mediterranean vineyard watershed from ASTER data. *Remote Sens. Environ.* **2011**, *115*, 1326–1340. [CrossRef]
49. Parkkari, M.; Parviainen, M.; Ojanen, P.; Tolvanen, A. Spatial modelling provides a novel tool for estimating the landscape level distribution of greenhouse gas balances. *Ecol. Indic.* **2017**, *83*, 380–389. [CrossRef]

50. Laib, M.; Kanevski, M. Spatial Modelling of Extreme Wind Speed Distributions in Switzerland. *Energy Procedia* **2016**, *97*, 100–107. [CrossRef]
51. Yan, G.; Wang, T.; Jiao, Z.; Mu, X.; Zhao, J.; Chen, L. Topographic radiation modeling and spatial scaling of clear-sky land surface longwave radiation over rugged terrain. *Remote Sens. Environ.* **2016**, *172*, 15–27. [CrossRef]
52. Pan, T.; Wu, S.; Dai, E.; Liu, Y. Estimating the daily global solar radiation spatial distribution from diurnal temperature ranges over the Tibetan Plateau in China. *Appl. Energy* **2013**, *107*, 384–393. [CrossRef]
53. Schweitzer, A.; Andrea, C. Elaboración de un modelo de estimación de la distribución espacial de la radiación solar global mensual para Chile central. *Geography* **2007**, *1*, 10. Available online: https://api.semanticscholar.org/CorpusID:128639310 (accessed on 6 August 2024).
54. Alarcón, J.; Pabón Caicedo, J.D. El cambio climático y la distribución espacial de las formaciones vegetales en Colombia. *Colomb. For.* **2013**, *16*, 171–185.
55. Lary, D.J.; Alavi, A.H.; Gandomi, A.H.; Walker, A.L. Machine learning in geosciences and remote sensing. *Geosci. Front.* **2016**, *7*, 3–10. [CrossRef]
56. Edgar, T.W.; Manz, D.O. Machine Learning. In *Research Methods for Cyber Security*; Elsevier: Amsterdam, The Netherlands, 2017; pp. 153–173. [CrossRef]
57. Zhang, W.; Gu, X.; Tang, L.; Yin, Y.; Liu, D.; Zhang, Y. Application of machine learning, deep learning and optimization algorithms in geoengineering and geoscience: Comprehensive review and future challenge. *Gondwana Res.* **2022**, *109*, 1–17. [CrossRef]
58. Yeom, J.-M.; Park, S.; Chae, T.; Kim, J.-Y.; Lee, C.S. Spatial Assessment of Solar Radiation by Machine Learning and Deep Neural Network Models Using Data Provided by the COMS MI Geostationary Satellite: A Case Study in South Korea. *Sensors* **2019**, *19*, 2082. [CrossRef]
59. Koo, C.; Li, W.; Cha, S.H.; Zhang, S. A novel estimation approach for the solar radiation potential with its complex spatial pattern via machine-learning techniques. *Renew. Energy* **2019**, *133*, 575–592. [CrossRef]
60. Núñez-Reyes, A.; Ruiz-Moreno, S. Spatial Estimation of Solar Radiation Using Geostatistics and Machine Learning Techniques. *IFAC-PapersOnLine* **2020**, *53*, 3216–3222. [CrossRef]
61. Du, P.; Bai, X.; Tan, K.; Xue, Z.; Samat, A.; Xia, J.; Li, E.; Su, H.; Liu, W. Advances of Four Machine Learning Methods for Spatial Data Handling: A Review. *J. Geovisualiz. Spat. Anal.* **2020**, *4*, 13. [CrossRef]
62. Kartal, S.; Sekertekin, A. Prediction of MODIS land surface temperature using new hybrid models based on spatial interpolation techniques and deep learning models. *Environ. Sci. Pollut. Res.* **2022**, *29*, 67115–67134. [CrossRef] [PubMed]
63. Liao, X.; Zhu, R.; Wong, M.S. Simplified estimation modeling of land surface solar irradiation: A comparative study in Australia and China. *Sustain. Energy Technol. Assess.* **2022**, *52*, 102323. [CrossRef]
64. Sun, Y.; Jia, L.; Chen, Q.; Lin, X.; Sude, B.; Quan, Z.; Hutjes, R.W. Construction of a spatially gridded heat flux map based on airborne flux Measurements using remote sensing and machine learning methods. *Agric. For. Meteorol.* **2023**, *334*, 109424. [CrossRef]
65. Lu, Y.; Zhang, R.; Wang, L.; Su, X.; Zhang, M.; Li, H.; Li, S.; Zhou, J. Prediction of diffuse solar radiation by integrating radiative transfer model and machine-learning techniques. *Sci. Total Environ.* **2023**, *859*, 160269. [CrossRef] [PubMed]
66. Lu, Y.; Wang, L.; Zhu, C.; Zou, L.; Zhang, M.; Feng, L.; Cao, Q. Predicting surface solar radiation using a hybrid radiative Transfer–Machine learning model. *Renew. Sustain. Energy Rev.* **2023**, *173*, 113105. [CrossRef]
67. Stadler, C.; Fusé, V.S.; Linares, S.; Juliarena, P. Evaluation of interpolation methods for the prediction of urban methane concentrations. *Urban Clim.* **2023**, *49*, 101556. [CrossRef]
68. Chirinos, L.R.; Mallqui, C.G. Comparación entre Interpoladores Espaciales en el Estudio de Distribución de Partículas Sedimentables Insolubles en la Cuenca Atmosférica de Lima y Callao. *Inf. Tecnol.* **2016**, *27*, 111–120. [CrossRef]
69. Pereira, F.; Silva, C. Machine learning for monitoring and classification in inverters from solar photovoltaic energy plants. *Sol. Compass* **2024**, *9*, 100066. [CrossRef]
70. Li, F.; Yigitcanlar, T.; Nepal, M.; Nguyen, K.; Dur, F. Machine learning and remote sensing integration for leveraging urban sustainability: A review and framework. *Sustain. Cities Soc.* **2023**, *96*, 104653. [CrossRef]
71. Hrisko, J.; Ramamurthy, P.; Gonzalez, J.E. Estimating heat storage in urban areas using multispectral satellite data and machine learning. *Remote Sens. Environ.* **2021**, *252*, 112125. [CrossRef]
72. Wang, Y.; Khodadadzadeh, M.; Zurita-Milla, R. Spatial+: A new cross-validation method to evaluate geospatial machine learning models. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *121*, 103364. [CrossRef]
73. Sharma, N.; Saharia, M.; Ramana, G.V. High resolution landslide susceptibility mapping using ensemble machine learning and geospatial big data. *Catena* **2024**, *235*, 107653. [CrossRef]
74. Manivannan, R. Research on IoT-based hybrid electrical vehicles energy management systems using machine learning-based algorithm. *Sustain. Comput. Inform. Syst.* **2024**, *41*, 100943. [CrossRef]
75. Ajibade, S.-S.M.; Zaidi, A.; Bekun, F.V.; Adediran, A.O.; Bassey, M.A. A research landscape bibliometric analysis on climate change for last decades: Evidence from applications of machine learning. *Heliyon* **2023**, *9*, e20297. [CrossRef]